Elucidation of amyloid $\beta$-protein oligomerization mechanisms: Discrete molecular dynamics study

B. Urbanc$^1$, M. Betnel$^2$, L. Cruz$^1$, G. Bitan$^{3,4}$, and D. B. Teplow$^{3,4}$

$^1$Department of Physics, Drexel University, Philadelphia, PA 19104, USA; $^2$Department of Physics, Boston University, Boston, MA 02215; $^3$Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095-7334, USA; $^4$Molecular Biology Institute and Brain Research Institute, University of California, Los Angeles, CA 90095-7334, USA.

Abstract

Oligomers of amyloid $\beta$-protein (A$\beta$) play a central role in the pathology of Alzheimer’s disease. Of the two predominant A$\beta$ alloforms, A$\beta_{1-40}$ and A$\beta_{1-42}$, A$\beta_{1-42}$ is more strongly implicated in the disease. We elucidated structural characteristics of oligomers of A$\beta_{1-40}$ and A$\beta_{1-42}$ and their Arctic mutants, [E22G]A$\beta_{1-40}$ and [E22G]A$\beta_{1-42}$. We simulated oligomer formation using discrete molecular dynamics (DMD) with a four-bead protein model, backbone hydrogen bonding, and residue–specific interactions due to effective hydropathy and charge. For all four peptides under study, we derived the characteristic oligomer size distributions that were in agreement with prior experimental findings. Unlike A$\beta_{1-40}$, A$\beta_{1-42}$ had a high propensity to form paranuclei (pentameric or hexameric) structures that could self-associate into higher-order oligomers. Neither of the Arctic mutants formed higher-order oligomers but [E22G]A$\beta_{1-40}$ formed paranuclei with a similar propensity as A$\beta_{1-42}$. While the best agreement with the experimental data was obtained when the charged residues

$^*$corresponding author; e-mail: brigita@drexel.edu
were modeled as solely hydrophilic, further assembly from spherical oligomers into elongated protofibrils was induced by non-zero electrostatic interactions among the charged residues. Structural analysis revealed that the C-terminal region played a dominant role in $\text{A}\beta_{1-42}$ oligomer formation, while $\text{A}\beta_{1-40}$ oligomerization was primarily driven by intermolecular interactions among the central hydrophobic regions. The N-terminal region A2–F4 played a prominent role in $\text{A}\beta_{1-40}$ oligomerization but did not contribute to oligomerization of $\text{A}\beta_{1-42}$ or the Arctic mutants. The oligomer structure of both Arctic peptides resembled more $\text{A}\beta_{1-42}$ than $\text{A}\beta_{1-40}$, consistent with their potentially more toxic nature.
Introduction

Alzheimer’s disease (AD), characterized by irreversible, progressive deterioration of learning and memory, is the dominant cause of dementia in the elderly. One of the hallmarks of AD is an accumulation of extracellular senile plaques, which contain fibrillar aggregates of the amyloid β-protein (Aβ). Genetic and experimental evidence strongly supports the hypothesis that low-order\(^1\) oligomeric assemblies of Aβ, rather than fibrils, are the proximate neurotoxic agents in AD.\(^2\)–\(^4\) Aβ is produced through cleavage of the amyloid precursor protein (APP) and is normally present in the body predominantly in two alloforms, Aβ\(_{1-40}\) and Aβ\(_{1-42}\), that differ structurally by the absence or presence of the two C-terminal amino acids Ile41–Ala42, respectively. Despite this relatively small difference in primary structure, Aβ\(_{1-42}\) oligomers are more toxic.\(^5\) Proper targeting of therapeutic agents requires elucidation of Aβ assembly dynamics and the determination of the site(s) responsible for imparting on Aβ\(_{1-42}\) its particular toxic potential. Different Aβ oligomeric assemblies may induce neurotoxicity by distinct mechanisms,\(^6\) thus it is important to examine oligomeric structure of specific order and under specific conditions when determining structure–toxicity correlations. A variety of oligomers have been reported to be toxic,\(^4\) but none have been characterized structurally at the atomic level.

Aβ folding and assembly are remarkably complex processes, which complicates the application of classical structure determination methods such as X-ray crystallography and solution state NMR to the oligomerization question.\(^7\) A powerful approach providing information on the initial Aβ oligomerization process has been chemical cross-linking. The method of photo-induced cross-linking of unmodified proteins (PICUP), combined with SDS-PAGE, was used by Bitan et al. to determine the Aβ\(_{1-40}\) and Aβ\(_{1-42}\) oligomer size frequency distributions, revealing that the two alloforms form oligomers through distinct pathways.\(^8\)–\(^10\) Aβ\(_{1-42}\) assembled into pentamer/hexamer units (paranuclei\(^11\)) and multiples of paranuclei (dodecamers and octadecamers), while Aβ\(_{1-40}\) only formed dimers, trimers, and tetramers that were in equilibrium with monomers. Recently, ion mobility spectrometry was applied to determine Aβ\(_{1-40}\) and Aβ\(_{1-42}\) oligomer distributions in vitro without the use of chemical cross-linking.\(^12\) Bern-
stein et al. found distinct $\text{A}\beta_{1-40}$ and $\text{A}\beta_{1-42}$ oligomer distributions,\textsuperscript{12} in agreement with PICUP/SDS-PAGE findings.\textsuperscript{9} Oligomer size distributions have been found to be sensitive to single amino acid substitutions,\textsuperscript{13} including those causing familial forms of AD and cerebral amyloid angiopathy. As such, the oligomer size distribution can be considered a “fingerprint” of a particular peptide and may be correlated with structural properties of the associated oligomers and protofibrils. However, neither PICUP/SDS-PAGE nor ion mobility spectrometry, by themselves, can reveal directly the interatomic interactions controlling the oligomerization process.

Molecular dynamics (MD) provides the means to visualize and quantify three-dimensional atomic structures of proteins on time scales and in detail not possible experimentally. The number of computational approaches aimed at understanding $\text{A}\beta$ folding and assembly has increased substantially over the past decade (see Ref.\textsuperscript{14} for a review). All-atom MD, which utilizes complete atomic models of $\text{A}\beta$, remains computationally untenable because of the large system size and simulation duration necessary to study $\text{A}\beta$ oligomerization.\textsuperscript{7} For this reason, discrete molecular dynamics (DMD) combined with simplified protein models (one or two beads per residue) were explored to study protein folding.\textsuperscript{15–21} However, these models require the native protein structure to be incorporated into interparticle interactions, which is undesirable for studies of natively disordered proteins such as $\text{A}\beta$. Intermediate-resolution models, such as the four-bead and united-atom models with backbone hydrogen bonding, which only require the knowledge of the protein sequence, have been found to yield promising results when combined with DMD\textsuperscript{22–26} or Monte Carlo dynamics.\textsuperscript{27,28}

Recently, the DMD approach with either the four-bead protein model or a more sophisticated united-atom protein model (for a review, see Ref.\textsuperscript{29}) was applied to the $\text{A}\beta$ system. The results agreed with existing in vitro findings or were amenable to in vitro confirmation.\textsuperscript{30–34} The DMD approach, which comprised the four-bead model, backbone hydrogen bonding, and the empirical hydropathy scale derived by Kyte and Doolittle,\textsuperscript{35} demonstrated that the key differences in oligomerization between $\text{A}\beta_{1-40}$ and $\text{A}\beta_{1-42}$ resulted from the hydrophobic nature of the amino acids I41 and A42 present in $\text{A}\beta_{1-42}$,\textsuperscript{31} in agreement with PICUP data.\textsuperscript{13} When the effective electrostatic interaction among the charged amino acids was included in addi-
tion to the backbone hydrogen bonding and effective hydropathy, significantly larger oligomer sizes in both Aβ1−40 and Aβ1−42 were observed. This is a particularly intriguing finding as a more recent work on Aβ1−40 and Aβ1−42 folding showed that the electrostatic interaction did not affect folded structures in a significant way. The oligomerization differences between the two alloforms were preceded by differences in the folded structures. Specifically, Aβ1−42, but not Aβ1−40, was characterized by a turn structure centered at G37–G38, a prediction consistent with findings of several independent experimental studies. Aβ1−42 and Aβ1−40 folded structures differed also at the N-terminal region, where Aβ1−40, but not Aβ1−42, showed an increased β-strand propensity at A2–F4. The DMD approach was applied to examine temperature-induced changes in the secondary structure of the folded Aβ1−40 and Aβ1−42 as well as to characterize folding differences induced by the Arctic mutation, E22G, which is associated with a familial form of AD characterized by an increased propensity for protofibril formation. Interestingly, the E22G substitution induced structural differences at the N-terminal region, resulting in the folded structure of both the Aβ1−40 and Aβ1−42 Arctic mutants showing a greater resemblance to the wild-type (WT) Aβ1−42.

In this work, we applied the *ab initio* DMD approach with two implicit solvent parameters to examine the early events in assembly pathways of Aβ1−40, Aβ1−42, [E22G]Aβ1−40, and [E22G] Aβ1−42. Adjusting the two solvent parameters to replicate experimental solvent conditions is important because Aβ folding and assembly strongly depend on the solvent. Earlier applications of the DMD approach to Aβ folding and oligomer formation and subsequent comparison of the results to *in vitro* PICUP data on oligomer size distributions and to the average amounts of β-strand in monomeric and oligomeric states showed a consistent agreement between computational and experimental results for the same narrow range of the two implicit solvent parameters. The present work builds upon the assumptions that the oligomer size distributions are correctly represented by experimental technique using PICUP/SDS-PAGE and that our DMD approach captures the key features relevant to folding, early assembly pathways, and the resulting structures of the four Aβ peptides under study. Here, relative to our earlier study, we simulated Aβ1−40 and Aβ1−42 oligomerization for twice as many time steps (20×10⁶), further extended the main production runs to 40×10⁶, and
we incorporated a more accurate estimate of the physiological temperature, changes that allowed us to obtain a closer agreement with experimentally derived oligomer size distributions.\textsuperscript{9} Using these changed simulation parameters, we then examined the role of weak electrostatic interactions between charged amino acids on the formation of both quasispherical oligomers and the more elongated protofibril-like assemblies into which they convert. Finally, we applied the DMD simulations with the same implicit solvent parameters and followed the same simulation protocol as for the WT peptides to examine the effects of the Arctic mutation on the oligomerization pathways and structural properties of the resulting oligomers.

**Methods**

The advantages and limitations of the DMD approach with the four-bead protein model, backbone hydrogen bonding, and amino acid-specific interactions modeling the solvent implicitly were described in detail in Ref.\textsuperscript{29} Below we briefly summarize key characteristics of this computational approach.

**Discrete Molecular Dynamics**

MD can be reduced to discrete molecular dynamics (DMD) whenever the interparticle potentials are approximated by a single or a combination of square wells.\textsuperscript{45} DMD is event-driven and each event corresponds to a collision between a pair of particles. Particles move along straight lines with constant speeds between two succeeding collisions. Instead of integrating Newton’s equations at every time step for every particle to obtain a complete trajectory, DMD only requires an efficient sorting of the collision times between pairs of particles and the selection of the shortest one, which determines the time of the next event/collision. Consequently, DMD is orders of magnitude faster than MD. In our DMD approach,\textsuperscript{29} the simulation volume and number of particles are fixed, the temperature is held constant by the Berendsen thermostat,\textsuperscript{46} and periodic boundary conditions are implemented.

In contrast to all-atom MD in explicit solvent, in which an explicit time step (e.g., 1 ps) ensures a direct comparison between simulation and experimental time scales, event-driven
DMD with implicit solvent lacks such an “absolute” time scale, because the collisions between the protein atoms and solvent and those among solvent atoms are not explicitly modeled. Consequently, the relationship between the DMD simulation and real time may not be a simple linear function. Thus, a direct comparison of DMD–derived kinetics of protein folding and assembly to experimentally obtained data is not straightforward. However, DMD is a dynamics in a true sense, in that it results in a causally related sequence of events. In our study, the sequence of events includes all stages of folding from unstructured random coil-like monomers, followed by assembly into oligomers, and potentially, larger aggregates.

The Four-Bead Protein Model with Backbone Hydrogen Bonding

In the four-bead protein model, the backbone is represented by three beads, corresponding to the amide (N), the α-carbon (Cα), and the carbonyl (C) groups. Each side-chain, except glycine, is represented by one side-chain bead (Cβ). Four beads per amino acid is the smallest number of beads that can capture the chiral nature of an individual amino acid, i.e., a lack of mirror symmetry in all amino acids except glycine. The lengths of bonds and the angular constraints are determined phenomenologically by calculating their distributions using the known folded protein structures of ~7,700 proteins from the Protein Data Bank (PDB). The values of the lengths of covalent bonds and the angular constraints are allowed to fluctuate around their average values with the fluctuation amplitude set to 2% of the average value. The effective backbone hydrogen bond is implemented between the nitrogen atom $N_i$ of the $i$–th amino acid and the carbon atom $C_j$ of the $j$–th amino acid. For the hydrogen bond to form, the neighboring backbone atoms need to be at correct distances from the atoms involved in hydrogen bond formation to account for the anisotropic character of the hydrogen bond. The strength of the backbone hydrogen bond interaction, $E_{HB}$, represents a unit of energy and the temperature is given in units $E_{HB}/k_B$. 


Amino Acid-Specific Interactions Due to Hydropathy and Charge

To account implicitly for the aqueous solvent, we previously implemented the effective amino acid-specific hydropathic interactions between the $C_\beta$ atoms\textsuperscript{31} using the hydropathy scale derived by Kyte and Doolittle.\textsuperscript{35} At neutral pH, the amino acids I, V, L, F, C, M, and A were considered hydrophobic, N, Q, and H were considered non-charged hydrophilic, and R, K, D, and E were considered charged hydrophilic. The remaining amino acids with absolute values of hydropathies below threshold values were treated as neutral. Only $C_\beta-C_\beta$ interactions within the group of hydrophobic residues and within the group of hydrophilic residues were allowed. The rest of $C_\beta-C_\beta$ interactions were only due to the excluded volume ("hard sphere") repulsion. Each $C_\beta$ atom was associated with a hydropathy parameter normalized between $-1$ (for the most hydrophobic residue) to $+1$ (for the most hydrophilic residue). The effective hydropathic interaction was modeled by a single square-well potential with the absolute value of the potential energy (or the strength of the interaction) equal to $E_{HP}$ and the interaction range of $0.75$ nm. When the $C_\beta-C_\beta$ distance of two hydrophobic/hydrophilic side-chain atoms decreased from above to below $0.75$ nm (but remained larger than the sum of their hard sphere radii), they experienced an attractive/repulsive interaction, equal to $E_{HP}$ multiplied by an average of their respective hydropathy parameters. The implicit solvent parameter $E_{HP}$, which strongly depends on the presence or absence of water in the solvent, was set to $E_{HP} = 0.3$ (in units of $E_{HB}$) as in the previous study.\textsuperscript{31} The interaction between two hydrophobic side chains was attractive, to model the tendency of each hydrophobic residue to decrease its solvent exposure. The interaction between two hydrophilic side chains was repulsive, to model the tendency of each hydrophilic residue to maximize its solvent exposure. The implementation of the effective electrostatic interaction between charged residues is given below. Note that in the absence ($E_{CH} = 0$) of the effective electrostatic interaction between the charged hydrophilic residues, two $C_\beta$ atoms of the same or of the opposite charge experience effective hydrophilic repulsion. This state corresponds to solvent conditions in which all charged groups are completely screened by solvent water molecules, disabling electrostatic interactions among them.

The effective electrostatic interaction between two charged side-chain atoms was imple-
mented using a double attractive/repulsive square well potential and the interaction range of 0.75 nm. The absolute value of the potential energy associated with the electrostatic interaction was $E_{CH}$ (in units of $E_{HB}$), which was experienced between two charged $C_\beta$ atoms at a distance < 0.6 nm (the soft interaction range). Here we used very small strengths of the effective electrostatic interactions, $E_{CH} = 10^{-6}$, or $E_{CH} = 10^{-2}$ (in units of $E_{HB}$). In contrast to the case of $E_{CH} = 0$, in which the charged amino acids R, K, D, and E were considered purely hydrophilic, when $E_{CH} = 10^{-6}$ or $E_{CH} = 10^{-2}$ the charged amino acids interacted among themselves through their electrostatic but not their hydrophilic properties, even though the corresponding strengths of the electrostatic interaction were small compared to $E_{HB}$. Note that even with $E_{CH} > 0$, there was an effective hydrophilic repulsion between two non-charged hydrophilic $C_\beta$ atoms as well as between a charged and a non-charged hydrophilic $C_\beta$ atom. The implicit solvent parameter values $E_{CH} = 10^{-6}$ and $E_{CH} = 10^{-2}$ would correspond to the solvent conditions, at which the charged groups were only partially screened by solvent water.

**Simulation Conditions**

The primary structure of $A\beta_{1-42}$ is:

\begin{align*}
1\text{DAEFRHDSGY}^{11}\text{EVHHQKLFF} & \quad 21\text{AEDVGSNKGA}^{31}\text{IIGLMVGGVV}^{41}\text{IA}.
\end{align*}

The shorter alloform, $A\beta_{1-40}$, lacks the C-terminal amino acids, I$_{41}$A$_{42}$. Following the previous notation,\textsuperscript{31,33,34} we will refer to the region L17–A21 as the central hydrophobic cluster (CHC), the region I31–V36 as the mid-hydrophobic region (MHR), and the region V39–V40/A42 as the C-terminal region (CTR).

Previously, we demonstrated that the four-bead DMD approach was successful in capturing oligomer size distribution differences between $A\beta_{1-40}$ and $A\beta_{1-42}$. $A\beta_{1-40}$ produced an oligomer size distribution with a high dimer frequency and decreasing amounts of higher-order oligomers, whereas $A\beta_{1-42}$ displayed a pentamer peak, in addition to a dimer peak.\textsuperscript{31} Although these simulation results were consistent with experimentally observed oligomer frequency distributions,\textsuperscript{9} they did not yield significant numbers of $A\beta_{1-42}$ hexamers nor oligomers of orders
that were found experimentally. To address this issue, we have extended the simu-
lation times and employed a more physiologic temperature, $T = 0.13$ (in units of $E_{HB}/k_b$) versus 0.15 used previously. Longer simulation times were employed because oligomeric states are metastable and are hypothesized to eventually evolve through protofibrils to fibril states. $T = 0.13$ was chosen because recent simulations of Aβ1–40 and Aβ1–42 folding \(^{34}\) have shown that when $T \in (0.12, 0.13)$, the population average β-strand ($\langle \beta\text{-strand} \rangle$) content most accurately represents that found experimentally.\(^{44}\)

We simulated oligomer formation starting from 32 spatially separated non-interacting and unfolded Aβ peptides in a cubic box of 25 nm. Initial distinct populations of conformers for 8 trajectories of each alloform were obtained by performing short high-temperature ($T = 4$) DMD simulation runs and saving configurations every $0.1 \times 10^6$ simulation steps. Eight production runs each for Aβ1–40 and Aβ1–42 then were performed for $20 \times 10^6$ and for the optimal implicit solvent parameters further extended to $40 \times 10^6$ simulation steps.

**Structural Analysis**

Quasispherical oligomers formed by Aβ1–40 and Aβ1–42, from dimers through pentamers, were structurally characterized in the previous study.\(^{31}\) The structures of Aβ1–40 and Aβ1–42 trimers and tetramers showed structural characteristics that overlapped partially with those of dimers and pentamers, suggesting a monotonic structural evolution from dimers through pentamers.

In the present study, eight trajectories of 32 peptides each were produced for the four peptides under study. In the oligomer size distribution calculation, we considered realizations at time frames of 19, 19.5, and $20 \times 10^6$ simulation steps to facilitate the comparison of the current and previous work.\(^{31}\) For the rest of the analyses, realizations at eleven time frames between 19 and $20 \times 10^6$ simulation steps (at $10^5$-step intervals) were considered. Each realization consisted of a mixture of monomers, as well as oligomers of various order. A criterion used to determine whether two peptide chains, A and B, belonged to the same oligomer, was based on the threshold distance of 0.75 nm between the centers of mass of an atom of the chain
A and an atom of the chain B. If any pair of such atoms was at a distance smaller than 0.75 nm, the chains A and B were assigned to the same oligomer.

Prior to the analysis, PDB\textsuperscript{47} files of individual monomers and oligomers were extracted from all realizations and sorted by oligomer order. Oligomers of a particular order that were most abundantly represented in the probability distribution of oligomer sizes were, by definition, those with the lowest free energies. In our structural analysis, we thus selected dimers and hexamers, which were relatively abundant in all peptides. We did so because the corresponding oligomer structures were representative, which was needed to highlight the differences among the four peptides, and the number of actual oligomer conformations\textsuperscript{48} was large enough to produce statistically reliable results. In addition, dimer formation is the first step in oligomerization, thus it is important to compare structurally dimers of the four peptides under study and hexamers, which are hypothesized to be among the most toxic A\textsubscript{B1-42} oligomers. In total, across 8 different trajectories and 11 time frames per peptide, there were 396 A\textsubscript{B1-40}, 215 A\textsubscript{B1-42}, 266 [E22G]A\textsubscript{B1-40}, and 129 [E22G]A\textsubscript{B1-42} dimers, and 24 A\textsubscript{B1-40}, 59 A\textsubscript{B1-42}, 69 [E22G]A\textsubscript{B1-40}, and 43 [E22G]A\textsubscript{B1-42} hexamers, yielding a set of conformers large enough to power statistical analyses of the structural data (see also Table 2).

**Oligomer Size Distribution**

Because oligomer size distributions evolved with time, we selected for the calculation of probability distributions a relatively narrow time window of $1 \times 10^6$ steps, within which the distributions did not change significantly. Within this window, we selected three time frames such that the corresponding populations of monomers and oligomers were as independent of each other as possible, to avoid biasing the averaging and standard error calculations. Thus, for each peptide, 24 distinct realizations from 8 trajectories (each comprising 32 peptides) at time frames of 19, 19.5, and $20 \times 10^6$ steps were analyzed. For each of these 24 realizations per peptide, the probability distributions of monomers and oligomers of all sizes were calculated. The final normalized probability distributions of monomers and oligomers of all orders and the standard errors of the mean (SEM) were calculated as averages over 24 individual probability distributions. All oligomer size distributions reported in this work were calculated using
the same protocol. To address the concern of 24 not being a sufficient number of frames, we also recalculated all of the size distributions by considering, instead of 3 time frames (at 19, 19.5, and 20 million time steps), 11 time frames (from 19 to 20 million time steps, taking into account conformations 100,000 steps apart)—in total 88 frames (instead of 24). The resulting size distributions were smoother, but essentially overlapping within the SEM values, with the original distributions (data not shown).

**Secondary Structure**

Amino acid-specific propensities for the secondary structure formation were calculated using the STRIDE program\textsuperscript{49,50} within the Visual Molecular Dynamics (VMD) software package.\textsuperscript{51} The secondary structure propensities included several types of $\alpha$-helical, $\beta$-strand, turn, and bridge structures. Here we analyzed in detail the dominant secondary structure elements, $\beta$-strand and turn.

**Solvent Accessible Surface Area (SASA)**

We calculated the solvent accessible surface area per amino acid (SASA) by using VMD.\textsuperscript{51} In addition to the four atoms per amino acid (three for glycine), we also considered the backbone carbonyl oxygen and amide hydrogen involved in hydrogen bond formation. This calculation created a spherical surface around each atom in the amino acid, 1.4 Å away from the atom’s van der Waals surface, resulting in a combined three-dimensional surface around each amino acid. SASA was then defined by calculating the part of this surface area that did not overlap with any other surface belonging to the neighboring amino acids and was thus be accessible to the implicit solvent. Amino acids that were buried inside an assembly had a smaller SASA, whereas amino acids on the surface assumed a higher SASA.

**Contact Maps with Hydrogen Bond Propensities**

By definition, two amino acids were deemed to be in contact if the distance between their centers of mass was $\leq 0.75$ nm. The contact map was defined as a matrix in which the value of
each \((i,j)\) element was equal to an average number of contacts between amino acids \(i\) and \(j\). We distinguished intra- and intermolecular contact maps. If amino acids \(i\) and \(j\) belonged to the same peptide, the corresponding contacts contributed to the intramolecular contact map. If not, the contacts contributed to the intermolecular contact map. The contact map of each assembly was normalized by the number of contributing peptide molecules and by the maximal number of contacts between two residues. Because of the four-bead representation, the maximal number of intramolecular contacts was \(4^2 = 16\). Consequently, the average number of intramolecular contacts between residues \(i\) and \(j\) normalized by the factor 16 could not surpass 1. Normalizing the intermolecular contact map in the same way, the maximal number of contacts between residues \(i\) and \(j\) could be larger than 1 because amino acid \(i\) of one molecule could be surrounded by several amino acids \(j\) belonging to different molecules.

The contact map described above did not distinguish between the types of interactions between the two amino acids in contact. The contact between two amino acids could be due to the effective hydrophobic attraction, attractive electrostatic interaction, backbone hydrogen bonding, or simply a result of proximity to neighboring amino acids involved in an attractive interaction. Because the backbone hydrogen bonding is key to the secondary structure formation, we focused specifically on amino acids involved in backbone hydrogen bond formation. We defined intra- and intermolecular hydrogen bond maps for each pair of amino acids as an average number of backbone hydrogen bonds between the two amino acids. The pairs of amino acids involved in backbone hydrogen bond formation were identified within each oligomer conformation (PDB file) using VMD. Depending on whether the amino acids belonged to the same or to different peptide chains, these pairs contributed to either intra- or intermolecular hydrogen bond maps. Each amino acid (except proline) can form up to two intermolecular backbone hydrogen bonds, as for example inside a \(\beta\)-sheet structure. Consequently, both intra- and intermolecular hydrogen bond map values can assume any value between 0 and 2. The final contact maps with the average numbers of intra- and intermolecular hydrogen bonds were then obtained by averaging the individual hydrogen bond maps over all oligomer conformations (PDB files) of specific oligomer order.
Relative distance per residue

By inspection, oligomers (dimers through hexamers) of the four peptides under study were quasispherical. To gain information on an average oligomer size and an average arrangement of individual residues within the oligomer with respect to its center of mass, we calculated the average relative distance from the center of mass per residue as in the prior work. For each oligomer, we first computed the center of mass and the distances of C\textsubscript{α} atoms of individual residues from the center of mass. We calculated separately the averages over all dimers and all hexamers for all four peptides, resulting in 8 different curves of the average relative distance per residue. We then could identify the peptide regions that were: (i) most distant from the center of mass and thus determined the average diameter of dimers and hexamers (the distance between the most distant C\textsubscript{α} atom from the center of mass was defined as one half of an average oligomer diameter); and (ii) closest to the center of mass. If these latter minimal relative distances (ii) were considerably larger than \(~0.5\) nm, this result would indicate that the shape of oligomers was more ellipsoidal than spherical, providing insight into their shape as well as the structure and assembly. In this case the estimation of the oligomer size as defined by (i) would correspond to the largest diameter of the ellipsoidally shaped oligomer.

Results

Initial conditions for our simulations included 32 non-interacting monomeric peptides per trajectory. These peptides were spatially separated and at the early stages of all production runs folding occurred prior to assembly into oligomers. Our initial study reported the sequence of folding events in A\textsubscript{β}1–40 and A\textsubscript{β}1–42, quantified the differences between A\textsubscript{β}1–40 and A\textsubscript{β}1–42 folded structures and was followed by a systematic study of monomer folding in A\textsubscript{β}1–40 and A\textsubscript{β}1–42 and their Arctic mutants. We focus here on the formation of oligomeric assemblies in A\textsubscript{β}1–40, A\textsubscript{β}1–42, [E22G]A\textsubscript{β}1–40, and [E22G]A\textsubscript{β}1–42.

Our DMD approach is based on two implicit solvent parameters, \(E_{HP}\) and \(E_{CH}\), which need to be adjusted to the specific solvent under study. Lam et al. used the DMD approach to study temperature dependence of the average β-strand in A\textsubscript{β}1–40 and A\textsubscript{β}1–42 folded monomers and
demonstrated that the average amount of $\beta$-strand in folded monomers critically depends on $E_{HP}$ (and not strongly on $E_{CH}$) and that the range for $E_{HP}$ that matches experimental circular dichroism intensities$^{44}$ best is $E_{HP} \in [0.3, 0.4]$, also in agreement with our earlier oligomerization study.$^{31}$ This comparison of the $\beta$-strand values obtained \textit{in silico} and \textit{in vitro} also resulted in an estimate of physiological temperature, 0.12-0.13 (in units of $E_{HB}/k_B$).$^{34}$ In the present study, we showed that by setting the two implicit solvent parameters $E_{HP} = 0.3$ and $E_{CH} = 0$, and using $T = 0.13$ for the physiological temperature, a \textit{simultaneous} match of both $\alpha\beta_{1-40}$ and $\alpha\beta_{1-42}$ oligomer size distributions to the experimental values was obtained. Then, we applied the same DMD simulation protocol with the \textit{same implicit solvent parameters and physiological temperature} to conduct the DMD simulations of two Arctic mutants to obtain $[\text{E22G}\alpha\beta_{1-40}]$ or $[\text{E22G}\alpha\beta_{1-42}]$ oligomer size distributions, which were experimentally shown to significantly differ from each other as well as from the distributions of the WT peptides.$^{13}$

The results of the present study are organized as follows. First, probability distributions of $\alpha\beta_{1-40}$ and $\alpha\beta_{1-42}$ conformers obtained by our DMD approach are analyzed and compared to the results of our earlier study.$^{31}$ Second, the role of effective electrostatic interactions on $\alpha\beta_{1-40}$ and $\alpha\beta_{1-42}$ assembly is studied. Third, using the same simulation parameters, probability distributions of $[\text{E22G}\alpha\beta_{1-40}]$ and $[\text{E22G}\alpha\beta_{1-42}]$ are derived and compared to \textit{in vitro} PICUP data.$^{13}$ Fourth, dimer and hexamer structures of all four peptides are characterized by calculating the $\beta$-strand and turn propensities, SASA per residue, and contact maps. Finally, formation of higher–order oligomers is addressed.

\textbf{Oligomer Size Distribution of Wild-Type $\alpha\beta$}

We found that the $\alpha\beta_{1-40}$ oligomer size distribution was not significantly different (Figure 1, thick black curve) from the one obtained in our previous work (Figure 1, thin black curve).$^{31}$ The new distribution for $\alpha\beta_{1-40}$ showed a slightly increased frequency of trimers and decreased frequencies of pentamers and hexamers. The $\alpha\beta_{1-42}$ oligomer size distribution, however, showed a significant amount of hexamers as well as an additional peak at dodecamer/tridecamer (Figure 1, thick red curve), which was not present in the previous simulations.
Doubling the simulation time (from $10 \times 10^6$ to $20 \times 10^6$) and using a more physiologic temperature estimate ($T = 0.13$ instead of $T = 0.15$) thus produced distributions that were more realistic, as determined by comparison with the *in vitro* oligomer size distributions obtained by PICUP.

### Role of Effective Electrostatic Interactions in Aβ Oligomerization

The results described above were obtained by using the same implicit solvent parameters, $E_{HP} = 0.3$ and $E_{CH} = 0$, as used previously, where all charged amino acids were treated as solely polar with no electrostatic charge. The reasoning behind this interaction scheme was an assumption that the charged amino acid side-chain groups are completely screened by surrounding water molecules at this early stage of assembly. In contrast to this assumption, substantial experimental evidence indicates that the D23–K28 salt bridge plays an important role in fibril formation of the WT peptides grown under agitated conditions. E22–K28 and D23–K28 salt bridges also were shown to stabilize a turn in the A21–A30 peptide segment that evidence suggests nucleates folding of full-length Aβ.

Prior DMD study on the effect of electrostatic interactions on Aβ$_{1-40}$ and Aβ$_{1-42}$ oligomer formation demonstrated that an effective electrostatic interaction potential energy $E_{CH} = 0.6$ promoted higher-order oligomer formation for both alloforms. The oligomer sizes were considerably larger than the sizes measured experimentally. This result could be due to the fact that in an aqueous solution water molecules interact with charged hydrophilic residues and thereby effectively screen the electrostatic interactions between pairs of charged residues. Consequently, the potential energy of effective electrostatic interactions of $E_{CH} = 0.6$ in the study of Yun *et al.* would be overestimated. However, a recent DMD study showed that Aβ$_{1-40}$ and Aβ$_{1-42}$ folding was almost unaltered over a wide temperature range for effective electrostatic interaction strengths $E_{CH} \leq 0.3$. It is thus important to address the question of whether subtle changes in the interaction between charged amino acids affects the oligomer size distribution.

To answer the question posed above, we studied three different cases: $E_{CH} = 0$, $E_{CH} = 10^{-6}$, and $E_{CH} = 10^{-2}$. The difference between the cases $E_{CH} = 0$ and $E_{CH} = 10^{-6}$ was in
the way the charged amino acids interacted among themselves. In the former case, a pair of charged amino acids experienced effective hydrophilic repulsion independent of the respective charges, while in the later case, a pair of charged amino acids experienced a small effective electrostatic attraction or repulsion, depending on their respective charges, but no effective hydrophilic repulsion. Because the potential energy associated with the later case was small ($10^{-6}$) compared to the thermal energy ($k_B T = 0.13$), charged amino acids effectively interacted only through repulsive hard-sphere excluded volume interactions. In the third case, $E_{CH} = 10^{-2}$, the strength of the effective electrostatic interactions among charged amino acids was still small but more comparable to the thermal energy ($\sim 10\%$ of $k_B T$).

We conducted simulations for both $\text{A}\beta_1-40$ and $\text{A}\beta_1-42$ at effective electrostatic potential energies $E_{CH} = 10^{-6}$ and $E_{CH} = 10^{-2}$ following the same protocol as for $E_{CH} = 0$. We then determined the oligomer size distributions and compared them to the case $E_{CH} = 0$ (Figure 2). Surprisingly, the slight changes in interaction schemes among the three cases resulted in significant changes in the oligomer size distributions. Decreasing the amount of hydrophilic repulsion among the charged hydrophilic residues by increasing $E_{CH}$ from 0 to $10^{-6}$ resulted in significantly increased numbers of larger oligomers for both $\text{A}\beta_1-40$ and $\text{A}\beta_1-42$. Increasing the effective electrostatic interaction energy further (to $10^{-2}$) resulted in another, though less pronounced, increase in the frequency of larger oligomers for both alloforms.

Why do effective electrostatic interactions have such a strong effect on oligomer size when they do not seem to affect monomer folding? If one compares the DMD results with experimental data, the best agreement is obtained with $E_{CH} = 0$, a condition under which charged residues are treated as non-charged and hydrophilic. This result can be understood if, at the initial stages of oligomer formation, the charged side-chain groups were effectively screened by water molecules, resulting in the absence of effective electrostatic interactions among the charged amino acids. These results demonstrate that initial oligomer formation is driven primarily by a hydrophobic collapse and that electrostatic interactions may only become important after desolvation of charged amino acid groups occurs. Note that neither $\text{A}\beta_1-40$ nor $\text{A}\beta_1-42$ oligomer size distributions demonstrated the presence of oligomers of order $n > 15$. In contrast, $\text{A}\beta_1-42$ size distribution in the cases $E_{CH} = 10^{-6}$ and $E_{CH} = 10^{-2}$ displayed a few
peaks in the range \( n > 15 \) (Figure 2b). Most of these peaks were associated with a standard error equal to the mean value of the occurrence probability, suggesting that these larger oligomers were a result of statistical fluctuations. It is significant, nonetheless, that such fluctuations were not present in the case \( E_{CH} = 0 \). In the case \( E_{CH} = 10^{-2} \), the \( \text{A}\beta_{1-42} \) oligomer size distribution exhibited a peak at \( n = 28 \) with the standard error somewhat smaller than the mean occurrence probability value, marking an onset of an \( \text{A}\beta_{1-42} \) assembly into larger oligomers (Figure 2b).

Inspection of the \( \text{A}\beta_{1-42} \) conformations at \( n = 28 \) confirmed that electrostatic interactions between charged residues increased the rate of conversion of quasispherical \( \text{A}\beta_{1-42} \) oligomers (characteristic of \( E_{CH} = 0 \)) to elongated protofibril-like (characteristic of \( E_{CH} = 10^{-2} \)) conformers (Figure 3).

Why was the formation of protofibril-like oligomers enhanced in the presence of electrostatic interactions? To answer this question, we first looked at the structural characteristics of monomeric and oligomeric assemblies regardless of particular assembly states and distributions of oligomer sizes. We analyzed the overall secondary structure to determine the structural changes resulting from different treatments of charged amino acids within the interaction schemes above. The major secondary structure component was \( \beta \)-strand. We thus compared \( \langle \beta \text{-strand} \rangle \) among all \( \text{A}\beta_{1-40} \) and \( \text{A}\beta_{1-42} \) trajectories at time frames from 19–20 × 10⁶ simulation steps for all \( E_{CH} \) three values (Table 1). Despite a trend towards higher \( \langle \beta \text{-strand} \rangle \) with higher \( E_{CH} \), the differences were not significant. The observed formation of larger oligomers thus was not accompanied by a significant increase in \( \langle \beta \text{-strand} \rangle \).

Dimer formation is the initial step in assembly. We thus explored the solvent exposed surface area (SASA) of \( \text{A}\beta_{1-40} \) and \( \text{A}\beta_{1-42} \) dimers that were present in all trajectories at time frames between 19 and 20 million simulation steps (Supporting Information, Figure S5). In all dimers, the amino acids or peptide segments most exposed to the solvent were D1, H6–S8, H13–Q15, E22, and S26–K28. Note that in all cases, with the exception of H13–Q15, one amino acid that is charged at neutral pH was present (D1, D7, E22, and K28). Among these, the only positively charged amino acid, K28, is in the turn region. The overall exposure to solvent was not drastically affected by the presence of electrostatic interactions. However, a more detailed comparison of the SASA for \( \text{A}\beta_{1-40} \) and \( \text{A}\beta_{1-42} \) dimers at \( E_{CH} = 0 \) showed that
the Aβ1−42 dimers had a significantly higher SASA in the A2-R5 region, whereas H6 and E11 in Aβ1−40 dimers were more exposed (Supporting Information, Figure S5b and Figure S5a, black curves). We point out that the inclusion of electrostatic interactions did not change the SASA of the Aβ1−40 dimers significantly (Supporting Information, Figure S5a), whereas it decreased the SASA of the Aβ1−42 dimers in the N-terminal region (Supporting Information, Figure S5b). This suggests that the charged N-termini of Aβ1−42 may participate in peptide assembly at $E_{CH} > 0$. These results also suggest that the formation of protofibrils (Figure 3), which was enhanced by the presence of electrostatic interactions, may be driven by a desolvation process that enables salt bridge formation between the most exposed charged amino acids of the proximate oligomers.

**Oligomer Size Distribution of Arctic Aβ**

The similarity of the oligomer frequency distributions produced computationally (Figure 1) and experimentally supports the use of our DMD approach to examine the effects of single amino acid Aβ substitutions causing familial forms of Alzheimer’s disease (FAD). The Arctic E22G amino acid substitution produces Aβ1−40 and Aβ1−42 alloforms whose folding pathways resemble that of Aβ1−42 but not that of Aβ1−40. Here, we explore the effects of the Arctic substitution on oligomerization. We find that [E22G]Aβ1−40 oligomerization involves an increased number of pentamers and hexamers (Figure 4, black dotted curve), comparable in number to those formed by Aβ1−42 (Figure 4, red solid curve). However, no significant peaks at higher oligomer order were found. In [E22G]Aβ1−42, the number of pentamers and hexamers was slightly decreased relative to Aβ1−42 and no peaks at decamers or higher order oligomers were observed (Figure 4, red dotted curve). However, at larger simulation times ($39-40 \times 10^6$) the relative number of [E22G]Aβ1−42 hexamers became comparable to the relative number of Aβ1−42 hexamers, mainly because the relative number of Aβ1−42 hexamers decreased as the relative number of Aβ1−42 dodecamers increased (Supporting Information, Figure S1). Dimers were the most abundant conformers in Aβ1−40, Aβ1−42, and [E22G]Aβ1−40, followed by trimers. Interestingly, [E22G]Aβ1−42 was characterized by an increased number of trimers, followed a slightly lower frequency of dimers. These size distribution characteristics agree
well with experimental findings.\textsuperscript{13}

We also addressed the time evolution of all four size distributions as well as compared
the individual statistical properties within and among distributions (Figures S1 and S2 and Ta-
ble S1, Supporting Information). The results displayed in Figures S1 and S2 and Table S1
demonstrate that the oligomer size distributions, which were initially the same (one peak at
oligomer size 1), developed into 4 statistically significantly different distributions, characteris-
tic of the peptide under study, within the first $20 \times 10^6$ time steps. On the time scale between
20 and $40 \times 10^6$ time steps, all distributions retained their characteristics while demonstrating
a slow evolution towards larger assemblies. The results of our analysis are consistent with a
view that oligomers represent metastable, slowly evolving states on the pathway to larger as-
semblies and not fixed states, “frozen” in time. Time scales that would allow for an observation
of statistically significant numbers of assemblies of order $n > 20$ might be larger than $100 \times
10^6$ time steps and might require larger scale simulations with an order of magnitude larger
number of peptides per trajectory.

\section*{Structural Characterization of Wild-Type and Arctic A\textsubscript{\textbeta} Oligomers}

We explored here the initial oligomerization of WT and Arctic (E22G) peptides of A\textsubscript{\textbeta}$_{1-40}$ and
A\textsubscript{\textbeta}$_{1-42}$ using the interaction scheme $E_{CH} = 0$. This scheme yielded oligomer size distributions
for A\textsubscript{\textbeta}$_{1-40}$ and A\textsubscript{\textbeta}$_{1-42}$ that agreed best with experimental results.\textsuperscript{9} The scheme also revealed
an initial hydrophobic collapse phase, during which low-order oligomers form, but before
desolvation of the charged amino acids promotes further assembly into protofibrils.

The structure of monomers and oligomers could be described predominately as collapsed
coils with $\beta$-strands separated by turns and loops as the dominant secondary structure ele-
ment. This result is consistent with the NMR data derived by Zhang et al., showing that A\textbeta
adopts a collapsed coil structure in water.\textsuperscript{40} For each of the four peptides, we calculated the
$\beta$-strand and turn propensities per residue for monomers alone and for each separate oligomer
order, taking into account all conformations at time frames between 19 and $20 \times 10^6$ simulation
steps (Figure 5). The average $\beta$-strand propensity across monomers and oligomers of order 2
(dimers) to 6 (hexamers) of all four peptides was 11–22\% (Figure 5a). For A\textsubscript{\textbeta}$_{1-40}$, A\textsubscript{\textbeta}$_{1-42}$,
and [E22G]Aβ₁₋₄₀. \( \langle \beta\text{-strand} \rangle \) was lowest in the monomer state and increased the most upon monomer to dimer transition. This observation is qualitatively similar to the observation of Ono et al., who studied cross–linked Aβ₁₋₄₀ monomers, dimers, trimers, and tetramers, and suggests that structural differences between monomers and dimers are larger than the differences among low-molecular weight oligomers of order \( n > 2 \). For each peptide individually, the relative increase in \( \langle \beta\text{-strand} \rangle \) from monomeric to hexameric oligomer state was significant as indicated by non-overlapping SEM bars and amounted to 32%, 39%, 70%, and 35% for Aβ₁₋₄₀, Aβ₁₋₄₂, and [E22G]Aβ₁₋₄₀, and [E22G]Aβ₁₋₄₂, respectively. The increase was the most prominent in [E22G]Aβ₁₋₄₀ even though the largest amount of \( \langle \beta\text{-strand} \rangle \) in hexamers was observed for [E22G]Aβ₁₋₄₂, which also had significantly larger amounts of \( \langle \beta\text{-strand} \rangle \) in a monomeric state compared to the other three peptides. The average turn propensity was 40–50% in all peptides (Figure 5b) and did not change significantly with oligomer order in Aβ₁₋₄₀, Aβ₁₋₄₂, or [E22G]Aβ₁₋₄₀. In [E22G]Aβ₁₋₄₂, however, we observed a significant (20%) decrease in the average turn propensity from the monomeric state (48.5%±3.5%) to the hexameric state (40.5%±2.9%).

Dimer formation is the first step in oligomer formation, but it has been suggested that pentamers and hexamers of Aβ₁₋₄₂ play a key role in Aβ₁₋₄₂ toxicity. We thus examined the β-strand propensity per amino acid in dimers (Figure 6a) and hexamers (Figure 6b) of all four peptides under study. The most striking difference between the β-strand propensities per residue among the four peptides was the presence of a short β-strand at A2–F4 in Aβ₁₋₄₀ oligomers (Figure 6a-b, black solid curves) that was completely absent in oligomers of the other three peptides (Figure 6, red solid and all dotted curves). This feature was shown previously to distinguish Aβ₁₋₄₀ folded monomer structure from the folded monomer structures of Aβ₁₋₄₂, [E22G]Aβ₁₋₄₀, and [E22G]Aβ₁₋₄₂. The present data demonstrated that in Aβ₁₋₄₀, β-strand propensities at A2–F4 increased upon oligomerization and were highest in hexamers, suggesting that this region plays a significant role in Aβ₁₋₄₀ oligomer assembly. We carried out a detailed pair-wise comparison of the \( \langle \beta\text{-strand} \rangle \) per residue and described it in Supplementary Information (Figure S4). To further examine the significance of the A2–F4 β-strand in Aβ₁₋₄₀ oligomer formation, we studied the time evolution of the interaction...
between a tetramer and dimer during \(A\beta_{1-40}\) hexamer formation (Figure 8). A short parallel intermolecular \(\beta\)-strand (A2–F4) in a dimer (upper right) interacts with one of the N-terminal regions in the tetramer (lower left). Initially, the N-terminal region docked onto the existing \(\beta\)-sheet (Figure 8A) and eventually adopted a \(\beta\)-strand structure (Figure 8B). This process was particular to \(A\beta_{1-40}\) hexamer formation and was observed in 13 out of 24 \(A\beta_{1-40}\) hexamers across different trajectories and time frames. Oligomerization of the other three peptides was not accompanied by formation of \(\beta\)-strands and the merging of two conformers into a larger one happened on time scales smaller than \(0.1 \times 10^6\) (time between the saved simulation frames).

To gain insight into the effect of the Arctic mutation on solvent exposure of different regions of the peptides, we calculated the average SASA per residue in dimers (Supporting Information, Figure S5a) and hexamers (Supporting Information, Figure S5b) for all four peptides. The results indicated that the overall structures of the C-terminal halves of the peptides comprising the oligomers were the same in all four peptides. The regions CHC, MHR, and CTR comprised the oligomer core, which was the most shielded from the solvent. In hexamers, the oligomer core had, on average, smaller SASAs than in dimers, suggesting that the hexamer core was more shielded from the solvent than the dimer core. SASAs at the N-terminal region were not affected by the higher oligomer order but differed among the four peptides. Differences in SASA among the four peptides were observed in the N-terminal region D1–R5, where \(A\beta_{1-40}\) oligomers showed significantly lower SASAs relative to oligomers of the other three peptides. This result was consistent with the existence of the \(\beta\)-strand structure at A2–F4 in only \(A\beta_{1-40}\) oligomer structures, resulting in lower SASAs caused by the involvement of the A2–F4 region in intra- and intermolecular contacts. An important observation was that SASAs of the N-terminal region D1–R5 within \(A\beta_{1-42}\), [E22G]\(A\beta_{1-40}\), and [E22G]\(A\beta_{1-42}\) were similar to each other and were all significantly higher than the corresponding SASAs in \(A\beta_{1-40}\). Unlike \(A\beta_{1-40}\), the other three peptides possessed N-termini that did not appear to be involved in intra- and intermolecular contacts and were thus mostly unstructured.

To gain additional insight into amino acid geometry within oligomers, we calculated the average relative distance from the center of mass per residue (Figure 9). First, we estimated the dimer and hexamer average sizes by identifying the largest relative distance from the center
of mass. Twice the value of this distance was used to estimate the average diameter of dimers (Figure 9a) $d_2 \approx 3$ nm and the average dimer of hexamers (Figure 9b) $d_6 \sim 4.6$ nm. The average hexamer diameter was slightly larger than the value $\sim 4.3$ nm calculated by using the dimer diameter (3 nm), assuming that the number of peptides $n$ in an oligomer scaled with its volume $V \propto d^3$ as $n \propto d^3$, where $d$ was the diameter of the oligomer of order $n$. This could be explained by taking into account the shape variability of hexamers, which was in many cases more elongated than spherical (in contrast to more consistently spherical dimers), resulting in increased relative distances from the center of mass. This particular scaling of the oligomer size from dimers to hexamers, resulting in less densely packed hexamers relative to dimers, is in agreement with the observed non-linear increase of oligomer sizes with the oligomer order $n$ in cross-linked $\text{A}\beta_{1-40}$ monomers, dimers, trimers, and tetramers in a recent study by Ono et al.$^{59}$ Our observation of oligomer shapes ranging from spherical to elongated ellipsoidal is also consistent with earlier neutron scattering data demonstrating a prolate ellipsoid shape for $\text{A}\beta$ oligomers.$^{60}$

Second, we compared the relative arrangement of different peptide regions with respect to the center of mass within dimers and hexamers. Except at the N-terminal region A1–Y10, the relative arrangement of residues was similar for all four peptides in dimers (Figure 9a). The region A1-R5 of $\text{A}\beta_{1-40}$ dimers was characterized by significantly lower relative distances than the ones of $\text{A}\beta_{1-42}$, [E22G]$\text{A}\beta_{1-40}$, and [E22G]$\text{A}\beta_{1-42}$ dimers (Figure 9a). In dimers of the two shorter peptides, $\text{A}\beta_{1-40}$ and [E22G]$\text{A}\beta_{1-40}$, the region V18-F20 was the closest to the center of mass (at $\sim 0.5$ nm). In dimers of the two longer peptides, $\text{A}\beta_{1-42}$ and [E22G]$\text{A}\beta_{1-42}$, I41 was the closest to the center of mass (at $\sim 0.4$ nm). The peptide region D7-G9 was characterized by the largest distance of $\sim 1.5$ nm from the center of mass in dimers of all four peptides.

Relative distances at A1-R5 were significantly larger in $\text{A}\beta_{1-42}$ hexamers than in $\text{A}\beta_{1-40}$ hexamers, while distances at A1-R5 of both Arctic hexamers fell between the ones of $\text{A}\beta_{1-40}$ and $\text{A}\beta_{1-42}$ hexamers (Figure 9b). The region D7-S8 had the largest relative distance from the center of mass in hexamers of all 4 peptides ($\sim 2.3$ nm). D1 of $\text{A}\beta_{1-42}$ hexamers was also at the same distance of $\sim 2.3$ nm. These differences in the relative distances between the four
peptides were consistent with the SASA results, which demonstrated significant differences in solvent exposure at the N-terminal region of dimers and hexamers (Supporting Information, Figures S5a and S5b).

\( \alpha \beta_{1-40} \) hexamers were characterized by the minimal distance of \( \sim1.4 \) nm at V40 and slightly larger distances of \( \sim1.5 \) nm at V18-F20, which were, unlike in dimers, significantly higher than in hexamers of the three other peptides (Figure 9b, black solid curve). The cause of relatively large distances at V18-F20 from the center of mass in \( \alpha \beta_{1-40} \) hexamers was the existence of a substantial number of \( \alpha \beta_{1-40} \) hexamer conformations similar to the one shown in Figure 8, where a hexamer consisted of two distinct interacting spherical oligomers of lower order (in Figure 8 a dimer and a tetramer). In \( \alpha \beta_{1-42} \) hexamers V40-I41 and L34-M35 were closest to the center of mass with relative distances of \( \sim1.2 \) nm (Figure 9b, red solid curve). The relative distances at I31-A42 of [E22G]\( \alpha \beta_{1-40} \) and [E22G]\( \alpha \beta_{1-42} \) hexamers fell between the \( \alpha \beta_{1-42} \) and \( \alpha \beta_{1-40} \) curves (Figure 9b, black and red dotted curves). Here the relative distances of [E22G]\( \alpha \beta_{1-40} \) hexamers were smaller than in [E22G]\( \alpha \beta_{1-42} \) hexamers and thus more similar to distances found in \( \alpha \beta_{1-42} \) hexamers (Figure 9b, red solid and black dotted curves). Relative distance values at the region I31-V40 (I31-A42) were correlated with the compactness of hexamers of the four peptides. The most compact were the \( \alpha \beta_{1-42} \) hexamers and the least \( \alpha \beta_{1-40} \) hexamers for the reasons discussed above. Interestingly, among the two Arctic peptides, [E22G]\( \alpha \beta_{1-42} \) hexamers were less compact, consistent with our observation that some amount of conformations showed [E22G]\( \alpha \beta_{1-42} \) hexamers composed of two interacting oligomers of lower order similar to \( \alpha \beta_{1-40} \) hexamers. Unlike \( \alpha \beta_{1-40} \) hexamers, however, the interaction between these two [E22G]\( \alpha \beta_{1-42} \) oligomers did not proceed through formation of \( \beta \)-strands at A2–F4.

**Peptide regions involved in \( \alpha \beta \) oligomerization**

To identify peptide regions that play key roles in oligomer formation, we quantified the tertiary and quaternary structures comprising \( \alpha \beta \) oligomers. Comparison of regions identified within oligomer populations formed by each of the four peptides provided the means to elucidate the distinct structures and peptide regions mediating common and distinct assembly paths of the
peptides. We computed the contact maps of dimers and hexamers for all four peptides (Supporting Information, Figures S6 and S7) and derived two types of contact maps (see Methods). Considering all contacts between each pair of amino acids, irrespective of the interaction type, we obtained intra- (Supporting Information, Figures S6, lower left triangles of all maps) and intermolecular (Supporting Information, Figures S7, lower left triangles of all maps) contact maps. In addition, we analyzed the backbone hydrogen bonding to be able to identify particular patterns of secondary structure formation within each molecule (intramolecular) or among molecules (intermolecular). These intra- (Supporting Information, Figures S6, upper right triangles of all maps) and intermolecular (Supporting Information, Figures S7, upper right triangles of all maps) maps contain hydrogen bond propensities for each pair of residues.

A detailed description of contact maps is given in Tertiary and quaternary structures of Aβ oligomers in Supporting Information. The comparisons made in Supporting Information provide a large amount of structural information. Aβ1−40 and Aβ1−42 oligomers were the most distinct structurally. [E22G]Aβ1−40 and [E22G]Aβ1−42 oligomers shared characteristics with oligomers formed by both WT peptides. Our analyses revealed a number of important structural characteristics of the oligomers:

1. The tertiary structures of peptides comprising oligomers of all four peptides were characterized by a hydrogen bonding pattern consistent with small, but non-zero, α-helix propensity in the G25–G33 region. The propensity was higher, and involved a more lengthy peptide segment (G22–G33), in both Arctic peptides.

2. Aβ1−40 oligomer formation proceeded predominantly through interactions among the CHC regions and intermolecular β-strand formation at the N-terminal region A2–F4.

3. Aβ1−42 oligomer formation proceeded predominantly through intermolecular interactions involving the C-terminal regions I31–A42 as well as interactions between I31–A42 and the CHC region.

4. Aβ1−40 dimers and hexamers had indistinguishable tertiary structures. In contrast, the average number of intramolecular (tertiary structure) contacts in Aβ1−42 hexamers was smaller than in Aβ1−42 dimers.
5. [E22G]Aβ1–40 oligomer formation proceeded predominantly through intermolecular interactions between the C-terminal region I31–V40 and the CHC. Intermolecular contacts among the CHC regions as well as contacts among the I31–V40 regions were stronger than in Aβ1–40 and weaker than in Aβ1–42.

6. [E22G]Aβ1–42 oligomerization involved intermolecular contacts among the C-terminal region I31–V40 as well as contacts between the CHC and I31–V40. In addition, the A2–F4 region participated in intermolecular contacts with the CHC and I31–A42 more than in Aβ1–42 and [E22G]Aβ1–40 but less than in Aβ1–40.

7. As with Aβ1–42, but not with Aβ1–40, a loss of tertiary structure contacts was observed during the assembly transition of [E22G]Aβ1–40 from dimer to hexamer, but the effect was smaller than that observed with Aβ1–42.

8. In contrast to the other three peptides, a small increase in the number of tertiary structure contacts was observed during the transition of [E22G]Aβ1–42 from dimer to hexamer.

These observations demonstrate that the balance of intra- and in particular intermolecular interactions among different peptide regions determines the oligomerization propensity of the four peptides. Aβ1–42, in which the center of intermolecular interactions was the C-terminal region, formed larger oligomers compared to Aβ1–40, where the center of intermolecular interactions was the CHC region (with additional strong participation of the A2–F4 region). The strongest intermolecular interactions in [E22G]Aβ1–40 hexamers also involved the C-terminal region I31–V40, but to a lesser extent than in Aβ1–42. Correspondingly, the oligomerization propensity was less than in Aβ1–42 but more than in Aβ1–40. The observed oligomerization tendencies of [E22G]Aβ1–42 were consistent with the above results. Here, the intermolecular interactions were centered more at the CHC than at I31–A42, and thus its oligomerization propensity was between that of [E22G]Aβ1–40 and that of Aβ1–40.

**Formation of higher order Aβ1–42 oligomers**

It has been suggested recently that Aβ dodecamers are the proximate neurotoxins in AD. We examined the structure of dodecamers of Aβ1–42, the only peptide in our study that dis-
played significant numbers of oligomers of order 11–13 at $E_{CH} = 0$, the implicit solvent parameter that best described Aβ oligomerization. We showed that even small effective electrostatic interactions ($E_{CH} = 10^{-2}$) induced formation of elongated protofibril-like assemblies. We thus asked whether Aβ$_{1-42}$ dodecamers obtained at $E_{CH} = 0$, or undecamers obtained at $E_{CH} = 10^{-2}$, were structurally similar or distinct. Assuming that undecamers at $E_{CH} = 10^{-2}$ were representative of the protofibrillar structure and dodecamers at $E_{CH} = 0$ representative of oligomeric structure, this comparison would help elucidate structural changes involved in assembly from oligomers to protofibrils.

Tertiary structures of Aβ$_{1-42}$ dodecamers obtained at $E_{CH} = 0$ (Figure 10a) and Aβ$_{1-42}$ undecamers obtained at $E_{CH} = 10^{-2}$ (Figure 10c) were similar in that the strongest intramolecular contacts involved the regions I31–A42. However, the contact strengths differed in all four regions indicated by boxes (1–4). Dodecamers were characterized by significantly less tertiary structure contacts than protofibril-like undecamers: (i) in the central folding region (Figure 10a and c, Box 4); and (ii) in the N-terminal region A2–F4 interacting with the CHC, MHR, and CTR (Figure 10a and c, Box 2).

Compared to Aβ$_{1-42}$ hexamers, Aβ$_{1-42}$ dodecamers had even less intramolecular contacts in the central folding region (Figure 10a, Box 4), suggesting that additional unfolding of individual Aβ$_{1-42}$ peptides was needed for their assembly into dodecamers. In contrast, tertiary contacts in the central folding region of protofibril-like undecamers resembled those of Aβ$_{1-40}$ dimers and hexamers (Figure 10c, Box 4) but were characterized by a larger number of hydrogen bonds (Figure 10c, Box 4, upper right triangle). Tertiary contacts involving A2–F4 (Figure 10a, Box 2) in Aβ$_{1-42}$ dodecamers were stronger than the ones in Aβ$_{1-42}$ hexamers but weaker than in protofibril-like undecamers (Figure 10c, Box 2).

Quaternary structures of Aβ$_{1-42}$ dodecamers obtained at $E_{CH} = 0$ (Figure 10b) and Aβ$_{1-42}$ undecamers obtained at $E_{CH} = 10^{-2}$ (Figure 10d) also were different. The intermolecular contacts in regions enclosed in Boxes 3–5 involving the CTR, MHR, and CHC were the strongest intermolecular contacts in both dodecamers and in protofibril-like undecamers. However, these contacts were generally stronger in the dodecamers than in the undecamers. The undecamers were characterized by additional intermolecular interactions involving the A2–F4 region (Fig-
ure 10 d, Boxes 1 and 2) that were also characteristic of Aβ<sub>1−40</sub> hexamer formation, i.e., the existence of an intermolecular parallel β-strand (Figure 10 d, Box 1).

We also examined all four contact maps for the occurrence of any regular arrangement of hydrogen bonding (Figure 10a-d, upper right triangle). An off-diagonal intramolecular hydrogen bond pattern consistent with non-zero α-helix propensities was found in both intramolecular contact maps of dodecamers (Figure 10a, upper right triangle) and protofibril-like undecamers (Figure 10c, upper right triangle) and was most prevalent in the G25–G33 regions. In addition, this pattern also was present in regions with high turn propensities (Figure 7), suggesting that the turn and α-helix propensities were correlated and were enhanced by the presence of glycines in the sequence. In addition, a diagonal intermolecular hydrogen bond pattern consistent with parallel in-register intermolecular hydrogen bonds was observed in undecamers at $E_{CH} = 10^{-2}$ (Figure 10d, upper right triangle). Even though the corresponding hydrogen bond propensities in this pattern were low, this feature might indicate the onset of a parallel cross-β structure that is characteristic for full-length Aβ and several other amyloid fibrils. The fact that this parallel intermolecular hydrogen bond pattern was most pronounced in protofibril-like undecamers suggested that electrostatic interactions that become effective upon desolvation significantly contribute to formation of the parallel in-register cross-β structure.

**Discussion**

Amyloid β-protein belongs to a class of amyloid-forming proteins whose assembly is thought to cause a variety of neurodegenerative diseases. It is hypothesized that less structured oligomeric assemblies, that may be on- or off-pathway for amyloid fibril formation, are the most toxic structures triggering the disease. Examples include but are not limited to α-synuclein, relevant to Parkinson’s disease, and islet amyloid polypeptide, relevant to type 2 diabetes. Because toxicity is by definition directly associated with structural properties of these protein assemblies, it is imperative to address the question of how such a mutation alters the structural properties of the protein under study and the pathway of protein assembly.
In the present work, we studied oligomer formation of Aβ_{1-40}, Aβ_{1-42}, and their Arctic mutants, [E22G]Aβ_{1-40} and [E22G]Aβ_{1-42}, using the DMD approach combined with a four-bead protein model and implicit solvent. We examined oligomerization pathways of the four peptides, derived oligomer size distributions, and compared the distributions to in vitro data obtained by PICUP/SDS-PAGE. In vitro oligomer size distributions of a series of Aβ peptides obtained by PICUP/SDS-PAGE were shown to be sensitive to the peptide length and sequence. The oligomer size distribution thus can be considered a fingerprint of the peptide under study and means to distinguish oligomerization pathways of different Aβ variants. We used these data to adjust the values of the two implicit solvent parameters (strengths of the effective hydropathic and electrostatic interactions) that would produce simulated oligomer distributions best matching those obtained experimentally. By incorporating a better estimate of a physiologic temperature and using longer simulation times relative to our earlier study, we observed an Aβ_{1-42} oligomer size distribution that included not only a peak at pentamers/hexamers but in addition a peak at oligomer orders 12–13, which was not observed in our prior work. 

The Arctic mutation causes a severe, early onset form of AD. The propensity of [E22G]-Aβ_{1-40} and [E22G]Aβ_{1-42} to form paranuclei or higher order oligomers may be a critical factor contributing to the pathogenesis of this form of AD. [E22G]Aβ_{1-42} toxicity also was demonstrated in a recent study by Luheshi et al., who used rational mutagenesis of Aβ_{1-42} to find a strong correlation between aggregation propensity of a series of Aβ_{1-42} mutants and dysfunction they caused in a Drosophila model of Alzheimer disease. Among several Aβ_{1-42} variants, the Arctic mutant [E22G]Aβ_{1-42} was characterized by strongly enhanced toxicity relative to Aβ_{1-42}, presumably caused by oligomeric forms of [E22G]Aβ_{1-42}. To study the oligomerization of the two Arctic peptides, we applied our DMD simulation protocol to obtain a large population of in silico [E22G]Aβ_{1-40} and [E22G]Aβ_{1-42} oligomers. None of the Arctic mutants produced significant amounts of oligomers larger than nonamers, but [E22G]Aβ_{1-40} displayed a high probability of forming both pentamers and hexamers, while [E22G]Aβ_{1-42} produced increased numbers of trimers and also small amounts of pentamers and hexamers. These results were consistent with PICUP/SDS-PAGE data, except for the fre-
quency of [E22G]Aβ₁₋₄₂ paranuclei, which was experimentally found to be similar to the frequency of Aβ₁₋₄₂ paranuclei. This apparent discrepancy was resolved by further extending the simulation time, during which the number of Aβ₁₋₄₂ hexamers slightly decreased and the number of Aβ₁₋₄₂ dodecamers increased, resulting in comparable relative numbers of Aβ₁₋₄₂ and [E22G]Aβ₁₋₄₂ paranuclei.

We also examined whether the probability distributions of oligomer sizes for each of the four peptides reached a steady–state. We quantified the changes in all four distributions due to time evolution on different time scales. Our data showed that by considering time scales of up to $5 \times 10^6$ simulation steps, all four size distributions reached a metastable steady state within the first $15 \times 10^6$ time steps. However, slower time evolution on time scales of $20 \times 10^6$ was still observed, consistent with a view that oligomeric states represent quasi-stable intermediates, possibly on the pathway to protofibril and fibril formation.

We explored the role of the electrostatic interaction on Aβ oligomerization and demonstrated that the oligomer size distribution was sensitive to electrostatic interactions among the charged amino acids. The best fit to experimental oligomer size distributions was obtained if the side chain charges were completely neglected and the charged side chains were modeled as pure hydrophiles. These findings are consistent with the hypothesis that formation of oligomers is driven by a hydrophobic collapse, during which water interacting with the hydrophilic side chains efficiently solvates the charged residues, inhibiting the electrostatic interactions among them. We further showed that if the charged side chains were allowed to interact electrostatically, the spherical oligomers readily evolved into elongated protofibrils, even at very small strength of effective electrostatic interaction. We showed that this mechanism was a consequence of structural properties of quasispherical oligomers, which were characterized by strong exposure to the solvent of charged residues at the N-terminus (D1, E3, R5, D7) and at the turn region A21–A30 (E22, D23, K28).

We then structurally analyzed spherical oligomers produced by the four different peptides. As shown previously, Aβ₁₋₄₀ and Aβ₁₋₄₂ oligomers were characterized by significant differences in β-strand and turn propensities at the N-terminal region D1–Q15, even though their sequences differ at the C-termini. We showed that the secondary structure propensities at the
N-termini of oligomers of both Arctic mutants resembled the propensities of Aβ1–42 but not those of Aβ1–40. Thus, an amino acid substitution at the CHC (E22G) gives rise to structural differences in another region, the N-terminus. These results are consistent with a view that Aβ conformations are significantly affected by the competition between the N- and C-termini to form a stable complex with the CHC.71 Aβ1–42 oligomer formation was predominantly driven by intermolecular interactions among the CTRs, while in Aβ1–40 assembly the intermolecular interactions among the CHC regions were dominant. For the Arctic peptides, intermolecular interactions among the CTRs and CHC regions were equally important. Of the four peptides, only Aβ1–40 oligomerization was characterized by involvement of the N-terminal region A2–F4 in intermolecular interactions. We showed that Aβ1–40 hexamers assembled from oligomers of lower order through formation of a parallel β-strand at A2–F4, which appeared to be the rate-limiting process. This observation is consistent with existing PICUP/SDS-PAGE data demonstrating that N-terminally truncated Aβ3–40 and Aβ5–40 (in which parallel β-strand at formation at A2–F4 would be hindered or non-existent) displayed increased amounts of pentamers/hexamers, while Aβ3–42 and Aβ5–42 oligomer size distributions were indistinguishable from that of Aβ1–42.13 Further examination of intra- and intermolecular contacts between different peptide regions in each peptide demonstrated that [E22G]Aβ1–40 hexamers were structurally similar to Aβ1–42 hexamers, implicating the [E22G]Aβ1–40 hexamers in this familial form of AD. Considering that the shorter Aβ alloform is typically more abundant in the human body, the ability of [E22G]Aβ1–40 to form toxic oligomers might play a critical role in this familial form of AD.

We systematically examined the intra- and intermolecular hydrogen bonds in oligomers of all four peptides. The tertiary structure of all four peptides was characterized by a hydrogen bonding pattern consistent with low α-helix propensities in the region G25–G33 that were more pronounced in the G22–G33 region of both Arctic peptides. Glycines were also associated with increased turn propensities, so our data showed a correlation between the turn and α-helix propensities in all four peptides. This observation is consistent with in vitro findings that Aβ assembly into protofibrils and fibrils is preceded by a transitory increase in α-helical structure.43 We also examined Aβ1–42 dodecamers, which recently were suggested to
be the proximate neurotoxic species in AD. Similar α–helix propensities were observed in Aβ\textsubscript{1-42} dodecamers and in elongated protofibrillar Aβ\textsubscript{1-42} undecamers obtained at protofibril-inducing non-zero electrostatic interaction potentials. Interestingly, in elongated protofibrillar Aβ\textsubscript{1-42} undecamers, we observed a parallel intermolecular hydrogen bond pattern that was consistent with an onset of parallel, in-register cross-β structure, characteristic of full-length Aβ fibrils.

Our computational findings are consistent with the existence of three distinct processes occurring sequentially within a single Aβ assembly pathway: (1) fast hydrophobic collapse resulting in a population of oligomers; (2) slow desolvation enabling charged residues to interact electrostatically and thus induce formation of elongated protofibrils; and (3) very slow emergence of parallel in-register intermolecular hydrogen bonding associated with the cross-β structure of the amyloid fibril. Whether such an on-pathway scenario exists in \textit{in vitro} or \textit{in vivo}, and whether this scenario represents the only possible pathway of assembly, may be critical to understanding all amyloid-forming proteins. Experimental testing of the assembly pathway might be possible by further developing spectroscopic techniques that would allow observation of Aβ oligomerization in an unperturbed solvent on sub–second time scales in a time–resolved manner and with spatial resolution of at least a nanometer, such as for example time–resolved liquid transmission electron microscopy.

The two processes we discussed within the scope of the present study, i.e., hydrophobic collapse and slow desolvation, originate in an explicit aqueous environment in the way water molecules interact with hydrophobic versus hydrophilic residues. Effective hydropathic and electrostatic interactions are caused or strongly affected by the polar nature of water molecules and their ability to form hydrogen bonds among themselves and with hydrophilic protein groups. Thus, both processes are directly related to the thermodynamics and dynamics of water. Here, hydrophobic collapse refers to a process during which the hydrophobic groups minimize their exposure to water in order to maximize the number of hydrogen bonds among water molecules, which is energetically favorable. Because water molecules do not form hydrogen bonds with hydrophobic protein groups, this process is expected to occur rapidly. On the other hand, desolvation of hydrophilic residues requires water molecules to break hydrogen
bonds with hydrophilic residues, which is energetically unfavorable, resulting in longer times needed for desolvation.

Our present results demonstrate that the time scale over which hydrophobic collapse occurs is sensitive to Aβ length and to single amino acid substitutions. The process of desolvation on Aβ_{10−35} monomer conformation was studied using all-atom MD by Massi and Straub, who showed that the amino acid substitution associated with the Dutch mutation (E22Q) induced changes in the first solvation layer of water that resulted in enhanced desolvation, a key process in assembly. Recently, all-atom MD simulations of two amyloidogenic protein fragments were carried out to monitor desolvation occurring upon addition of a monomer to a preformed fibril. Consistent with our data, Reddy et al. suggest that the desolvation process, which is protein-sequence dependent, plays an important role in fibril formation. We demonstrated that our DMD approach captures the majority of distinct characteristics of Aβ_{1−40}, Aβ_{1−42}, [E22G]Aβ_{1−40} and [E22G]Aβ_{1−42} oligomer size distributions, provides mechanistic insights into elongated protofibril formation, and provides structural information on toxic Aβ oligomers relevant to structure-based drug design.

This research was supported by NIH grant AG027818.
Table 1: The average amounts of the β-strand propensity in $A\beta_{1-40}$ and $A\beta_{1-42}$ at simulation times between 19 and 20 million steps for three different interaction schemes. The error bars correspond to SEM.

<table>
<thead>
<tr>
<th>(β-Strand)</th>
<th>$A\beta_{1-40}$</th>
<th>$A\beta_{1-42}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{CH} = 0$</td>
<td>$17.2 \pm 1.5%$</td>
<td>$18.6 \pm 1.5%$</td>
</tr>
<tr>
<td>$E_{CH} = 10^{-6}$</td>
<td>$16.9 \pm 1.5%$</td>
<td>$19.1 \pm 1.6%$</td>
</tr>
<tr>
<td>$E_{CH} = 0.01$</td>
<td>$18.6 \pm 1.6%$</td>
<td>$20.6 \pm 1.6%$</td>
</tr>
</tbody>
</table>

Table 2: Total number of conformers from monomers (n=1) to oligomers of order n=2 through n=6 used in the structural analysis. The conformers were obtained using 8 trajectories per peptide and 11 time frames between 19 and $20 \times 10^6$ time steps per trajectory.

<table>
<thead>
<tr>
<th>Number of Conformers</th>
<th>n=1</th>
<th>n=2</th>
<th>n=3</th>
<th>n=4</th>
<th>n=5</th>
<th>n=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A\beta_{1-40}$</td>
<td>108</td>
<td>396</td>
<td>336</td>
<td>110</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>$A\beta_{1-42}$</td>
<td>56</td>
<td>215</td>
<td>177</td>
<td>54</td>
<td>56</td>
<td>59</td>
</tr>
<tr>
<td>[E22G]$A\beta_{1-40}$</td>
<td>61</td>
<td>266</td>
<td>189</td>
<td>83</td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>[E22G]$A\beta_{1-42}$</td>
<td>69</td>
<td>129</td>
<td>230</td>
<td>144</td>
<td>57</td>
<td>43</td>
</tr>
</tbody>
</table>
Figure 1: Oligomer size distributions for $A\beta_{1-40}$ (black curve) and $A\beta_{1-42}$ (red curve) obtained in silico using DMD with the four-bead protein model at $T = 0.130$. Each distribution is an average over eight $20 \times 10^6$ simulation steps long trajectories. All conformations at time frames at 19, 19.5, and $20 \times 10^6$ simulation steps were included in the analysis. Each trajectory involved 32 peptides, initially spatially separated and each in random coil conformation, enclosed in a 25 nm length cubic box. The error bars represent SEM. For comparison, old simulation data reported by Urbanc et al. for $A\beta_{1-40}$ (thin black curve) and $A\beta_{1-42}$ (thin red curve), obtained at $T = 0.150$ by averaging over eight $10 \times 10^6$ simulation steps long trajectories, are shown.
Figure 2: Oligomer size distributions for (a) Aβ_{1−40} and (b) Aβ_{1−42} obtained by DMD simulations using the four-bead protein model at $T = 0.130$ at different effective electrostatic interactions. Each distribution is an average over eight $20 \times 10^6$ simulation steps long trajectories. All conformations at time frames at 19, 19.5, and $20 \times 10^6$ simulation steps were included in the analysis. Each trajectory involved 32 peptides, initially spatially separated and each in random coil conformation, enclosed in a 25 nm length cubic box. The error bars represent SEM.
Figure 3: Effective electrostatic interactions speed up the conversion from hexamers ($n=6$) to elongated protofibrillar assemblies (e.g., $n=28$). β-Strands are depicted as yellow ribbons, turns as light blue and random coil as white. The amino acid D1, marking the N-termini, is represented by red spheres.
Figure 4: Oligomer size distributions for Arctic mutants [E22G]Aβ₁–₄₀ (dotted black curve) and [E22G]Aβ₁–₄₂ (dotted red curve) obtained by DMD simulations using the four-bead protein model at \( T = 0.130 \). Each distribution is an average over eight \( 20 \times 10^6 \) simulation steps long trajectories. The oligomer size distributions for Aβ₁–₄₀ (solid black curve) and Aβ₁–₄₂ (solid red curve) are shown for comparison. All conformations at time frames at 19, 19.5, and \( 20 \times 10^6 \) simulation steps were included in the analysis. Each trajectory involved 32 peptides, initially spatially separated and each in random coil conformation, enclosed in a 25 nm length cubic box. The error bars represent SEM. The inset adapted from Bitan et al.\(^{13}\) shows the experimental data obtained by PICUP/SDS-PAGE for (a) Aβ₁–₄₀, (b) Aβ₁–₄₂, (c) [E22G]Aβ₁–₄₀, and (d) [E22G]Aβ₁–₄₂.
Figure 5: Average propensities for (a) $\beta$-strand and (b) turn formation in monomers, dimers, trimers, tetramers, pentamers, and hexamers for $A\beta_{1-40}$ (solid black curve), $A\beta_{1-42}$ (solid red curve), $[E22G]A\beta_{1-40}$ (dotted black curve), and $[E22G]A\beta_{1-42}$ (dotted red curve). The error bars correspond to SEM.
Figure 6: The average β-strand propensity per amino acid for (a) dimers and (b) hexamers of Aβ₁₋₄₀ (solid black curve), Aβ₁₋₄₂ (solid red curve), [E22G]Aβ₁₋₄₀ (dotted black curve), and [E22G]Aβ₁₋₄₂ (dotted red curve). The error bars correspond to SEM.
Figure 7: The average turn propensity per amino acid for (a) dimers and (b) hexamers of $A\beta_{1-40}$ (solid black curve), $A\beta_{1-42}$ (solid red curve), $[E22G]A\beta_{1-40}$ (dotted black curve), and $[E22G]A\beta_{1-42}$ (dotted red curve). The error bars correspond to SEM.
Figure 8: Formation of an Aβ₁−₄₀ hexamer from a tetramer and a dimer at (A) $19.0 \times 10^6$ and (B) $19.6 \times 10^6$ simulation steps. β-Strands are depicted as yellow ribbons, turns as light blue and random coil as white. The amino acid D1, marking the N-termini, is represented by red spheres, V39 and V40 are marked as orange spheres.
Figure 9: Relative distances of individual residues from the center of mass of $\mathrm{A}^\beta_{1-40}$ (solid black curve), $\mathrm{A}^\beta_{1-42}$ (solid red curve), $\mathrm{[E22G]}A^\beta_{1-40}$ (dotted black curve), and $\mathrm{[E22G]}A^\beta_{1-42}$ (dotted red curve) (a) dimers and (b) hexamers. The error bars correspond to SEM.
Figure 10: **Intra- and intermolecular contact maps of Aβ1–42 oligomers, (a-b) n = 12 for E_{CH} = 0** (average taken over 12 conformers) and (c-d) n = 11 for E_{CH} = 10^{-2} (average taken over 30 conformers). The lower triangle contains the average number of contacts between two amino acids and the upper triangle contains the average number of hydrogen bonds for each pair of amino acids. The scale on the right shows the color mapping. The two types of maps have different scales, the scale on the left corresponds to the average number of contacts and the scale on the right corresponds to the average number of hydrogen bonds. The two thin diagonal lines are drawn through the diagonal elements of the two types of contact maps. The rectangular gray boxes with numbers mark regions of interest.
(1) We note that “order” generally is used herein to define the number of Aβ monomers comprising a particular assembly. Unless stated otherwise, “order” does not refer to secondary or tertiary structure characteristics.


(3) Hardy, J. Ann Neuro 2003, 54, 143-144.


(11) A paranucleus is defined by its secondary structure (~18–24% β-sheet/β-turn and ~3–7% α-helix), morphology (quasispherical by electron microscopy or atomic force microscopy), and quaternary structure (pentamer or hexamer).


(48) We use the term “oligomer conformation” to refer to a collection of all spatial coordinates of all atoms of all peptide chains comprising a single oligomer.


(52) An ellipsoidal shape of an oligomer of the n-th order would indicate that the oligomer was formed by merging two spherical oligomers of orders $n_1$ and $n_2$, $n=n_1+n_2$.


(62) Under these conditions, an insufficient number of dodecamers were observed to allow further analysis.

(63) [E22G]Aβ1-42 hexamers showed a similar though less pronounced pattern, see Supporting Information, Figure S5h, upper right triangle.


While a true steady state of Aβ monomers and oligomers of different order may not exist, oligomer size distributions were experimentally observed to be stable in the time period prior to protofibril formation.
TOC Graphics
Supporting Information. *Elucidation of amyloid* 

**β-protein oligomerization mechanisms: Discrete Molecular Dynamics Study**

B. Urbanc\(^1\)*, M. Betnel\(^2\), L. Cruz\(^1\), G. Bitan\(^3\),\(^4\), and D. B. Teplow\(^3\),\(^4\)

\(^1\)Department of Physics, Drexel University, Philadelphia, PA 19104, USA; \(^2\)Department of Physics, Boston University, Boston, MA 02215; \(^3\)Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095-7334, USA; \(^4\)Molecular Biology Institute and Brain Research Institute, University of California, Los Angeles, CA 90095-7334, USA.

**Time Evolution of Oligomer Size Distributions**

Initially, each DMD simulation trajectory consisted of 32 spatially–separated peptides in random-coil–like conformations. Thus, the probability distributions of oligomer sizes for all four peptides (A\(β\)\(_{1–40}\), A\(β\)\(_{1–42}\), [E22G]A\(β\)\(_{1–40}\), and [E22G]A\(β\)\(_{1–42}\)) at the simulation time \(t = 0\) were equivalent and characterized by a probability \(P(n) = 1\) at the oligomer size \(n=1\) (and \(P(n) = 0\) for all \(n > 1\)). Figs. S1a-d show probability distributions, \(P(n)\), for each of the four peptides at 4 different time windows: (1) 9-10 \(\times\) 10\(^6\), (2) 19-20 \(\times\) 10\(^6\), (3) 29-30 \(\times\) 10\(^6\), and (4) 39-40 \(\times\) 10\(^6\) simulation time steps. Within each time window, we selected frames that were as independent of each other as possible, to avoid biasing the calculation of the standard error of the mean (SEM) for each oligomer size. Within each time window, 3 time frames (e.g., at 9, 9.5, and 10 \(\times\) 10\(^6\) time steps) were thus selected for each of the 8 trajectories—in total 24 different populations of oligomers for each of the

\(^{\ast}\)Corresponding Author: brigita@drexel.edu
four peptides to calculate the four oligomer size distribution probabilities \( P(n) \). The results demonstrate how the characteristic peaks in \( P(n) \) change with simulation time on a \( 10 \times 10^6 \) time scale. These results show that the main characteristics of the probability distributions already develop within the first \( 10 \times 10^6 \) time steps. However, the distributions keep evolving with time. Of the four peptides, only \( \text{A} \beta_{1-40} \) was characterized by a probability distribution that remained the same on a long time scale, between 20 and \( 40 \times 10^6 \) time steps (Figure S1a). \( \text{A} \beta_{1-42} \) oligomers with a characteristic size 5-6 were already present at \( 10 \times 10^6 \) time steps. However, larger oligomers with sizes \( \sim 12 \) appeared at \( 20 \times 10^6 \) time steps and became more abundant between 20 and \( 40 \times 10^6 \) time steps, at the expense of \( \text{A} \beta_{1-42} \) hexamers (Figure S1b). The characteristic peak at oligomer sizes 5-6 in the \([\text{E22G}]\text{A} \beta_{1-40}\) distribution appeared at \( 20 \times 10^6 \) time steps and became more pronounced between 20 and \( 40 \times 10^6 \) time steps (Figure S1c). Interestingly, the \([\text{E22G}]\text{A} \beta_{1-42}\) distribution did not show very significant changes after \( 10 \times 10^6 \) time steps (Figure S1d). However, because the relative number of \( \text{A} \beta_{1-42} \) hexamers (Figure S1b) decreased between 30 and \( 40 \times 10^6 \) time steps, the relative numbers of \( \text{A} \beta_{1-42} \) and \([\text{E22G}]\text{A} \beta_{1-42}\) hexamers were similar at these longer time scales, which is a result consistent with PICUP observations. \(^1\)

Based on finite time-scale simulations, we cannot conclude whether any of four probability distributions reached a true steady state. However, we explored whether the distributions changed significantly on time scales of \( 1 \times 10^6 \), \( 2 \times 10^6 \), and \( 5 \times 10^6 \) simulation time steps by applying the chi–square test that compared two subsequent distributions and gave the p-value, i.e. the probability that the two distributions were statistically equivalent. For each of the four peptides, we first calculated histograms of oligomer sizes using a sliding time window: \( 1-2 \times 10^6 \), \( 2-3 \times 10^6 \), \( 39-40 \times 10^6 \) time steps. Within each of the sliding time windows, we selected populations of monomers and oligomers of all 8 trajectories at 3 time frames, e.g., at 1, 1.5, and \( 2 \times 10^6 \), to calculate the histograms of oligomer sizes for each of the four peptides. We then calculated the p-values of two subsequent histograms of oligomer sizes of the same peptide, which were separated by time lags, \( \Delta t \), of 1, 2, and \( 5 \times 10^6 \) time steps. The time evolution of the p-values is presented in Fig. S2 for each of the four peptides at the three different values of \( \Delta t \). The distribu-
tions were changing significantly (p-value < 0.05) within $\Delta t = 1 \times 10^6$ only within the first $5 \times 10^6$ time steps (Figure S2a). A similar conclusion can be made for changes within $\Delta t = 2 \times 10^6$ (Figure S2b), even though the p-value fluctuations were larger than in Figure S2a. However, the changes within $\Delta t = 5 \times 10^6$ (Figure S2c) were significant within the first 10-15 $\times 10^6$ time steps and were different for each of the four different peptides. The steady state of the probability distribution on time scales of $\Delta t = 5 \times 10^6$ (Figure S2c) was first reached by [E22G]A$\beta_1$-$42$ (at 5-10 $\times 10^6$), followed by A$\beta_1$-$40$ (at $\sim$10 $\times 10^6$), while A$\beta_1$-$42$ and [E22G]A$\beta_1$-$42$ needed the longest time to reach the steady state (at 15-20 $\times 10^6$), consistent with observations in Figure S1. The p-values in Figure S2c also showed the largest fluctuations, consistent with the largest time lag of $\Delta t = 5 \times 10^6$. Interestingly, on a time scale 20-40 $\times 10^6$, the two Arctic peptides displayed the largest fluctuations in p-values ([E22G]A$\beta_1$-$40$ even more so than [E22G]A$\beta_1$-$42$), most likely indicating an onset of assembly into larger structures, also consistent with emergence of small peaks at oligomer sizes $\geq 13$ in Figure S1, corresponding to elongated protofibril-like oligomers.

Because the time evolution of the probability distributions in Figure S1 showed some temporal changes between 20 and 40 $\times 10^6$ time steps, we next asked whether the probability distributions calculated within the time window 19-20 $\times 10^6$ significantly differed from those calculated within the time window 39-40 $\times 10^6$. The p-values of the corresponding chi–square test (along the diagonal elements of Table S1) demonstrated that: (1) the A$\beta_1$-$40$ distribution did not change significantly (p-value > 0.05); (2) the changes in the A$\beta_1$-$42$ and [E22G]A$\beta_1$-$42$ distributions were on the border of significance (0.01 < p-value < 0.05) and (3) [E22G]A$\beta_1$-$40$ distribution changed significantly (p-value = 0.0089). Despite these changes, the overall characteristics of distributions belonging to the four peptides did not change. For example, [E22G]A$\beta_1$-$40$ distribution, which was the only one that changed significantly, showed at later times a smaller relative numbers of dimers and trimers and a larger relative number of hexamers. Thus, the relative number of [E22G]A$\beta_1$-$40$ paranuclei, which was comparable to the relative number of A$\beta_1$-$42$ paranuclei within the time window 19-20 $\times 10^6$, increased and surpassed the relative number of A$\beta_1$-$42$ paranuclei within the time window 39-40 $\times 10^6$ (Figure S1). We also examined the temporal changes in individual
distributions by quantifying the distribution differences among the four peptides at a fixed time window. The results of this cross-comparison are reported in Table S1 for each of the two time windows: $19-20 \times 10^6$ (the p-values below the diagonal) and $39-40 \times 10^6$ (the p-values above the diagonal). These off-diagonal p-values were one or more orders of magnitude smaller than the p-values along the diagonal, demonstrating that the distribution differences among the four different peptides were significantly larger than the time–induced distribution differences of the same peptide.

**β-Strand Propensity Per Residue**

Our results showed that the β-strand secondary structure was the most prominent secondary structure in oligomers of all four peptides under study. Different peptide regions were shown to have distinct propensities to form the β-strand structure (Figure 6), depending both on the specific peptide as well as the oligomer order. To elucidate these differences in detail, we systematically compared the β-strand propensity per residue for $\text{A}\beta_{1-40}$, $\text{A}\beta_{1-42}$, $[\text{E22G}]\text{A}\beta_{1-40}$, and $[\text{E22G}]\text{A}\beta_{1-42}$ dimers and hexamers, by plotting all 28 pairs of β-strand propensity versus residue curves (Figure S4). The graphs in Figure S4 are arranged such that the first three rows include 6 graphs corresponding to data of all possible dimer pairs. Analogously, the last three columns include 6 graphs corresponding to data of all possible hexamer pairs. The remaining 16 graphs (a square matrix of 16 graphs comprising the last 4 rows and the first 4 columns) contain data of dimers-hexamers pairs of the four peptides.

**Dimer-Dimer Pairs.** Comparing the data of all dimer pairs (the first 3 rows of graphs in Figure S4), $\text{A}\beta_{1-40}$ and $\text{A}\beta_{1-42}$ dimers showed overall the largest differences in the β-strand propensities per residue curves ($P(2, 1)$). $\text{A}\beta_{1-40}$ and $\text{A}\beta_{1-42}$ β-strand propensities differed at the N-terminal region A2-F4, where $\text{A}\beta_{1-40}$ but not $\text{A}\beta_{1-42}$ showed β-strand propensity of up to 0.3 ($P(2, 1)$). Of all dimers, only $\text{A}\beta_{1-40}$ dimers were characterized by the β-strand propensity at A2-F4. In contrast, $\text{A}\beta_{1-42}$ dimers showed a significantly higher β-strand propensity at R5-H6, G9-E11, L17-A21, and V39-I41 ($P(2, 1)$). Significant β-strand propensities at R5-H6 and G9-V11 were observed
in all but Aβ1−40 dimers. [E22G]Aβ1−40 dimers showed increased β-strand propensities at F19-V24 of similar values to those in Aβ1−42 dimers at the CHC (L17-A21). β-Strand propensities at the CTR (V39-I41) were non-zero in Aβ1−42 and [E22G]Aβ1−42 dimers but were zero in Aβ1−40 and [E22G]Aβ1−40 dimers. Except for this difference at the CTR, [E22G]Aβ1−40 and Aβ1−42 dimers were characterized with the most similar β-strand propensities \((P(3,2))\). [E22G]Aβ1−42 and Aβ1−42 propensities were also quite similar except at R5-H6, H13-H14, and L17-A21, where [E22G]Aβ1−42 dimers showed significantly lower propensities than Aβ1−42 dimers \((P(4,2))\). Similarly, [E22G]Aβ1−42 dimers showed lower propensities than [E22G]Aβ1−40 dimers at R5-H6, Y10-V24, and N27-G29 but showed higher propensities at M35 and the CTR \((P(4,3))\).

**Hexamer-Hexamer Pairs.** Comparing the data of all hexamer pairs (the last 3 columns of graphs in Figure S4), β-strand propensities among pairs of the 4 peptides displayed similar tendencies as in dimers. Only Aβ1−40 hexamers were characterized by high β-strand propensities (0.4-0.6) at A2-F4. Aβ1−42 hexamers were characterized by higher β-strand propensities than Aβ1−40 hexamers at R5-E11, H13-F20, G25-S26, G29-I31, L34, and at the CTR \((P(6,5))\). Only Aβ1−42 and [E22G]Aβ1−42 hexamers had a non-zero propensity (0.2-0.3) at the CTR. The two most distinct features between the β-strand propensity curves were between Aβ1−40 and the two Arctic peptides, [E22G]Aβ1−40 \((P(7,5))\) and [E22G]Aβ1−42 \((P(8,5))\). Hexamers of both Arctic peptides were characterized by significantly higher β-strand propensities along the entire peptide sequence relative to hexamers of the wild-type peptides (except at A2-F4). The two most similar β-strand propensity curves were those belonging to Aβ1−42 and [E22G]Aβ1−42 hexamers \((P(8,6))\), where [E22G]Aβ1−42 had slightly higher β-strand propensities at R5-H6, E11-Q15, L17-F19, S26-K28, and M35-V36. The β-strand propensities of Aβ1−42 and [E22G]Aβ1−40 hexamers \((P(7,6))\) were more distinct than in dimers: [E22G]Aβ1−40 hexamers had higher β-strand propensities at R5-H6, Y10-H13, F20-V24, and N27-I31 but none at the CTR (in contrast to Aβ1−42 hexamers). [E22G]Aβ1−40 hexamers had higher propensities at G9-H13, A21-G25, N27, and A30-I31 relative to [E22G]Aβ1−42 hexamers, which were characterized by higher propensities at L17-F19, M35-V36, and at the CTR \((P(8,7))\).
**Dimer-Hexamer Pairs.** We finally compared pairs of β-strand versus residue curves of dimers and hexamers, including all 4 peptides (the last 4 rows and the first 4 columns of graphs in Figure S4). Each of the 4 diagonal graphs, $P(5,1)$, $P(6,2)$, $P(7,3)$, and $P(8,4)$, contained the β-strand propensities of dimers and hexamers of the same peptide. These diagonal graphs thus elucidated the differences in β-strand propensities due to different oligomer order. In Aβ$^{1-40}$, the β-strand propensities at A2-F4 increased substantially from ∼0.3 in dimers to ∼0.55 in hexamers ($P(5,1)$). In the rest of the sequence (except at A2-F4), β-strand propensities were decreased in hexamers relative to dimers ($P(5,1)$). This decrease of the overall β-strand propensity in hexamers relative to dimers was characteristic also of Aβ$^{1-42}$ ($P(6,2)$). In [E22G]Aβ$^{1-40}$, however, the changes in the β-strand propensity induced by higher oligomer order (dimers to hexamers) were minimal ($P(7,3)$). Interestingly, in [E22G]Aβ$^{1-42}$ a different tendency was observed: while β-strand propensities remained unchanged at the C-terminal region I31-A42, the values at R5-H6, H13-F20, N27-G29 were increased in hexamers relative to dimers ($P(8,4)$).

β-Strand propensities in Aβ$^{1-40}$ dimers were compared to those of hexamers of Aβ$^{1-42}$ ($P(6,1)$), [E22G]Aβ$^{1-40}$ ($P(7,1)$), and [E22G]Aβ$^{1-42}$ ($P(8,1)$). The overall β-strand propensities in Aβ$^{1-42}$ hexamers were similar to those in Aβ$^{1-40}$ dimers, where Aβ$^{1-40}$ dimers showed higher propensities at A2-F4, V12-H14, and L34-V36 and Aβ$^{1-42}$ hexamers displayed higher propensities at R5-E11, G25-S26, and the CTR ($P(6,1)$). Similar tendencies were observed when Aβ$^{1-40}$ dimers were compared to [E22G]Aβ$^{1-40}$ hexamers: [E22G]Aβ$^{1-40}$ hexamers had higher propensities at R5-V12 and A21-I31 ($P(7,1)$). [E22G]Aβ$^{1-42}$ had higher propensities at R5-E11, H14-F19, S26-G29, and M35-V40 compared to Aβ$^{1-40}$ dimers ($P(8,1)$).

We compared β-strand propensities in Aβ$^{1-42}$ dimers to those of hexamers of Aβ$^{1-40}$ ($P(5,2)$), [E22G]Aβ$^{1-40}$ ($P(7,2)$), and [E22G]Aβ$^{1-42}$ ($P(8,2)$). The largest differences between the β-strand versus residue curves were observed between Aβ$^{1-42}$ dimers and Aβ$^{1-40}$ hexamers, where the former were characterized by significantly higher propensities at H5-H14, V18-F20, A30-I31, L34-G37, and the CTR than the latter ($P(5,2)$). Overall amounts of the β-strand propensity were comparable between Aβ$^{1-40}$ dimers and [E22G]Aβ$^{1-40}$/[E22G]Aβ$^{1-42}$ hexamers ($P(7,2)$ and
β-Strand propensities in [E22G]Aβ₁₋₄₀ dimers were then compared to those of hexamers of Aβ₁₋₄₀ (P(5,3)), Aβ₁₋₄₂ (P(6,3)), and [E22G]Aβ₁₋₄₀ hexamers, which were characterized by a decreased propensities along most of the peptide (except at A2-F4), at R5-Q15 and L17-G37 (P(5,3)). The propensities of [E22G]Aβ₁₋₄₀ dimers and Aβ₁₋₄₀ hexamers were following similar trends as [E22G]Aβ₁₋₄₀ dimers showing consistently increased propensities at Y10-H14, F19-V24, and N27-G37 compared to Aβ₁₋₄₀ hexamers (P(6,3)). [E22G]Aβ₁₋₄₀ dimers and [E22G]Aβ₁₋₄₂ hexamers had comparable β-strand propensities with some variations along the sequence: [E22G]Aβ₁₋₄₂ hexamers had decreased propensities at Y10-E11, F20-V24, and A30-L34 as well as increased propensities at R5-H6, H13-F19, M35-V36, and the CTR relative to [E22G]Aβ₁₋₄₀ dimers (P(8,3)). Further, we compared β-strand propensities of [E22G]Aβ₁₋₄₂ dimers and hexamers of Aβ₁₋₄₀ (P(5,4)), Aβ₁₋₄₂ (P(6,4)), and [E22G]Aβ₁₋₄₀ (P(7,4)). [E22G]Aβ₁₋₄₂ dimers showed increased propensities at R5-V12, L17-F19, G25-S26, A30-I31, L34-V36, and the CTR as well as decreased propensities at A2-F4 (P(5,4)). β-Strand propensities of [E22G]Aβ₁₋₄₀ dimers and Aβ₁₋₄₂ hexamers were quite similar: Aβ₁₋₄₂ hexamers showed higher propensities at R5-H6, Q15-V18, and G25-K28 as well as lower propensities at A30-I31 and L34-V36 relative to [E22G]Aβ₁₋₄₂ dimers (P(6,4)). [E22G]Aβ₁₋₄₀ hexamers had higher propensities at the N-terminal two-thirds of the sequence at R5-H6, G9-H13, K16-L17, F20-V24, and N27-I31 as well as lower propensities at G33-V36 and the CTR (P(7,4))

Tertiary and quaternary structures of Aβ oligomers

To systematically examine the contact maps, we defined several regions enclosed in boxes on (a) intramolecular and (b) intermolecular contact maps with a high density of contacts. On each intramolecular (Figures S6a-h) and intermolecular (Figures S7a-h) contact map, we depicted 4 boxes marked by numbers (1-4). In the following we describe the differences between these maps,
referring to these boxes by their numbers. Note that the boxes were defined for each of the two Figures S6 and S7 and box definitions were valid only within each of the figures separately.

**Tertiary structures.** Tertiary structures of dimers and hexamers for all four peptides are depicted in Figures S6a-h. When compared among dimers and hexamers of all 4 peptides, these structures had some basic common features. They were rather similar to folded structures of the four individual peptides. This result suggests that oligomer formation may not be accompanied by major structural changes of individual peptides. In oligomers of all four peptides, the region with the largest density of intramolecular contacts was the central folding region, which included contacts between the CHC and MHR (Figures S6a-h, Box 4, lower left triangle). Similarly, intramolecular contacts at the CTR with the strongest contact V36–V39 were observed in all oligomers under study (Figures S6a-h, Box 3, lower left triangle). As expected, the intramolecular contacts around V36–V39 contact were strongest and more numerous in oligomers of Aβ1–42 and [E22G]Aβ1–42. The N-terminal E2–F4 region was involved, to different degrees, in intramolecular contacts with the CHC, MHR, and CTR in all 4 peptides (Figures S6a-h, Box 2, lower left triangle). These contacts varied among the four peptides and between dimers and hexamers of the same peptide.

Intramolecular contact maps of Aβ1–40 dimers and hexamers (Figures S6a and e) were indistinguishable from each other, demonstrating that tertiary structures of Aβ1–40 dimers and hexamers were identical. Box 1 includes all intramolecular contacts within the CHC (lower left triangle) as well as the hydrogen bond propensities (upper right triangle). Among dimers and hexamers of all 4 peptides, Aβ1–40 dimers and hexamers were the only oligomers characterized by an off–diagonal hydrogen bonding pattern indicating non-zero α-helix propensity in the CHC.

A hydrogen bonding pattern indicating non-zero α-helix propensity was observed in the region G25-G33, which contains 3 glycines in a GXXXGXXXG motif, of all oligomers of all 4 peptides (Figures S6a-h, Box 4, upper right triangle) and was stronger and comprised a wider region (G22-G33) in oligomers of the Arctic peptides (Figures S6c-d and g-h, Box 4, upper right triangle) with the GXXGXXXGXXXG motifs.

Aβ1–42 dimers were characterized by the strongest (Figure S6b, Box 1, lower left triangle) and
$\text{A}\beta_{1-42}$ hexamers by the weakest (Figure S6b, Box 1, lower left triangle) intramolecular contacts in the CHC. This result showed that $\text{A}\beta_{1-42}$ oligomerization from dimers to hexamers was accompanied by a partial loss of intramolecular contacts within the CHC. No such loss of contacts in the CHC upon oligomer formation from dimers to hexamers was observed for the other three peptides. Contacts of A2–F4 with the CHC, MHR, and CTR were stronger and more numerous in $\text{A}\beta_{1-42}$ dimers than in $\text{A}\beta_{1-42}$ hexamers (Figures S6b and f, Box 2, lower left triangle). The intramolecular contacts involving the MHR and CTR (Figures S6b and f, Box 3, lower left triangle) were somewhat stronger in $\text{A}\beta_{1-42}$ hexamers than in $\text{A}\beta_{1-42}$ dimers, possibly due to a significantly higher hydrogen bond propensity V36–V39 (Figure S6f, Box 3, upper right triangle), which might stabilize the tertiary structure at the CTR. This region was the only one that showed increased intramolecular contacts in $\text{A}\beta_{1-42}$ hexamers relative to $\text{A}\beta_{1-42}$ dimers. Among oligomers of all 4 peptides under study, $\text{A}\beta_{1-42}$ oligomers were characterized by the least strong intramolecular contacts within the central folding region, which included contacts between the CHC and MHR (Figures S6a-h, Box 4, lower left triangle). Interestingly, $\text{A}\beta_{1-42}$ dimers had somewhat lower hydrogen bond propensities in this region than $\text{A}\beta_{1-42}$ hexamers (Figures S6b and f, Box 4, upper right triangle). This result suggests that partial destabilization of intramolecular contacts within the central folding region might promote $\text{A}\beta_{1-42}$ hexamer formation.

$[\text{E22G}]\text{A}\beta_{1-40}$ hexamers had somewhat weaker contacts in the central folding region relative to $[\text{E22G}]\text{A}\beta_{1-40}$ dimers (Figures S6c and g, Box 4, lower left triangle), even though these contacts were stronger than in $\text{A}\beta_{1-42}$ hexamers (Figures S6f, Box 4, lower left triangle). The hydrogen bond pattern in $[\text{E22G}]\text{A}\beta_{1-40}$ dimers, which indicated a $\beta$-hairpin structure (Figures S6c, Box 4, upper right triangle), was not present in $[\text{E22G}]\text{A}\beta_{1-40}$ hexamers (Figures S6g, Box 4, upper right triangle). As in $\text{A}\beta_{1-42}$, this observation was consistent with a hypothesis that partial destabilization of intramolecular contacts within the central folding region is needed to form $[\text{E22G}]\text{A}\beta_{1-40}$ hexamers. Consistent with this hypothesis was also the observation that the intramolecular contacts in $[\text{E22G}]\text{A}\beta_{1-42}$ hexamers, which formed with lower probability just like $\text{A}\beta_{1-40}$ hexamers, were almost the same as in $[\text{E22G}]\text{A}\beta_{1-42}$ dimers (Figures S6d and h). Instead of destabilizing the
contacts in the central folding region upon hexamer formation (a feature characteristic in particular of Aβ₁₋₄₂ hexamers), these contacts were somewhat stronger in [E22G]Aβ₁₋₄₂ hexamers than in [E22G]Aβ₁₋₄₂ dimers (Figures S6d and h, Box 4, lower left triangle) and the hydrogen bonding propensities were somewhat increased as well (Figures S6d and h, Box 4, upper right triangle). In contrast, a relatively high hydrogen bond propensity at V36–V39 that was characteristic for Aβ₁₋₄₂ hexamers was present in [E22G]Aβ₁₋₄₂ dimers (Figure S6d, Box 3, upper right triangle). These results combined suggest that (i) stabilization of the CTR-MHR intramolecular contacts through formation of the hydrogen bond V36–V39 and (ii) destabilization of intramolecular contacts in the central folding region might be two key factors associated with oligomer formation.

**Quaternary structures.** Quaternary structures of dimers and hexamers for all four peptides are depicted in Figures S7a-h. These structures when compared among dimers and hexamers of all 4 peptides had some common basic features. We observed an overall increase in intermolecular contact strengths between dimers (Figures S7a-d) and hexamers (Figures S7e-h). This was a consequence of the fact that in a hexamer, where a peptide was in contact with five other equivalent peptides, the number of possible contacts between any two residues that belong to different peptides would be larger than in a dimer. Basic quaternary structure differences between Aβ₁₋₄₀ and Aβ₁₋₄₂ oligomers of order 5 (or higher) were described in prior work³,⁴ that demonstrated that Aβ₁₋₄₀ oligomer formation was driven by intermolecular interactions between the CHC region, while Aβ₁₋₄₂ oligomerization was dominated by intermolecular interactions of CTRs with CTR, MHR, and CHC regions of other peptides within an oligomer.

The intermolecular contact maps of Aβ₁₋₄₀, Aβ₁₋₄₂, [E22G]Aβ₁₋₄₀, and [E22G]Aβ₁₋₄₂ dimers (Figures S7a-d) were relatively similar to each other, suggesting that dimer formation was similar in all four peptides. Here we compared mostly the quaternary structures of Aβ₁₋₄₀, Aβ₁₋₄₂, [E22G]Aβ₁₋₄₀, and [E22G]Aβ₁₋₄₂ hexamers, which showed some distinct features (Figures S7e-h). The region A2–F4 was highly involved in hexamer formation in only Aβ₁₋₄₀ (Figures S7e-h, Box 1, lower left triangle) consistent with our data on β-strand propensity (Figures S4). Analysis of the hydrogen bonding propensities showed a pattern consistent with parallel intermolecular
\( \beta \)-strands at A2–F4 (Figure S7e, Box 1, upper right triangle). As shown previously,\(^3,^4\) \( \text{A}\beta_{1-40} \) oligomer formation was dominated by interactions among the CHC regions, between the CHC and MHR regions, and between the CHC and CTR regions (Figure S7e, Boxes 3 and 4, lower left triangle). These contacts were present also in intermolecular contact maps of \( \text{A}\beta_{1-42}, [\text{E}22\text{G}]\text{A}\beta_{1-40}, \) and \( [\text{E}22\text{G}]\text{A}\beta_{1-42} \) hexamers, but were mostly weaker, in particular in \( \text{A}\beta_{1-42} \) hexamers (Figures S7f-h, Boxes 3 and 4, lower left triangle). In \( \text{A}\beta_{1-42} \) hexamers, the strongest contacts were formed among the CTR regions, between the CTR and MHR regions, and between CTR and CHC regions (Figure S7f, Boxes 4 and 5, lower left triangle). Intermolecular contacts between A2–F4 and CHC, MHR, and CTR were present in \( \text{A}\beta_{1-42}, [\text{E}22\text{G}]\text{A}\beta_{1-40}, \) and \( [\text{E}22\text{G}]\text{A}\beta_{1-42} \) hexamers but were significantly weaker than in \( \text{A}\beta_{1-40} \) hexamers (Figures S7e-h, Box 2, lower left triangle). On the other hand, the peptide region A30-V40 (A30-A42) was characterized by significantly larger number of intermolecular contacts in \( \text{A}\beta_{1-42}, [\text{E}22\text{G}]\text{A}\beta_{1-40}, \) and \( [\text{E}22\text{G}]\text{A}\beta_{1-42} \) hexamers relative to \( \text{A}\beta_{1-40} \) hexamers. These contacts were strongest in \( \text{A}\beta_{1-42} \) hexamers. Intermolecular contact maps of \( [\text{E}22\text{G}]\text{A}\beta_{1-40}, \) and \( [\text{E}22\text{G}]\text{A}\beta_{1-42} \) hexamers were more similar to the contact maps of \( \text{A}\beta_{1-42} \) hexamers but showed more intermolecular contacts among CHC regions than \( \text{A}\beta_{1-42} \) hexamers, suggesting that the quaternary structure of the Arctic peptides had most features of \( \text{A}\beta_{1-42} \) hexamers but also some characteristics of \( \text{A}\beta_{1-40} \) hexamers.
Table 1: The chi–square test results for the probability that two histograms of oligomer size are equal. The p-values result from a comparison between the histograms of oligomer size of the same peptide obtained by averaging over time frames at (i) 19, 19.5, and $20 \times 10^6$ steps and (ii) 39, 39.5, and $40 \times 10^6$ steps. The p-values below the diagonal p-values in the table are a result of comparisons between the histograms of oligomer size of two different peptides obtained by averaging over time frames at 19, 19.5, and $20 \times 10^6$ steps. The p-values above the diagonal p-values in the table are a result of comparisons between the histograms of oligomer size of two different peptides obtained by averaging over time frames at 39, 39.5, and $40 \times 10^6$ steps.

<table>
<thead>
<tr>
<th>p-values</th>
<th>$\beta_{1-40}$</th>
<th>$\beta_{1-42}$</th>
<th>[E22G]$\beta_{1-40}$</th>
<th>[E22G]$\beta_{1-42}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{1-40}$</td>
<td>$1.3 \times 10^{-1}$</td>
<td>$1.5 \times 10^{-12}$</td>
<td>$7.6 \times 10^{-9}$</td>
<td>$2.1 \times 10^{-9}$</td>
</tr>
<tr>
<td>$\beta_{1-42}$</td>
<td>$6.6 \times 10^{-11}$</td>
<td>$2.2 \times 10^{-2}$</td>
<td>$2.2 \times 10^{-4}$</td>
<td>$1.1 \times 10^{-9}$</td>
</tr>
<tr>
<td>[E22G]$\beta_{1-40}$</td>
<td>$7.4 \times 10^{-7}$</td>
<td>$1.4 \times 10^{-3}$</td>
<td>$8.9 \times 10^{-3}$</td>
<td>$1.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>[E22G]$\beta_{1-42}$</td>
<td>$5.1 \times 10^{-10}$</td>
<td>$3.9 \times 10^{-7}$</td>
<td>$4.6 \times 10^{-6}$</td>
<td>$2.4 \times 10^{-2}$</td>
</tr>
</tbody>
</table>
**FIGURE CAPTIONS**

Fig. S1: Time evolution of probability distributions of oligomer sizes for (a) $\text{A}\beta_{1-40}$, (b) $\text{A}\beta_{1-42}$, (c) $\text[E22G]}\text{A}\beta_{1-40}$, and (d) $\text[E22G]}\text{A}\beta_{1-42}$. The scale on the x-axis was adjusted to the probability distribution of each peptide to best display the differences among the curves obtained at 4 different time windows: $t = 9 - 10 \times 10^6$ (black solid lines), $t = 19 - 20 \times 10^6$ (red solid lines), $t = 29 - 30 \times 10^6$ (black dashed lines), and $t = 39 - 40 \times 10^6$ (red dashed lines). The error bars correspond to SEM.

Fig. S2: Time progression of the p-values obtained by applying the chi–square test to quantify whether two histograms of oligomer sizes that were obtained at two different time windows, $\Delta t$ apart, were statistically equivalent for (a) $\Delta t = 1 \times 10^6$, (b) $\Delta t = 2 \times 10^6$, and (c) $\Delta t = 5 \times 10^6$. The p-values were calculated for each peptide $\text{A}\beta_{1-40}$ (solid black lines), $\text{A}\beta_{1-42}$ (solid red lines), $\text[E22G]}\text{A}\beta_{1-40}$ (dashed black lines), and $\text[E22G]}\text{A}\beta_{1-42}$ (dashed red lines), separately.

Fig. S3: Solvent accessible surface area (SASA) per residue for (a) $\text{A}\beta_{1-40}$ and (b) $\text{A}\beta_{1-42}$ dimers at $E_{CH} = 0$, $E_{CH} = 10^{-6}$, and $E_{CH} = 10^{-2}$. The error bars correspond to SEM.

Fig. S4: Pair plots of $\langle \beta\text{-strand} \rangle$ per residue for dimers and hexamers of all four peptides: $\text{A}\beta_{1-40}$, $\text{A}\beta_{1-42}$, $\text[E22G]}\text{A}\beta_{1-40}$, and $\text[E22G]}\text{A}\beta_{1-42}$. The error bars correspond to SEM. The solid curves correspond to the wild-type peptides and the dotted curves correspond to the Arctic peptides. The curves for the shorter peptides, $\text{A}\beta_{1-40}$ and $\text[E22G]}\text{A}\beta_{1-40}$, are plotted in black and the curves for $\text{A}\beta_{1-42}$ and $\text[E22G]}\text{A}\beta_{1-42}$ are plotted in red. Circles (open for $\text{A}\beta_{1-40}$ and $\text[E22G]}\text{A}\beta_{1-40}$ and filled for $\text{A}\beta_{1-42}$ and $\text[E22G]}\text{A}\beta_{1-42}$) correspond to dimers while squares (crossed for $\text{A}\beta_{1-40}$ and $\text[E22G]}\text{A}\beta_{1-40}$ and filled for $\text{A}\beta_{1-42}$ and $\text[E22G]}\text{A}\beta_{1-42}$) correspond to hexamers. Labels on the y-axes, $P(i,j)$, are $\beta$-strand propensities (with values $\in [0,1]$) for oligomer types i and j, where the types are defined by: 1- $\text{A}\beta_{1-40}$ dimers, 2-$\text{A}\beta_{1-42}$ dimers, 3-$\text[E22G]}\text{A}\beta_{1-40}$ dimers, 4-$\text[E22G]}\text{A}\beta_{1-42}$ dimers, 5-$\text{A}\beta_{1-40}$ hexamers, 6-$\text{A}\beta_{1-42}$ hexamers, 7-$\text[E22G]}\text{A}\beta_{1-40}$ hexamers, and 8-$\text[E22G]}\text{A}\beta_{1-42}$ hexamers.
Fig. S5: The average SASA per amino acid for (a) dimers and (b) hexamers of Aβ1−40 (solid black curves), Aβ1−42 (solid red curves), [E22G]Aβ1−40 (dotted black curves), and [E22G]Aβ1−42 (dotted red curves). The error bars correspond to SEM.

Fig. S6: Intra-molecular contact maps for (a-d) dimers and (e-h) hexamers of Aβ1−40 (a,e), Aβ1−42 (b,f), [E22G]Aβ1−40 (c,g), and [E22G]Aβ1−42 (d,h). The lower triangle contains the average number of contacts between two amino acids and the upper triangle contains the average number of hydrogen bonds for each pair of amino acids. The scale on the right shows the color mapping. The two types of maps have different scales, the scale on the left corresponds to the average number of contacts and the scale on the right corresponds to the average number of hydrogen bonds. The two thin diagonal lines are drawn through the diagonal elements of the two types of contact maps. The rectangular gray boxes with numbers mark regions of interest.

Fig. S7: Inter-molecular contact maps for (a-d) dimers and (e-h) hexamers of Aβ1−40 (a,e), Aβ1−42 (b,f), [E22G]Aβ1−40 (c,g), and [E22G]Aβ1−42 (d,h). The lower triangle contains the average number of contacts between two amino acids and the upper triangle contains the average number of hydrogen bonds for each pair of amino acids. The scale on the right shows the color mapping. The two types of maps have different scales, the scale on the left corresponds to the average number of contacts and the scale on the right corresponds to the average number of hydrogen bonds. The two thin diagonal lines are drawn through the diagonal elements of the two types of contact maps. The rectangular gray boxes with numbers mark regions of interest.
Fig. S2
Fig. S3
Fig. S5
Fig. S6

