# Minimal Model of Self-Assembly: Emergence of Diversity and Complexity

Bogdan Barz[‡] and Brigita Urbanc*

Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, United States

**ABSTRACT:** Molecular self-assembly is ubiquitous in nature, yet prediction of assembly pathways from fundamental interparticle interactions has yet to be achieved. Here, we introduce a minimal self-assembly model with two attractive and two repulsive beads bound into a tetrahedron. The model is associated with a single parameter $\eta$ defined as the repulsive to attractive interaction ratio. We explore self-assembly pathways and resulting assembly morphologies for different $\eta$ values by discrete molecular dynamics. Our results demonstrate that $\eta$ governs the assembly dynamics and resulting assembly morphologies, revealing an unexpected diversity and complexity for $0.5 \leq \eta < 1$. One of the key processes that governs the assembly dynamics is assembly breakage, which emerges spontaneously at $\eta > 0$ with the breakage rate increasing with $\eta$. The observed assembly pathways display a broad variety of assembly structures characteristic of aggregation of amyloidogenic proteins, including quasi-spherical oligomers that coassemble into elongated protofibrils, followed by a conversion into ordered polymorphic fibril-like aggregates. We further demonstrate that $\eta$ can be meaningfully mapped onto amyloidogenic protein sequences, with the majority of amyloidogenic proteins characterized by $0.5 \leq \eta < 1$. Prion proteins, which are known to form highly breakage−prone fibrils, are characterized by $\eta > 1$, consistent with the model predictions. Our model thus provides a theoretical basis for understanding the universal aspects of aggregation pathways of amyloidogenic proteins relevant to human disease. As the model is not specific to proteins, these findings represent an important step toward understanding and predicting assembly dynamics of not only proteins but also viruses, colloids, and nanoparticles.

Self-Assembly Pathways
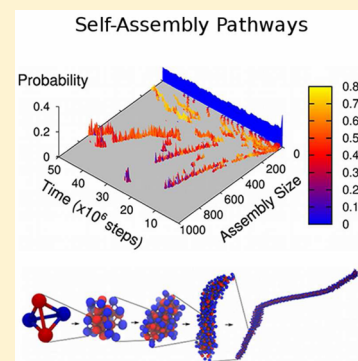
## ■ INTRODUCTION

Molecular self-assembly, a spontaneous association of disordered components into an ordered supramolecular structure, is responsible for the formation of complex biological systems and is becoming increasingly important in material sciences striving to develop novel biomaterials with a great diversity of biochemical and physical properties. Very little is known about the mechanisms underlying the emergence of ordered structures from disordered components. Recently, a remarkable diversity of polyhedra that self-assemble through excluded volume interactions at high packing fractions into entropic crystals with various degrees of crystalline order was reported.[1] At the opposite spectrum of a high packing fraction is protein aggregation, which typically occurs in vivo at nanomolar and in vitro at micromolar concentrations and must be consequently facilitated by attractive intermolecular interactions. Aberrant protein aggregation is at the core of many age-triggered diseases, such as Alzheimer's, Parkinson's, and Huntington's disease, amyotrophic lateral sclerosis, type II diabetes, systemic amyloidoses, and others.[2] These amyloid proteins do not share any obvious aspects of the primary structure yet they self-assemble into cytotoxic low-molecular weight oligomers and form fibrils with a common cross-$\beta$ structure.[3] Inherent toxicity of amyloid assemblies that cause the disease implies a common assembly mechanism;[4] however, assembly pathways are not well understood.[5] Here, we introduce a minimal model of self-assembly that unifies common features of protein amyloidogenesis and can be meaningfully mapped onto sequences of

amyloid proteins. Because the model is not protein specific, the model predictions extend to self-assembly systems beyond the aggregation of amyloidogenic proteins.

## ■ METHODS

**Model Construction.** A self-assembly model with a minimal number of beads and interparticle interactions is constructed as following. A one-bead molecule does not allow for the implementation of both attractive and repulsive interactions simultaneously. A two- or three-bead molecule is either anisotropic or planar, which imposes *a priori* geometric restrictions on self-assembling molecules. A three-dimensional molecule can be formed by a minimum of four beads, which we place at the four vertices of a tetrahedron (Figure 1A). Each tetrahedron molecule thus represents a molecule (protein monomer) comprising four beads each of a diameter $D$ connected by four covalent bonds of identical lengths $d$ (Figure 1A,B). Each bead within a tetrahedron molecule is assigned either an attractive (hydrophobic) or repulsive (hydrophilic) character, resulting in three possible model variants: (i) one hydrophilic and three hydrophobic beads, (ii) three hydrophilic and one hydrophobic bead, and (iii) two hydrophobic and two hydrophilic beads. Consistent with discrete molecular dynamics
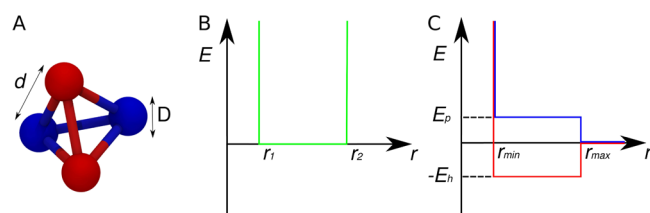
**Figure 1.** Definition of the tetrahedron model. (A) A tetrahedron molecule comprises two hydrophobic (red) and two hydrophilic (blue) beads of a diameter $D$ located at the vertices of a tetrahedron with the edge length $d$. (B) An "infinite" square-well *intra*molecular potential models covalent bonds among the four beads in the tetrahedron molecule. The bond length $d$ can vary between $r_1 = 1.56D$ and $r_2 = 1.61D$. (C) Effective attractive (red) and repulsive (blue) *inter*molecular potentials between pairs of hydrophobic and hydrophilic beads, respectively. The distance $r_{min} = D$ corresponds to a sum of the van Der Waals radii of the two interacting beads and $r_{max} = 2.22D$ is the interaction range distance. The strengths of the effective hydrophobic attractive and hydrophilic repulsive potential are denoted by $E_h$ and $E_p$, respectively.

(DMD) requirements (see below), the effective intermolecular hydrophobic attraction among hydrophobic beads and effective intermolecular hydrophilic repulsion among hydrophilic beads of different molecules are modeled by square-well potentials with a depth $E_h$ and a height $E_p$, respectively, (Figure 1C). A hydrophobic and a hydrophilic bead interact through an excluded volume only. The ratio of the two potential energies, $\eta = E_p/E_h$, is hereafter referred to as the hydropathy ratio. In variant (i), molecules aggregate into a single densely packed quasi-spherical assembly (data not shown). In variant (ii), molecules form stable dimers and trimers that do not assemble further (data now shown). In variant (iii) (Figure 1A), molecules exhibit complex assembly dynamics, which we examine for several values of the hydropathy ratio.

**Discrete Molecular Dynamics.** When interparticle potentials are approximated by a square-well or a combination of square-well potentials, molecular dynamics can be reduced to an efficient, event-driven discrete molecular dynamics (DMD). DMD was initially applied to simulate a system of hard spheres.[6,7] More recently, several low, intermediate, and high resolution protein models combined with DMD provided important insights into protein folding and assembly.[8–17]

Here, we apply DMD to examine pathways of self-assembly of the tetrahedron model. In each simulation, 1000 tetrahedron molecules occupy the simulation box with an edge length of $111D$, which corresponds to a volume fraction of $2.80 \pm 0.13 \times 10^{-3}$. Periodic boundary conditions are used and temperature is controlled by the Berendsen thermostat.[18] We examine self-assembly at five hydropathic ratios $\eta = E_p/E_h = 0, 0.5, 0.75, 1, 1.25$. For each $\eta$, 1000 tetrahedron molecules are placed into a cubic lattice in the simulation box, followed by high-temperature ($k_B T \gg E_h$) DMD simulations, which produce five distinct initial ensembles of randomly distributed non-interacting tetrahedron molecules. Thus obtained ensembles, five for each of the five $\eta$ values, resulting in 25 ensembles in total, are then used as initial configurations for the production runs. Here, a configuration is defined as an ensemble of coordinates of all tetrahedron molecules in a given trajectory at a given simulation time. Production runs are $10 \times 10^6$ ($\eta = 0$) or $50 \times 10^6$ ($\eta > 0$) simulation steps long, and configurations are recorded every $10^4$ simulation steps, resulting in $5 \times 1000 =$

5000 configurations ($\eta = 0$) or $5 \times 5000 = 25\,000$ configurations ($\eta > 0$).

The time scale of one simulation step, $\Delta t$, can be expressed in terms of the spatial resolution, $\Delta x \sim D \approx 10^{-9}$ m (thus approximating a protein by a tetrahedron with a circumsphere radius of $\sim 1$ nm), and thermal energy using the equipartition theorem,

$$\Delta t = \frac{\Delta x}{v} \approx D \sqrt{\frac{m_B}{k_B T}}$$

where we chose for the mass of one bead $m_B \sim 10^{-22}$ kg, equivalent to a sequence of $\sim 50$ amino acids or a $\sim 200$-residue protein corresponding to each tetrahedron molecule. We use the thermal energy at a physiological temperature 310 K, $k_B T \approx 4 \times 10^{-21}$ J, resulting in $\Delta t \approx 16 \times 10^{-9}$ s and the total simulation time per trajectory at nonzero $\eta$ values ($50 \times 10^6$ simulation steps) of $\sim 8 \times 10^{-3}$ s. Note that the unit of length, $D$, and the mass of each bead, $m_B$, can be adjusted to the size and molecular mass of the protein of interest, which consequently affects the time scale, $\Delta t \propto D(m_B)^{1/2}$, such that without considering the actual protein sequences, self-assembly of a larger protein occurs on a longer time scale than self-assembly of a smaller protein.

**Probability Distribution of Assembly Sizes.** The assembly size is defined as the number of molecules within a connected cluster. A recursive algorithm is used to identify tetrahedron molecules that belong to each cluster. A tetrahedron molecule is identified as a part of the cluster if any of its four beads is at a distance $\leq 2.22D$ from at least one of the beads of the molecules within the cluster. To derive three-dimensional plots displaying the time evolution of assembly sizes for each trajectory, assembly size distributions are calculated by using a binning interval of $5 \times 10^5$ simulation steps. The probability distribution of assembly sizes is calculated by using a $10^6$ simulation steps-wide window, resulting in 100 ensembles, each comprising 1000 tetrahedron molecules in various assembly states. Then, the average over the five probability distributions is taken, and the corresponding standard error of the mean (SEM) values are calculated. If an assembly size appears only once in a single trajectory, the corresponding SEM value is set to zero. By moving the $10^6$ simulation steps-wide window along the assembly trajectories between 0 and $50 \times 10^6$ simulation steps, time evolution of the size distribution for each $\eta$ is derived.

**Order Parameter.** To quantify the degree of orientational order within the assemblies of different sizes acquired for different $\eta$ values, we calculate the order parameter $S$ of each assembly. Each molecule in the assembly is characterized by an axis that connects its two hydrophobic (red) beads. The angle $\theta$ between the axis of each molecule and the long axis of the elongated assembly, determined by the principal component analysis,[19] is then used to calculate $S$ as the average:

$$S = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle \tag{1}$$

over all molecules in the assembly. Note that due to periodic boundary conditions, some assemblies are split in two or more clusters, which are separated by the linear size of the cubic simulation box in either of the three directions. To accurately calculate the order parameter, these split clusters are merged into a single cluster prior to the order parameter calculation. The order parameter calculation is performed for each assembly

of size larger than 5 within 5000 ($\eta = 0$) or 25 000 ($\eta > 0$) recorded configurations, each containing 1000 tetrahedron molecules. For each ensemble of assemblies of a size $\geq 6$, the average order parameter and SEM values are calculated by taking into account the order parameter values of all assemblies of a given size. For assemblies comprising 100 or more tetrahedron molecules, a running window along the assembly axis containing 50 molecules is used to calculate the local order parameter, followed by an average over all local order parameters.

## ■ RESULTS

Here a minimal model of self-assembling molecules, each characterized by a combination of attractive and repulsive interactions, is examined. Generally, any molecule prone to aggregation is characterized by attractive and repulsive interactions due to van der Waals and electrostatic interactions. We adopt in the following protein folding and aggregation terminology when discussing the origin of attractive and repulsive interactions. In proteins, attractive and repulsive interactions stem from the hydropathic (polar versus nonpolar) and electrostatic (charged versus uncharged) nature of individual residues. We posit that in addition to the effective hydrophobic attraction, the effective hydrophilic and/or electrostatic repulsion due to solvation of polar side chains by water molecules plays a critical role in prediction of protein assembly pathways and structures. We then construct a tetrahedron model with two hydrophobic and two hydrophilic beads as described in the Model Construction section (Methods), in which the hydropathy ratio $\eta = E_p/E_h$ between the repulsive $E_p$ and attractive $E_h$ potential energies is introduced as the only model parameter (Figure 1).

Self-assembly of the tetrahedron model is studied in the absence of repulsion among the hydrophilic beads ($\eta = 0$) and at four different hydropathy ratios ($\eta = 0.5, 0.75, 1,$ and $1.25$). To avoid a comparison of the assembly pathways of proteins with vastly different solubilities, the energy scale $E_h$ is adjusted for each $\eta$ to keep the final monomer concentration within the range of 10−16% (Table 1). For each $\eta$, five trajectories of self-assembly from initially monomeric ensembles of 1000 tetrahedron molecules are acquired and analyzed. For each trajectory, time evolution of assembly sizes is monitored and displayed as a three-dimensional plot (Figure 2), in which each assembly is also characterized by an order parameter as

**Table 1. Energy Unit and Steady-State Monomer, Dimer, and Trimer Concentrations**[a]

| $\eta = E_p/E_h$ | $E_h$ [$k_BT$] | [monomer] [%] | [dimer] [%] | [trimer] [%] |
|---|---|---|---|---|
| 0.00 | 0.97 | 14.027 ± 0.229 | 0.408 ± 0.028 | 0.014 ± 0.004 |
| 0.50 | 1.16 | 11.000 ± 0.383 | 0.272 ± 0.024 | 0.007 ± 0.003 |
| 0.75 | 1.27 | 10.737 ± 0.412 | 0.249 ± 0.016 | 0.007 ± 0.001 |
| 1.00 | 1.37 | 12.768 ± 0.268 | 0.430 ± 0.031 | 0.025 ± 0.005 |
| 1.25 | 1.45 | 15.687 ± 0.095 | 0.886 ± 0.016 | 0.052 ± 0.009 |

[a]The potential energy $E_h$, associated with the effective hydrophobic attraction, which represents a unit energy in our model is adjusted for each hydropathy ratio $\eta$ to yield comparable steady-state monomer concentration (solubility). For each $\eta$, steady-state monomer, dimer, and trimer concentrations and their SEM values are calculated by averaging over monomer, dimer, and trimer concentrations for each trajectory within 9−10 × 10⁶ steps ($\eta = 0$) or 49−50 × 10⁶ steps ($\eta > 0$), followed by averaging over the five trajectories.

described in the Order Parameter section (Methods), measuring a degree of the orientational order of molecules within the assembly (Figure 2). Below, we refer to small, quasi-spherical assemblies as oligomers to distinguish them from elongated, curvilinear assemblies (protofibrils) and more ordered assemblies (fibrils).

**Amorphous Assembly at $\eta = 0$.** Figure 3A shows time evolution of assembly sizes for four representative trajectories corresponding to $\eta = 0, 0.5, 1,$ and $1.25$, respectively. At $\eta = 0$, the initial hydrophobic collapse driving the assembly occurs rapidly, bringing the monomer concentration from the initial 100% to the final value of $\sim 14\%$ (Table 1). Within the first $10^6$ simulation steps (Figure 4A), initial monomers rapidly assemble into an aggregate comprising of 800−900 molecules, which remains in a steady-state equilibrium with monomers (Figure 3A, left) and small populations of dimers and trimers (Table 1). All assemblies have a relatively low order parameter as color-coded in Figure 3A (left). The assembly proceeds through formation of quasi-spherical oligomers, further elongating into larger curved tubules, which collapse onto themselves, exposing nonhydrophobic (blue) and shielding hydrophobic (red) beads from the "solvent", and eventually form a large quasi-spherical aggregate with an overall porous morphology (Figure 3A, right). The high bending propensity of these aggregates is reflected in their low order parameter for all five trajectories, which display almost identical assembly pathways (Figure 2).

**Transient Oligomers and Protofibril-like Assemblies at $\eta = 0.5$.** Simulations at $\eta = 0.5$ can be characterized by two assembly stages. The initial assembly stage lasts $\sim 10 - 20 \times 10^6$ simulation steps (Figure 3B, left), which is an order of magnitude longer than the hydrophobic collapse observed for $\eta = 0$. During this initial stage, smaller quasi-spherical oligomers coexist with larger elongated aggregates and a rapid elongation of large aggregates is mostly driven by oligomer addition. The assembly size distribution averaged over all five trajectories (Figure 2) reveals a large number of oligomers comprising 30−70 tetrahedron molecules forming within $\sim 10^6$ simulation steps (Figure 4B, black curve). These oligomers then coalesce into larger assemblies comprising 100−140 tetrahedron molecules at $1-2 \times 10^6$ simulation steps (Figure 4B, blue curve). After this initial assembly stage (at $\sim 20 \times 10^6$ simulation steps), oligomers disappear, marking the onset of the late assembly stage. In the late assembly stage, the elongated aggregates either increase or decrease in length at a rate distinctly lower than the aggregation rate of the initial assembly stage, while persisting in a steady-state equilibrium with the monomer population and minor populations of dimers and trimers (Table 1). Large aggregates at this late stage increase or decrease in size predominantly by monomer association and dissociation. In addition to this predominant growth dynamics, two additional albeit less frequent processes are noted: (a) merging of two elongated assemblies into a single larger elongated assembly and (b) breaking of an elongated aggregate into two smaller aggregates, which continue to either increase or decrease in length. The second process of aggregate breakage results in smaller assemblies of various sizes (Figure 4B, cyan, green, orange, and red curves). The morphology of the assemblies observed in simulation at $\eta = 0.5$ differs from those observed for $\eta = 0$. All assemblies at $\eta = 0.5$ have a considerably larger order parameter than those observed in simulations at $\eta = 0$. The order parameter of assemblies increases with their size (Figure 3B, left). Elongated aggregates are considerably less
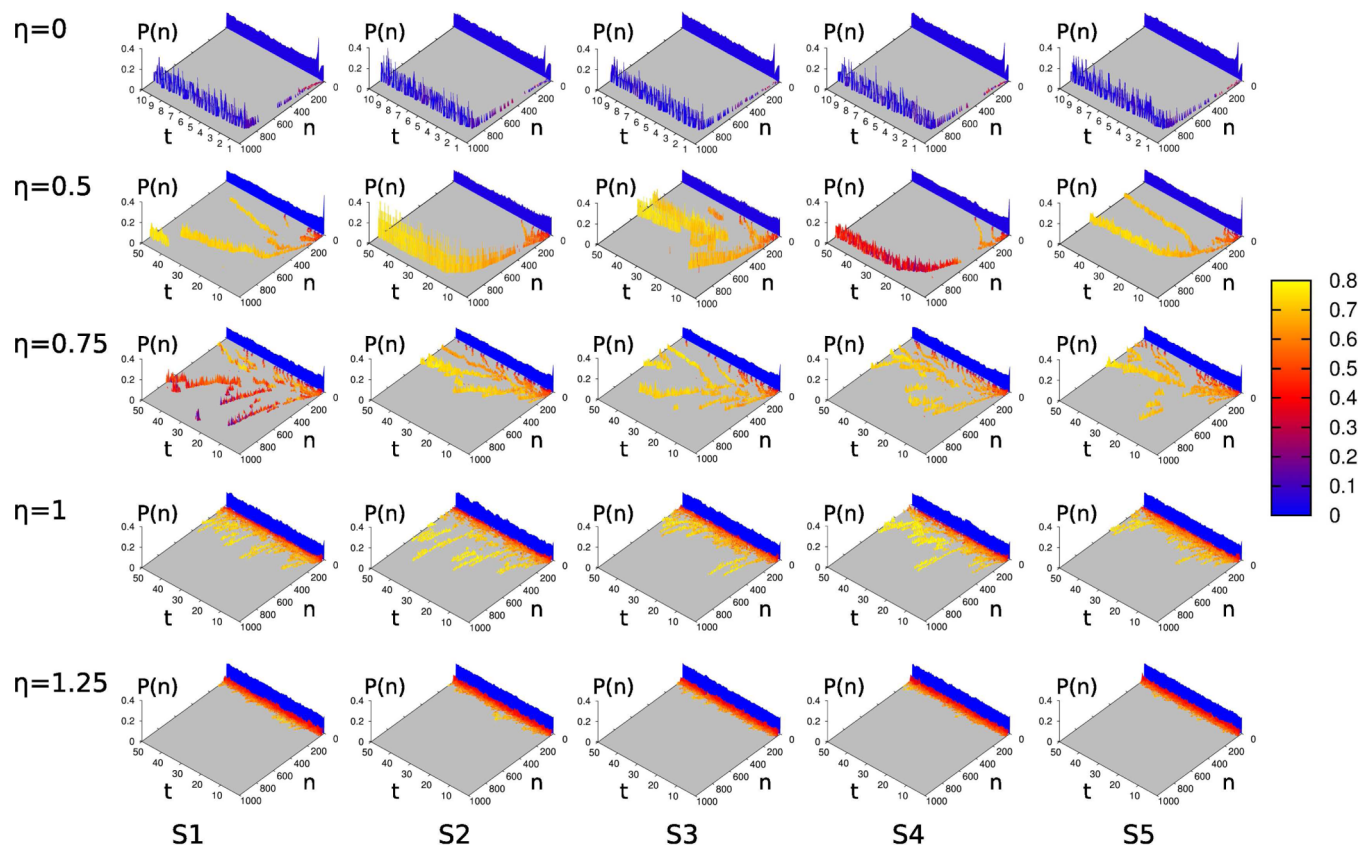
**Figure 2.** Assembly pathways of all trajectories. Three-dimensional plots showing probability distributions $P(n)$ of different assembly sizes $n$ as a function of the simulation time $t$ during the process of assembly starting from an ensemble of spatially separated (monomeric) tetrahedron molecules. Three−dimensional plots for each of the five trajectories, labeled as S1, S2, S3, S4, and S5, for each of the five $\eta$ values are displayed. The average order parameter $S$ that characterizes each assembly is color coded as shown on the color scale.

prone to bending than those observed for $\eta = 0$, but display kinks and can be slightly bent (Figure 3B, right). Unlike trajectories at $\eta = 0$, trajectories at $\eta = 0.5$ display significant variations (Figure 2).

**Breakage-Dominated Assembly at $\eta \geq 1$.** A two-stage assembly process observed for $\eta = 0.5$ is absent from simulations at $\eta = 1$. All trajectories acquired at $\eta = 1$ reach a steady state rapidly within $10^6$ simulation steps. The assembly pathways appear to be characterized by a single assembly growth rate (Figure 2). This rate is slightly lower than the one characterizing the initial assembly stage but notably larger than the elongation rate observed in the late assembly stage at $\eta = 0.5$. Importantly, the assembly dynamics at $\eta = 1$ is dominated by assembly breakage, which critically reduces the size of the largest aggregate. Consequently, the largest aggregate in the trajectory shown in Figure 3C (left) comprises ~400 tetrahedron molecules. Time evolution of the assembly size distributions reveals an abundance of quasi-spherical oligomers made of 20−30 molecules that form within the first $10^6$ simulation steps (Figure 4C, black curve). The corresponding peak in the assembly size distribution decreases and broadens at longer simulation times (Figure 4C, blue, cyan, green, orange, and red curves). Unlike in simulations at $\eta = 0.5$, where the oligomer assemblies disappear after the initial assembly stage, the assemblies of ~20−30 molecules persist at all simulation times, and give rise to assemblies of ~60 − 80 and ~100 − 120 tetrahedron molecules, respectively (Figure 4C, red curve). This multimodal character of the assembly size distribution indicates that larger assemblies form through coalescence of

smaller oligomers rather than through monomer addition. The steady−state pool of assemblies smaller than ~120 is a cause of a rapid growth of larger aggregates. A large breakage rate characteristic for $\eta = 1$ results in a broad distribution of assembly sizes up to ~550 (Figure 4C, green, orange, red curves). Similar to simulations at $\eta = 0$ and $\eta = 0.5$, molecules self-assemble initially into smaller quasi-spherical oligomers, followed by formation of elongated aggregates. The largest elongated aggregates show several kinks (Figure 3C, right). Some trajectory-to-trajectory variability is observed at $\eta = 1$, reflected in a variable size of the largest aggregate (Figure 2).

Simulations at $\eta = 1.25$ are characterized by a steady state dynamics, which is reached within $\leq 10^6$ simulation steps (Figure 3D, left). The rate of breakage is significantly higher than at $\eta = 1$, thus abolishing aggregates larger than ~120. Consequently, a steady-state assembly size distribution centered at the assembly size 25 with additional broader and less significant peaks centered at sizes ~50 and ~90, respectively, is observed (Figure 4D, red curve). The assembly proceeds through formation of smaller oligomers, followed by formation of elongated aggregates without any notable kinks or bending (Figure 3D, right). No significant trajectory-to-trajectory variability is observed (Figure 2).

**Complex and Diverse Assembly Pathways at $\eta = 0.75$.** The most variable assembly kinetics is observed for $0.5 \leq \eta \leq 1$, where the competition between the effective hydrophobic attraction and hydrophilic repulsion combined with thermal fluctuations results in a complex competing dynamics, as demonstrated by simulations at $\eta = 0.75$, which display the
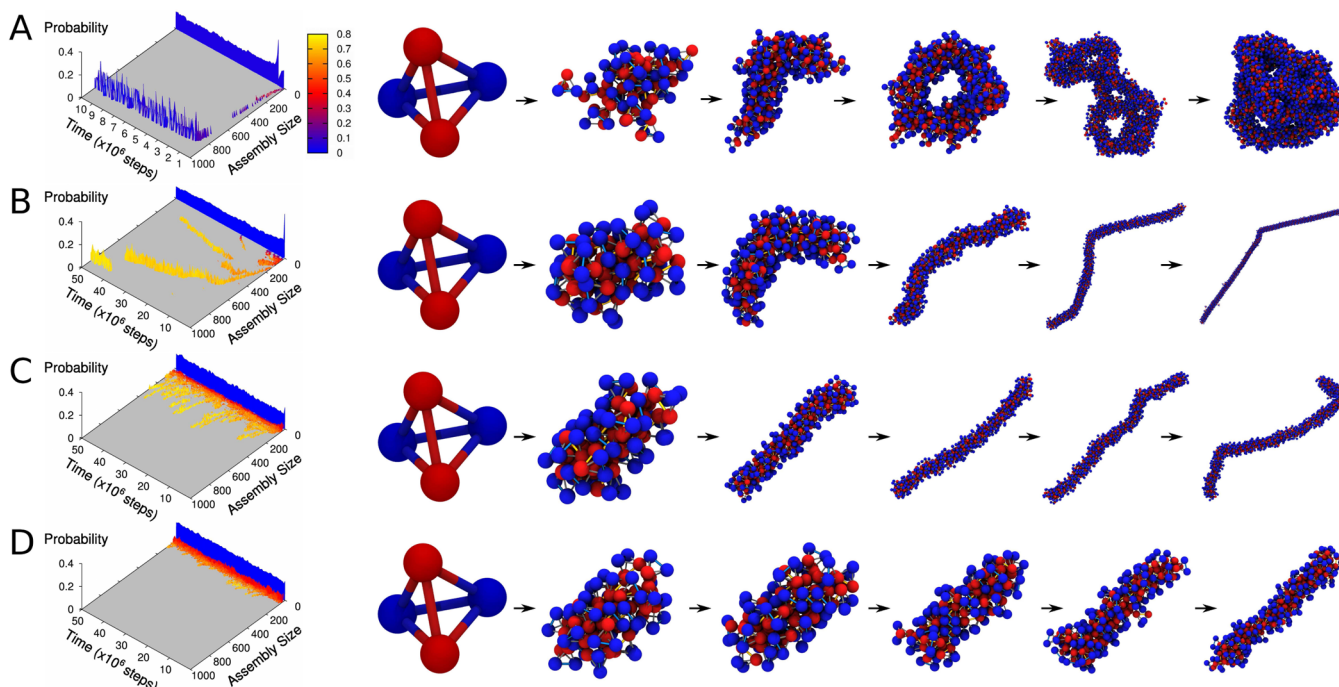
**Figure 3.** Time evolution of assembly pathways. Three-dimensional plots showing the assembly size distribution or probability (*z*-axis) of all assembly sizes (*x*-axis) at different times (*y*-axis) for representative trajectories (left) and their assembly pathways with characteristic morphologies (right) for simulations at (A) $\eta = 0$; (B) $\eta = 0.5$; (C) $\eta = 1$; and (D) $\eta = 1.25$. The color code in (A) is used to characterize the order parameter of each assembly in the three−dimensional plots. Note that the conformations are not displayed on the same scale: The final aggregates along assembly pathways for $\eta = 0$, 0.5 (A,B, right) are considerably larger than those for $\eta = 1$ and 1.25 (C,D, right), however, their lateral dimensions are comparable.
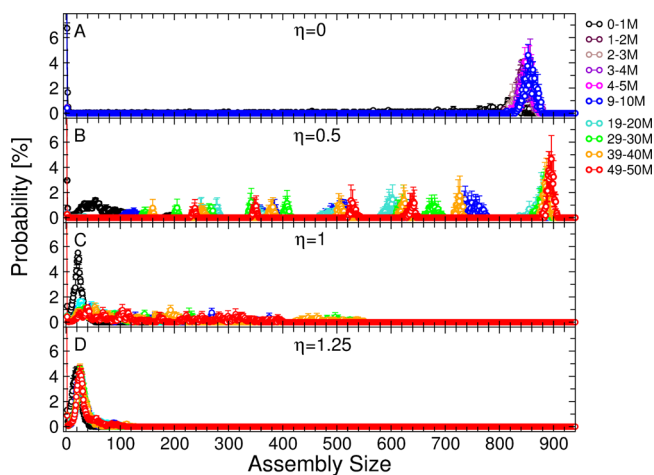


**Figure 4.** Time evolution of assembly size distributions. The average assembly size distributions at $\eta$ equal to (A) 0, (B) 0.5, (C) 1, and (D) 1.25 steadily evolve from the initial to the later assembly stages. Note that for $\eta = 0$, the size distribution reaches steady state within the initial $10^6$ simulation steps. The error bars correspond to SEM values.

largest trajectory-to-trajectory variability (Figure 2). Assembly dynamics shows characteristics of a two-stage aggregation process (observed for $\eta = 0.5$) as well as a relatively high breakage rate (observed for $\eta = 1$). Unique to simulations at $\eta = 0.75$ is a simultaneous occurrence of the fast and slow assembly dynamics, where one aggregate displays a fast growth through oligomer addition and the other aggregate displays a slow dynamics through monomer association and dissociation (Figure 5A, black double-sided arrow). Although the fast aggregation is more frequent earlier in the assembly process, it

is sporadically observed at later assembly stages as well, depending on the amount of oligomers available for coalescence with larger elongated aggregates (Figure 2). Figure 6 shows a merging process, during which the smaller assembly attaches itself to the end of the larger aggregate (C), temporarily forms a kink (D), and is subsequently integrated into the larger aggregate, eventually eliminating the kink (F).

In the early assembly stage, we observe oligomers comprising ~30 molecules (Figure 5B, top graph), which gradually decrease in abundance at later assembly stages at the expense of assemblies comprising up to 150 molecules as well as considerably larger aggregates comprising up to ~750 molecules (Figure 5B, bottom graph). This broad distribution of assembly sizes is caused by breakage, which is more frequent than for $\eta = 0.5$ but less frequent than for $\eta = 1$. Assembly into smaller quasi-spherical oligomers is followed by formation of elongated curvilinear protofibrils that eventually evolve into large multi-domain aggregates, separated by kinks (Figure 5C). Figure 5D shows four structurally distinct domains of a polymorphic aggregate that forms following the pathway described above. Domains I, III, and IV are cylindrically symmetric yet display variable degrees of lateral order from the most disordered (IV) to the most ordered (I), as shown in the corresponding profiles (Figure 5D, bottom). Domain II, which is not cylindrically symmetric, appears to be the most ordered of the four domains and displays a sandwich-like profile. Molecules in domain II are arranged into a double layer of 6−8 molecules in width. This highly structured domain emerges from a less ordered domain and spreads to neighboring regions in the course of the simulation (Figure 5E).

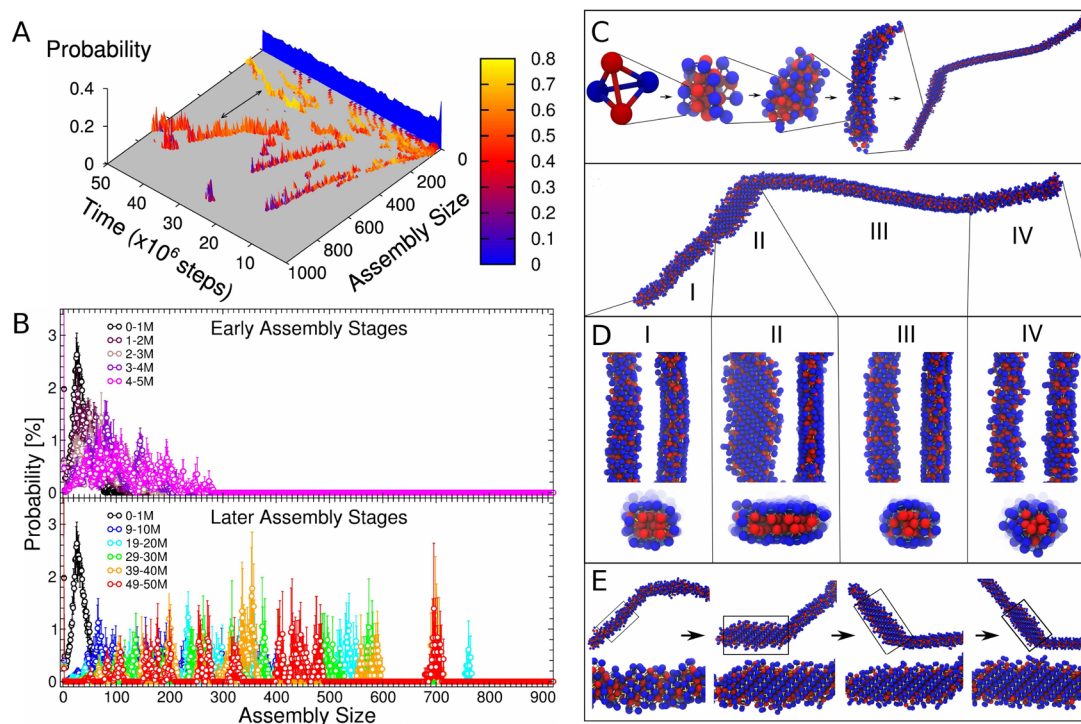**Morphology of the Assemblies Depends on the Hydropathy Ratio.** The assembly kinetics in our model is

**Figure 5.** Characterization of the assembly at $\eta = 0.75$. (A) A three-dimensional plot showing the assembly size distribution or probability ($z$-axis) of all assembly sizes ($x$-axis) at different times ($y$-axis) of the assembly process at $\eta = 0.75$. The color code on the right denotes the order parameter of each assembly. (B) Assembly size distributions at several early (top) and later (bottom) assembly stages are displayed. (C−E) The assembly pathway and corresponding morphologies observed in simulations at $\eta = 0.75$. (C) Initially, tetrahedron molecules form small quasi-spherical oligomers, which further assemble into elongated protofibril-like aggregates. Finally, a large elongated assembly with distinct structural domains (I−IV), including a sandwich-like ordered structure (II), emerges. (D) A closeup of the four distinct domains with the respective cross sections, displaying distinct degrees of lateral ordering. (E) Time evolution of the sandwich-like ordered structure (II).
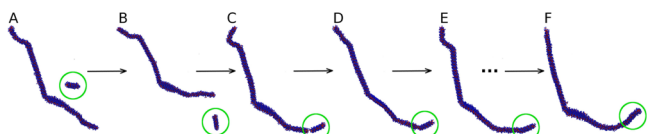


**Figure 6.** Time evolution of two merging assemblies for $\eta = 0.75$. (A−C) Large elongated multi-domain assembly and a small aggregate (inside a green circle) that are initially spatially separated, (D−F) merge into a single aggregate. The time frames correspond to (A) $42.33 \times 10^6$; (B) $42.34 \times 10^6$; (C) $42.35 \times 10^6$; (D) $42.36 \times 10^6$; (E) $42.38 \times 10^6$; and (F) $44.37 \times 10^6$ simulation steps. Note that within the $2 \times 10^6$ simulation steps between the last two time frames (E and F), the initially formed kink between the two merged assemblies smooths out as the smaller assembly integrates into the structure of the larger aggregate.

driven by a competition between spontaneous aggregation and disaggregation, controlled by the hydropathic ratio $\eta$, which affects the assembly size distribution, breakage rate, and assembly morphologies. The orientational order parameter of assemblies increases with the assembly size for each $\eta$ (Figure 7A). In addition, pentacosamer (25-mer) conformations are more ordered and less prone to bending as $\eta$ increases (Figure 7B).

**Mapping to Natively Unfolded Amyloid Proteins.** Complex and diverse self-assembly dynamics observed for $0.5 < \eta < 1$ results in multi-domain aggregates, consistent with in vitro observed molecular-level polymorphism of amyloid fibrils,[20] with bends and kinks (Figure 8A) that strongly resemble the morphology of in vitro amyloid fibers (Figure 8B). Can our minimal protein model thus provide insights into
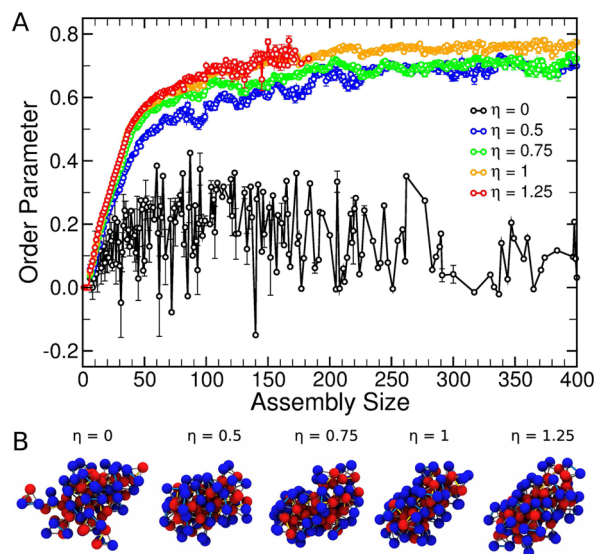


**Figure 7.** Order parameter and characteristic pentacosamer conformations. (A) The average order parameter as a function of the assembly size for the five different hydropathy ratios $\eta$. For each assembly size, the order parameter is calculated for a window containing a connected cluster of 50 tetrahedron molecules and then averaged over all windows along the assembly axis. The error bars correspond to SEM values. (B) Characteristic pentacosamer (25-mer) conformations at different values of $\eta$.

in vitro amyloid formation? Amyloid proteins assemble into cross-$\beta$ fibrils, stabilized by intermolecular hydrogen bonding
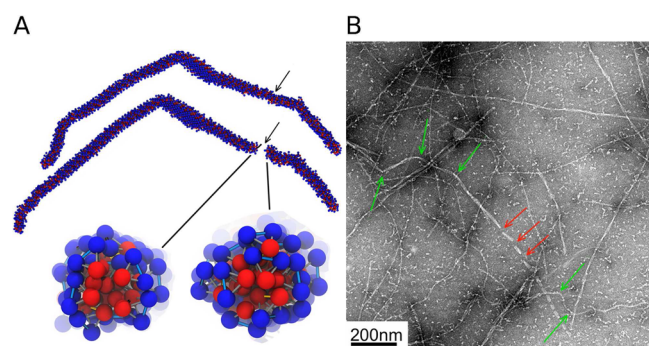
**Figure 8.** Comparison of in silico and in vitro morphologies. (A) A long in silico aggregate before (upper) and after (lower) the breakage occurring at the position marked by black arrows (DMD simulations at $\eta = 0.75$). Cross-sections to the left and to the right of the breakage with mismatched morphologies are displayed at the bottom. (B) Transmission electron micrograph of A$\beta$42 fibers characterized by several kinks (green arrows) and breakages (red arrows). The image is a courtesy of Drs. Louise C. Serpell, Julian Thorpe, and Thomas L. Williams.

along the fibril axis.[21−23] Yet, protein aggregation can occur in the absence of hydrogen bonding. In sickle cell anemia, sickle hemoglobin aggregates into polymers, causing a characteristic sickle-cell shape of red blood cells.[24] Whereas free energy calculations need to account for hydrogen bonding, our model reveals that a multi-stage assembly process that proceeds through formation of quasi-spherical and elongated mesoscopic-size aggregates, which eventually undergo a structural conversion into polymorphic aggregates, occurs also in the absence of hydrogen bonding.

The multi-stage assembly mechanism observed in our model is consistent with previously reported theoretical and experimental findings on amyloidogenic proteins. A recent theoretical study postulated both association and breakage processes to predict a variety of distinct aggregation pathways in agreement with experimental data.[25] Experimental studies on amyloid $\beta$-protein (A$\beta$), one of the most studied proteins due

to its relevance to Alzheimer's disease, which is the leading cause of dementia in elderly, showed that its two predominant alloforms, A$\beta$40 and A$\beta$42, form oligomers, which further assemble into protofibrils and finally convert into fibrils.[26,27] The existence of a nucleated conversion from prefibrillar into fibrillar assemblies was recently confirmed through FlAsH monitoring of Cys-Cys-A$\beta$ aggregation.[28] A structural conversion from within a molten collapsed intermediates into a structurally ordered fibril was also shown to be consistent with experimental data on yeast Sup35 prion protein.[29]

A large group of amyloidogenic proteins is natively unfolded.[30] A tetrahedron molecule in our model represents a monomer that retains its tertiary structure throughout the assembly process. As such, it can be identified as an aggregation-prone monomeric state of a natively unfolded protein. Whereas the internal degrees of freedom associated with intrinsically disordered nature of natively unfolded protein monomers are not incorporated in our model, the existence of such disorder implies that most amino acids in the sequence will be for some time exposed to the solvent, analogous to the four beads in the tetrahedron model. Consequently, it is reasonable to map the ratio of hydrophilic to hydrophobic amino acids, $N_p/N_h$, in the sequence of natively unfolded amyloidogenic proteins to the hydropathy ratio $\eta$ of our model. In this mapping, we implicitly assume that hydrophobic collapse dominates protein self-assembly and that hydrogen bonds, which are strongly directional, form among peptide regions only after these regions have been driven into proximity by effective hydrophobic interactions. This mapping also does not distinguish between charged and noncharged hydrophilic amino acids and thus all charged hydrophilic amino acids effectively contribute to repulsive interactions. In aqueous solutions, electrostatic interactions among charged amino acids depend on pH and ionic concentrations that may under specific conditions modify the overall repulsive to attractive interaction ratio. However, most protein sequences are dominated by noncharged amino acids and salt bridges between oppositely charged amino acids tend to be exposed to the solvent and rarely stabilize protein structures in water.[31,32] Thus, attractive

**Table 2. Natively Unfolded Amyloidogenic Proteins and Their $N_p/N_h$ Values**[a]

| protein | ID | length | Chothia[33] | Janin[34] | Kyte-Doolittle[35] | Eisenberg[36] | Engleman[37] | average |
|---|---|---|---|---|---|---|---|---|
| ABri[45] | − | 34 | 0.60 | 0.93 | 1.00 | 1.08 | 1.00 | 0.92 |
| ADan[46] | − | 34 | 0.56 | 0.75 | 0.88 | 0.86 | 0.88 | 0.79 |
| $\alpha$-Synuclein | 1XX8[43] | 140 | 0.68 | 0.96 | 0.98 | 0.71 | 0.98 | 0.86 |
| Amylin | 2KB8[43] | 37 | 0.71 | 0.64 | 0.71 | 0.64 | 0.71 | 0.68 |
| A$\beta$(1−40) | 1BA4[43] | 40 | 0.53 | 0.65 | 0.82 | 0.48 | 0.82 | 0.66 |
| A$\beta$(1−42) | 1IYT[43] | 42 | 0.47 | 0.58 | 0.74 | 0.44 | 0.74 | 0.59 |
| Apolipoprotein-A1 | P02647[44] | 242 | 1.03 | 1.26 | 1.38 | 1.12 | 1.32 | 1.22 |
| Calcitonin | 2GLH[43] | 33 | 1.00 | 1.00 | 1.12 | 0.89 | 1.12 | 1.03 |
| Human Prion | 1QLX[43] | 210 | 1.71 | 1.35 | 1.86 | 0.70 | 1.57 | 1.44 |
| Huntingtin | 3IO4[43] | 449 | 0.88 | 0.90 | 0.97 | 0.76 | 0.92 | 0.89 |
| Tau[47] | − | 441 | 1.13 | 1.33 | 1.44 | 0.93 | 1.44 | 1.25 |
| Transthyretin | 1BM7[43] | 127 | 0.58 | 0.71 | 0.84 | 0.59 | 0.80 | 0.70 |
| Yeast-Sup35 | P05453[44] | 685 | 1.24 | 1.32 | 1.41 | 1.04 | 1.38 | 1.28 |
| | | | | | | | | **0.86 ± 0.36** |

[a]The ratio of hydrophilic to hydrophobic amino acids ($N_p/N_h$) for each protein is calculated within each of the five distinct hydropathy scales. Proteins are identified by their names and/or ID codes from Protein Data Bank[43] or Universal Protein Resource.[44] For the three proteins without the ID codes, the source references are cited next to the protein name in the first column. The length of proteins is given in terms of the number of amino acids in the sequence (third column). The averages over the five $N_p/N_h$ values corresponding to each of the five hydropathy scales for each protein are shown in the last column. The average and standard deviation of the average $N_p/N_h$ values per protein of the last column are displayed at the bottom right (bold font).

3767

dx.doi.org/10.1021/jp412819j | J. Phys. Chem. B 2014, 118, 3761−3770

electrostatic interactions are mostly not expected to significantly contribute to inter-residue interactions during protein self-assembly.

By employing several hydropathy scales previously reported by Chothia,[33] Janin,[34] Kyte and Doolittle,[35] Eisenberg et al.,[36] and Engelman et al.,[37] we identified the hydrophilic and hydrophobic amino acids within each scale, counted the number of hydrophobic and hydrophilic amino acids in the sequence, and calculated $N_p/N_h$ for several natively unfolded proteins, known to form amyloid fibrils (Table 2). A histogram of the $N_p/N_h$ values from Table 2 demonstrates that the $N_p/N_h$ values are distributed in the range $0.5 < N_p/N_h < 2$ (Figure 9). The highest $N_p/N_h$ values are found for the human prion, prion-like yeast-Sup35, and tau (Table 2).
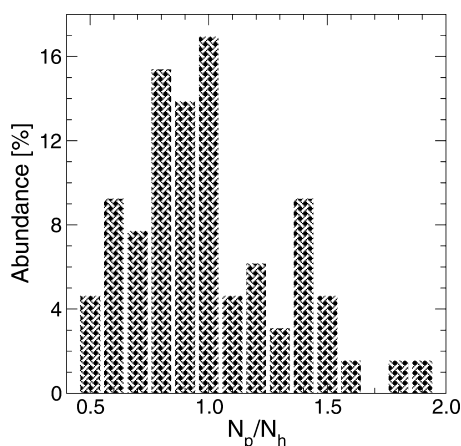


**Figure 9.** Histogram of the hydropathic ratios for natively unfolded amyloid proteins. Histogram of the ratio of hydrophilic to hydrophobic number of residues, $N_p/N_h$, in the sequence of natively unfolded amyloidogenic proteins.

High values of $\eta > 1$ in our model result in a high breakage rate and relatively short and ordered aggregates. Prion proteins are known to form brittle aggregates, prone to breakage, which is associated with their infectious nature.[38] Substantial evidence indicates that tau protein displays prion-like characteristics by initially aggregating in a few nerve cells in discrete brain areas and then self—propagating and spreading to distant brain regions.[39] Most $N_p/N_h$ values from Table 2 fall into the range $0.5 < N_p/N_h < 1$, for which our model predicts complex and diverse assembly dynamics with a broad assembly size distribution. These proteins include A$\beta$40 and A$\beta$42 associated with Alzheimer's disease, amylin associated with diabetes mellitus type 2, transthyretin associated with systemic amyloidosis, $\alpha$-synuclein associated with Parkinson's disease, and huntingtin associated with Huntington's disease. Interestingly, A$\beta$42, which aggregates faster and forms larger assemblies than A$\beta$40 under equivalent in vitro conditions, has a lower $N_p/N_h$ value of 0.59 than A$\beta$40 with $N_p/N_h = 0.66$, consistent with our model's maximal assembly size decreasing with $\eta$ in the range $0.5 < \eta < 1$. For $\eta \geq 0.75$, assembly breakage in our model creates a steady-state pool of oligomers, which form from smaller and merge into larger assemblies. Assuming that oligomers are the key toxic species responsible for pathogenesis, even if lifetimes of oligomers are short, the assembly dynamics that produces a constant pool of oligomers is a persistent source of toxicity, which provides a plausible mechanism through which a metastable assembly with a short lifetime might exert toxicity over a longer time period.

## ■ CONCLUSIONS AND DISCUSSION

In summary, our tetrahedron model predicts complex self-assembly into quasi-spherical oliogmers, curvilinear protofibrils, and multi-domain aggregates, characteristic of in vitro observed amyloid formation by natively unfolded amyloidogenic proteins. Spontaneous assembly breakage and merging, assembly size distributions, and a diversity of assembly morphologies in our model are controlled by a single parameter, the hydropathic ratio $\eta$, which can be meaningfully mapped onto the hydrophilic to hydrophobic ratio of the sequence of amyloidogenic proteins. We show that most natively unfolded amyloidogenic proteins are characterized by $0.5 < \eta < 1$, for which our model predicts the most complex and diverse assembly dynamics.

There is no consensus on the size and structure of toxic oligomers. In Alzheimer's disease, different A$\beta$ pre-fibrillar species were reported in vitro and in vivo, depending on the methodologies and experimental settings.[40] Moreover, relatively minor changes in the sequence between A$\beta$40 and A$\beta$42 result in distinct assembly pathways[27] and oligomer structures[10] that likely underlie distinct cytotoxic properties of the two peptides, with A$\beta$42 that is genetically more strongly associated with Alzheimer's disease than A$\beta$40. Similarly, a single amino acid mutation of A$\beta$ can cause earlier onset of the disease and/or altered pathology.[41] A seeming contradiction between two sides of the protein aggregation puzzle: (a) a large class of proteins with variable sequences that aggregate into amyloid fibrils and (b) small changes in the sequence and/or solvent conditions that significantly alter the assembly pathways and structures, can be reconciled within our model by noting that both sequence and solvent modifications alter an overall hydropathic nature of the protein, thus $\eta$, in the range, where the assembly dynamics is the most sensitive.

Short hydrophobic peptides incubated with A$\beta$42 were recently reported to inhibit A$\beta$42-induced toxicity in cell cultures[42] and subsequent computer simulations of A$\beta$42 assembly in the presence of these inhibitors elucidated the structure of the resulting large amorphous hetero-oligomers.[16] By binding hydrophobic peptides to A$\beta$42, the overall hydropathic ratio of hetero-assemblies, $\eta$, is lowered to $\eta \ll 0.5$, thereby biasing aggregation toward presumably less toxic amorphous aggregates. The hydropathy parameter might thus be critical for a deeper understanding of protein aggregation dynamics and developing drugs that aim to alter aggregation pathways in a way to minimize the damage caused by cytotoxic oligomeric assemblies. While the discussion described above is specific to protein aggregation, our model and its predictions are applicable to a wide range of self—assembling systems and elucidate a critical role of the repulsive to attractive intermolecular interaction ratio in supramolecular structure prediction.

## ■ AUTHOR INFORMATION

**Corresponding Author**
*E-mail: brigita@drexel.edu. Phone: (215) 895-2726. Fax: (215) 895-5934.

Article

## Present Address

‡(B.B.) Forschungszentrum Jülich GmbH Institute of Complex Systems: Structural Biochemistry (ICS6), 52425 Jülich, Germany.

## Notes

The authors declare no competing financial interest.

## ■ REFERENCES

(1) Damasceno, P.; Engel, M.; Glotzer, S. Predictive Self-Assembly of PAlyda into Ccomplex Structures. *Science* **2012**, *337*, 453−457.

(2) Walsh, D. M.; Selkoe, D. J. Oligomers on the brain: The emerging role of soluble protein aggregates in neurodegeneration. *Protein Pept Lett* **2004**, *11*, 213−228.

(3) Dobson, C. M. Protein folding and misfolding. *Nature* **2003**, *426*, 884−890.

(4) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **2002**, *416*, 507−511.

(5) Schnabel, J. Amyloid: Little proteins, big clues. *Nature* **2011**, *475*, S12−S14.

(6) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. I. General method. *J. Chem. Phys.* **1959**, *31*, 459−466.

(7) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. II. Behaviour of a small number of elastic spheres. *J. Chem. Phys.* **1960**, *33*, 1439−1451.

(8) Zhou, Y.; Karplus, M. Folding thermodynamics of a model three-helix-bundle protein. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 14429−14432.

(9) Ding, F.; Borreguero, J. M.; Buldyrev, S. V.; Stanley, H. E.; Dokholyan, N. V. Mechanism for the $\alpha$-Helix to $\beta$-Hairpin Transition. *Proteins: Structure, Function, and Genetics* **2003**, *53*, 220−228.

(10) Urbanc, B.; Cruz, L.; Yun, S.; Buldyrev, S. V.; Bitan, G.; Teplow, D. B.; Stanley, H. E. *In silico* study of amyloid $\beta$-protein folding and oligomerization. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17345−17350.

(11) Nguyen, H.; Hall, C. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16180−16185.

(12) Borreguero, J. M.; Urbanc, B.; Lazo, N. D.; Buldyrev, S. V.; Teplow, D. B.; Stanley, H. E. Folding events in the 21−30 region of amyloid $\beta$-protein (A$\beta$) studied *in silico*. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6015−6020.

(13) Nguyen, H.; Hall, C. Kinetics of fibril formation by polyalanine peptides. *J. Biol. Chem.* **2005**, *280*, 9074−9082.

(14) Ding, F.; Tsao, D.; Nie, H.; Dokholyan, N. *Ab initio* folding of proteins with all-atom discrete molecular dynamics. *Structure* **2008**, *16*, 1010−1018.

(15) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. Elucidation of Amyloid $\beta$-Protein Oligomerization Mechanisms: Discrete Molecular Dynamics Study. *J. Am. Chem. Soc.* **2010**, *132*, 4266−4280.

(16) Urbanc, B.; Betnel, M.; Cruz, L.; Li, H.; Fradinger, E.; Monien, B. H.; Bitan, G. Structural Basis of A$\beta_{1-42}$ Toxicity Inhibition by A$\beta$ C-Terminal Fragments: Discrete Molecular Dynamics Study. *J. Mol. Biol.* **2011**, *410*, 316−328.

(17) Meral, D.; Urbanc, B. Discrete Molecular Dynamics Study of Oligomer Formation by N-Terminally Truncated Amyloid $\beta$-Protein. *J. Mol. Biol.* **2013**, *425*, 2260−2275.

(18) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *8*, 3684−3690.

(19) R Development Core Team, *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008, http://www.R-project.org.

(20) Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W.-M.; Mattson, M. P.; Tycko, R. Self-Propagating, Molecular-Level Polymorphism in Alzheimer's $\beta$-Amyloid Fibrils. *Science* **2005**, *307*, 262−265.

(21) Serpell, L. C. Alzheimer's amyloid fibrils: Structure and assembly. *Biochim. Biophys. Acta* **2000**, *1502*, 16−30.

(22) Nelson, R.; Eisenberg, D. Recent atomic models of amyloid fibril structure. *Curr. Opin. Struct. Biol.* **2006**, *16*, 260−265.

(23) Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **2011**, *62*, 279−299.

(24) Ferrone, F. A. Crowding and the polymerization of sickle hemoglobin. *Methods. Enzymol.* **2006**, *12*, 285−299.

(25) Schreck, J.; Yuan, J. A kinetic study of amyloid formation: Fibril growth and length distributions. *J. Phys. Chem. B* **2013**, *117*, 6574−6583.

(26) Bitan, G.; Lomakin, A.; Teplow, D. B. Amyloid $\beta$-Protein Oligomerization: Prenucleation Interactions Revealed by Photo-Induced Cross-linking of Unmodified Proteins. *J. Biol. Chem.* **2001**, *276*, 35176−35184.

(27) Bitan, G.; Kirkitadze, M. D.; Lomakin, A.; Vollers, S. S.; Benedek, G. B.; Teplow, D. B. Amyloid $\beta$-protein (A$\beta$) assembly: A$\beta$40 and A$\beta$42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 330−335.

(28) Lee, J.; Culyba, E.; Powers, E.; Kelly, J. Amyloid-$\beta$ Forms Fibrils by Nucleated Conformational Conversion of Oligomers. *Nat. Chem. Biol.* **2011**, *7*, 602−609.

(29) Krishnan, R.; Lindquist, S. Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* **2005**, *435*, 765−772.

(30) Chiti, F.; Dobson, C. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333−366.

(31) Finkelstein, A. V.; Ptitsyn, O. B. *Protein Physics: A Course of Lectures*; Academic Press (An imprint of Elsevier Science): London, 2002.

(32) Straub, J. E.; Thirumalai, D. Principles governing oligomer formation in amyloidogenic peptides. *Curr. Opin. Struct. Biol.* **2010**, *20*, 187−195.

(33) Chothia, C. The Nature of the Accessible and Buried Surfaces in Proteins. *J. Mol. Biol.* **1976**, *105*, 1−14.

(34) Janin, J. Surface and inside volumes of globular proteins. *Nature* **1997**, *277*, 491−492.

(35) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157*, 105−132.

(36) Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of Membrane and Surface Protein Sequences with the Hydrophobic Moment Plot. *J. Mol. Biol.* **1984**, *179*, 125−142.

(37) Engelman, D. M.; Steitz, T. A.; Goldman, A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* **1986**, *15*, 321−353.

(38) Tanaka, M.; Collins, S.; Toyama, B.; Weissman, J. The physical basis of how prion conformations determine strain phenotypes. *Nature* **2006**, *442*, 585−589.

(39) Spillantini, M.; Goedert, M. Tau pathology and neurodegeneration. *Lancet Neurol.* **2013**, *12*, 609−622.

(40) Roychaudhuri, R.; Yang, M.; Hoshi, M. M.; Teplow, D. B. Amyloid $\beta$-protein assembly and Alzheimer disease. *J. Biol. Chem.* **2008**, *284*, 4749−53.

(41) Attar, A.; Meral, D.; Urbanc, B.; Bitan, G. Assembly of Amyloid $\beta$-Protein Variants Containing Familial Alzheimer's Disease-Linked Amino Acid Substitutions. In *Bio-nanoimaging: Protein Misfolding and Aggregation*; Uversky, V., Lyubchenko, Y., Eds.; Elsevier: London, 2014; Chapter 38, pp 429−442.

(42) Fradinger, E.; Monien, B. H.; Urbanc, B.; Lomakin, A.; Tan, M.; Li, H.; Spring, S. M.; Condron, M. M.; Cruz, L.; Xie, C.-W.; Benedek, G. B.; Bitan, G. C-terminal peptides coassemble into A$\beta$42 oligomers

and protect neurons against Aβ42-induced neurotoxicity. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 14175−14180.

(43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242, http://www.rcsb.org/pdb.

(44) The UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2014**, *42*, D191−D198, http://www.uniprot.org.

(45) Vidal, R.; Frangione, B.; Rostagno, A.; Mead, S.; Révész, T.; Plant, G.; Ghiso, J. A stop-codon mutation in the BRI gene associated with familial British dementia. *Nature* **1999**, *399*, 776−781.

(46) Vidal, R.; Révész, T.; Rostagno, A.; Kim, E.; Holton, J.; Bek, T.; Bojsen-Møller, M.; Braendgaard, H.; Plant, G.; Ghiso, J.; Frangione, B. A decamer duplication in the 3′ region of the BRI gene originates an amyloid peptide that is associated with dementia in a Danish kindred. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 4920−4925.

(47) von Bergen, M.; Barghorn, S.; Biernat, J.; Mandelkow, E.; Mandelkow, E. Tau aggregation is driven by a transition from random coil to beta sheet structure. *Biochim. Biophys. Acta* **2005**, *1739*, 158−166.

3770

dx.doi.org/10.1021/jp412819j | *J. Phys. Chem. B* 2014, 118, 3761−3770