

Cite this: *Chem. Sci.*, 2022, 13, 2649 All publication charges for this article have been paid for by the Royal Society of Chemistry

# Anatomy and formation mechanisms of early amyloid- $\beta$ oligomers with lateral branching: graph network analysis on large-scale simulations†

Miao Yuan, Xuan Tang and Wei Han \*

Oligomeric amyloid- $\beta$  aggregates (A $\beta$ O) effectively trigger Alzheimer's disease-related toxicity, generating great interest in understanding their structures and formation mechanisms. However, A $\beta$ O are heterogeneous and transient, making their structure and formation difficult to study. Here, we performed graph network analysis of tens of microsecond massive simulations of early amyloid- $\beta$  (A $\beta$ ) aggregations at near-atomic resolution to characterize A $\beta$ O structures with sizes up to 20-mers. We found that A $\beta$ O exhibit highly curvilinear, irregular shapes with occasional lateral branches, consistent with recent cryo-electron tomography experiments. We also found that A $\beta$ 40 oligomers were more likely to develop branches than A $\beta$ 42 oligomers, explaining an experimental observation that only A $\beta$ 40 was trapped in network-like aggregates and exhibited slower fibrillization kinetics. Moreover, A $\beta$ O architecture dissection revealed that their curvilinear appearance is related to the local packing geometries of neighboring peptides and that A $\beta$ 40's greater branching ability originates from specific C-terminal interactions at branching interfaces. Finally, we demonstrate that whether A $\beta$  oligomerization causes oligomers to elongate or to branch depends on the sizes and shapes of colliding aggregates. Collectively, this study provides bottom-up structural information for understanding early A $\beta$  aggregation and A $\beta$ O toxicity.

Received 15th November 2021  
Accepted 8th February 2022

DOI: 10.1039/d1sc06337e

rsc.li/chemical-science

## Introduction

The onset and development of Alzheimer's disease is widely believed to be linked with abnormal aggregation of amyloid- $\beta$  peptides (A $\beta$ ).<sup>1</sup> Aggregation of this peptide eventually leads to the formation of amyloid plaques in patients' brains and can also yield other intermediate aggregates, such as A $\beta$  oligomers (A $\beta$ O). Evidence suggests that A $\beta$ O have key roles in A $\beta$ 's neurotoxicity, which has shifted biomedical research focus from amyloid plaques to A $\beta$ O.<sup>2,3</sup> A $\beta$ s are comprised of various dynamically interconverting species differing in size, structure, biophysical properties, and toxicity,<sup>4,5</sup> with certain types of A $\beta$ O, such as those with high molecular weight (HMW) (typically MW > 50 kDa), being particularly toxic.<sup>5–8</sup> Knowledge about their structures and assembly pathways is crucial for understanding A $\beta$  oligomerization mechanisms, the molecular basis of A $\beta$ O toxicity, and its interplay with other aggregation diseases,<sup>9</sup> based on which aggregation inhibition strategies may be designed.<sup>10,11</sup> There are two major A $\beta$  alloforms that contain

40 and 42 amino acids, respectively. A $\beta$ 42, which has two more amino acids at the C-terminus, oligomerizes more easily and is more neurotoxic.<sup>12,13</sup> Comparing how the two alloforms oligomerize would further our understanding of A $\beta$ O pathology.

A $\beta$  oligomerization has been difficult to characterize due to A $\beta$ O metastability and heterogeneity. Several spectroscopic and imaging techniques have been applied<sup>14</sup> but provided only low-resolution structural information, such as secondary structure contents, overall morphologies and, in some cases, the time evolution of these properties with low<sup>8,15,16</sup> to medium resolutions.<sup>17,18</sup> Intermediate-size and HMW A $\beta$ O that rapidly formed during early aggregation stages were found to assume spherical or elongated, curvilinear shapes,<sup>19–24</sup> as well as annular structures.<sup>13,25–27</sup> Recently, three-dimensional imaging of elongated A $\beta$ O with high-resolution cryo-electron tomography (cryoET) revealed that the curvilinear aggregates were more irregular than previously thought and exhibited branched structures.<sup>28</sup> Despite these structural insights, it remains largely unclear how A $\beta$  peptides are spatially organized within A $\beta$ O, how they are packed against each other, and how these structural properties are realized through self-assembly. Addressing these questions is key to elucidating A $\beta$ O assembly mechanisms.

Molecular dynamic (MD) simulations are useful for deciphering the details of self-assembled supramolecular structures that are experimentally inaccessible.<sup>14,29,30</sup> Both atomic and coarse-grained (CG) simulations have been applied to study

State Key Laboratory of Chemical Oncogenomics, Guangdong Provincial Key Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, 518055, China. E-mail: hamw@pkusz.edu.cn

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1sc06337e



small A $\beta$ O<sub>s</sub> ranging from dimers to hexamers.<sup>14,31</sup> Recently, Zheng *et al.* employed a CG model and the umbrella sampling technique to determine the octamer formation free energy surface from which fibrillar assembly pathways were inferred.<sup>32</sup>

Computational investigation of larger assembly systems is significantly more challenging due to the rapidly growing number of degrees of freedom that need to be sampled and analyzed.<sup>33</sup> Only a few computational studies have employed CG models or atomic models in implicit solvent to probe the formation of A $\beta$ O<sub>s</sub> containing more than ten full-length A $\beta$  chains. These studies revealed the importance of hydrophobic and charged interactions for early oligomerization of A $\beta$ 40 and A $\beta$ 42,<sup>34,35</sup> and demonstrated that A $\beta$ O<sub>s</sub> with extended shapes have greater ability to aggregate than compact ones.<sup>36,37</sup> Large A $\beta$ O<sub>s</sub> of up to  $\sim$ 20 chains were occasionally observed, but a thorough examination of the structures and assembly pathways of large A $\beta$ O<sub>s</sub> has not been possible, probably due to insufficient sampling (*e.g.*, a simulation length of 2–5  $\mu$ s for assembly of each alloform) and the small number of chains (typically 20–30 chains) included in simulations, which does not considerably outnumber large A $\beta$ O size.<sup>14</sup>

In this study, we performed extensive simulations of A $\beta$ 40 and A $\beta$ 42 oligomerization. For each alloform, we followed the aggregation of 100 interacting chains using 30  $\mu$ s MD simulations with an efficient model that simplifies solvent representation but retains peptides' atomic details.<sup>38,39</sup> This model has proven to be reliable in predicting native structures of several small proteins,<sup>38,39</sup> determining the structures and assembly pathways of A $\beta$  dimers,<sup>40,41</sup> and elucidating the mechanisms of A $\beta$  fibril growth.<sup>42</sup> Moreover, graph networks<sup>43</sup> were adopted to represent an A $\beta$ O assembly hierarchy. This representation abstracts essential topological characteristics of nanostructured assembly, quantitatively describes an assembly hierarchy, and enables advanced graph theory to recognize structural patterns. It has been applied to self-assembly systems such as colloidal clusters, DNA-bricks, and metallic glasses to elucidate their assembly mechanisms.<sup>44–50</sup> Here, we further developed a network decomposition algorithm to facilitate the understanding of assembly hierarchy.

With these computational techniques, we investigated the anatomy of early A $\beta$ O<sub>s</sub> with sizes up to  $\sim$ 20-mers. We found that these A $\beta$ O<sub>s</sub> displayed curvilinear, irregular shapes with occasional lateral branches, consistent with A $\beta$ O morphologies reported by recent cryo-ET experiments.<sup>28</sup> A $\beta$ 40 oligomers were more likely to develop branches than A $\beta$ 42 oligomers, which may explain the tendency of A $\beta$ 40, but not A $\beta$ 42, to be trapped in network-like aggregates, as reported previously,<sup>20,25</sup> potentially explaining the slower fibrillization kinetics of A $\beta$ 40. Our graph network analysis further revealed the molecular basis of the observed A $\beta$ O morphologies and the difference in branching propensities between A $\beta$ 40 and A $\beta$ 42 oligomers. Finally, we found that A $\beta$  elongation and lateral branching are affected by the topologies and geometries of interacting aggregates. Taken together, our atomistic simulations provide bottom-up structural information for elucidating A $\beta$ O structures and the difference in early A $\beta$  aggregation of the two alloforms.

## Results

### Sampling a dynamic equilibrium of early metastable A $\beta$ oligomers using ten-microsecond simulations

To investigate early-stage A $\beta$ 40 and A $\beta$ 42 oligomerization, we performed simulations starting with 100 randomly dispersed and oriented peptide chains. The chain conformations were taken from monomeric state conformational ensembles (see Methods). For each alloform, we conducted two independent 15  $\mu$ s-long simulations to enhance sampling of A $\beta$ O<sub>s</sub> with large sizes. As shown in Fig. S1 in the ESI,<sup>†</sup> the A $\beta$ O size distribution ( $n$ ) evolved quickly during the early stage ( $t < \sim 2$   $\mu$ s) of simulations from predominantly monomeric states to larger oligomers. Similar oligomer size evolution has also been reported in previous CG and atomic simulations on a time scale of 0.5–1  $\mu$ s.<sup>34,36</sup> The size distributions continued to evolve and reached a steady state after  $t = 5$   $\mu$ s. At this stage, the largest A $\beta$ 40 and A $\beta$ 42 oligomers observed were 27-mers and 33-mers, respectively. Dynamic oligomerization events during  $t = 5$ –15  $\mu$ s were further analyzed. These results are presented in detail in the last part of the Results section; however, it should be noted that A $\beta$ O association and dissociation were frequently observed at steady state (Fig. S2<sup>†</sup>). The numbers of these association and dissociation events were roughly equal, indicating that the observed steady distribution in oligomer size is an outcome of dynamic equilibrium between oligomeric species rather than certain trapped oligomeric states.

Conformational dynamics of A $\beta$  during  $t = 5$ –15  $\mu$ s were also examined. A $\beta$  monomers exhibited a significant amount of turn structures (43–47%) but few helical (2–3%) or  $\beta$ -sheet (14–15%) structures, which agrees with previous CD measurements and atomic simulations.<sup>51,52</sup> Nonetheless, the conformational ensemble of A $\beta$  monomers covers dynamically interconverted conformations varying in secondary and tertiary structures (Fig. S3<sup>†</sup>), including both the ones lacking in any helical or sheet structures observed also in previous simulations,<sup>53</sup> and the partially folded ones with their helical and sheet contents up to  $\sim$ 15% and  $\sim$ 35%, respectively, some of which, such as multi-stranded conformations<sup>54</sup> or conformations with a partial helix around the central region,<sup>55</sup> have been reported experimentally.

Similar A $\beta$  peptide secondary structural contents were also observed in oligomeric states, and there was no significant increase in  $\beta$ -sheet content with oligomer size (Table S1<sup>†</sup>). Thus, our observations most likely corresponded to disordered intermediates at the very early stage of the aggregation process preceding any slow conformational conversion needed for fibril pathways.<sup>56</sup> The presence of such oligomeric intermediates has also been observed in recent microsecond atomistic simulations of A $\beta$  oligomerization with implicit solvent<sup>36,37</sup> and supported by experimental evidence for both small and large A $\beta$ O<sub>s</sub>.<sup>57,58</sup> Nonetheless, we did occasionally observed slow conformational change of A $\beta$ 's tertiary structures in oligomers based on our tertiary structural state analysis described in the ESI (Fig. S4<sup>†</sup>).<sup>59</sup> In this analysis, we identified for A $\beta$ 40 and A $\beta$ 42 several major tertiary structural states, each of which is



substantially different from the other in tertiary contact pattern. Based on how soon A $\beta$  peptides in oligomers evolved from its current tertiary structural state into a different one in the simulations, we employed a two-state Poisson process model (see the ESI†) to estimate the timescale of global reconfiguration of these peptides. The reconfiguration timescales for A $\beta$ 40 and A $\beta$ 42 were calculated to be  $\sim 66$  ms and  $\sim 39$  ms, respectively. This result suggested that the observed oligomers rearranged their structures very slowly and may even be off the pathways. It has been shown that high A $\beta$  concentrations retard fibrilization due to the rapid formation of off-pathway oligomers exhibiting high conversion barriers.<sup>62</sup> This finding is corroborated with our result given that A $\beta$  concentration in our simulations was rather high (3.9 mM). Regardless of whether these oligomers are on or off the pathways, their conversion timescales are far too long for our simulations.

### Both A $\beta$ 40 and A $\beta$ 42 form flexible elongated oligomers with branches, but A $\beta$ 42 oligomers appear more extended and less branched

With the data collected from  $t = 5\text{--}15$   $\mu\text{s}$  of the simulations, we first examined the overall shapes of A $\beta$ Os using topological networks. In these networks, nodes and edges indicate monomeric building blocks and their physical contacts, respectively.<sup>44,45,47,48,63</sup> This representation allows us to quantitatively describe the global topological features of supramolecular structures. For instance, the longest dimension of a topological network can be measured with the network diameter, defined as the length ( $l$ ) of the longest shortest path traversing the network. Fig. 1a and b show the diameter distributions  $P_n(l)$  with respect to oligomer size  $n$ . A wide oligomer diameter distribution was observed, especially for A $\beta$ Os with large sizes, indicating that A $\beta$ Os can assume distinct topologies varying in diameter. However, the average diameter  $\bar{l}$  of oligomeric topologies grew linearly with  $n$ , suggesting that A $\beta$ Os in general grow linearly and that there is an elongated backbone going across the entire oligomer.

We then analyzed A $\beta$ O backbone structures. An A $\beta$ O backbone is thought to be comprised of the monomers present on the longest shortest path, as well as those in direct contact with the path (see Methods). If monomers have a spherical shape, the elongated structure formed by the monomers on the path should accommodate exactly one monomer in its cross-section. When additional monomers are attached laterally to the elongated structure without increasing its length  $l$ , the cross-section of the structure would on average accommodate more monomers. Thus, if there are  $m$  monomers included in the backbone, the  $m/l$  ratio corresponds to the average number of monomers that can be accommodated in the cross-section of the backbone, thereby providing an estimate of its thickness. As shown in Fig. S5a,† this ratio was calculated to be  $\sim 1.3$  for both A $\beta$ 40 and A $\beta$ 42 oligomers and remained largely constant with oligomer size. The average A $\beta$ O thickness  $h$  was estimated according to  $h \approx 2R_g m/l$ , where  $R_g = \sim 1.0$  nm is the average radius of gyration of an A $\beta$  chain packed in oligomers. Our calculation indicated that A $\beta$ Os are  $\sim 2.6$  nm in thickness, which roughly agrees with

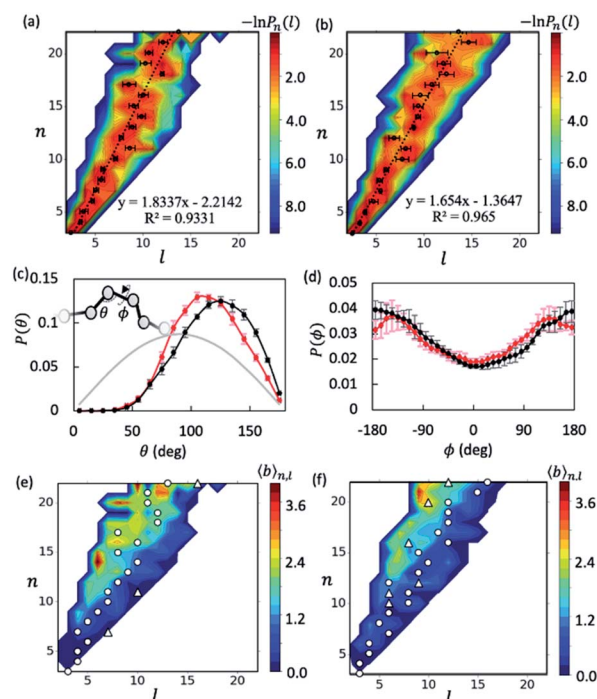


Fig. 1 Elongated, curvilinear topologies of A $\beta$ Os with lateral branching. Heat maps show the length distribution  $P_n(l)$  of A $\beta$ 40 (a) and A $\beta$ 42 (b) oligomers of a given size  $n$ . Circles denote the average  $\bar{l}$  for each  $n$ . Dashed lines show the linear correlation between  $\bar{l}$  and  $n$ . (c and d) Distributions of angles  $\theta$  (c) and dihedral angles  $\phi$  (d) formed by consecutive nodes along the longest shortest paths of A $\beta$ 40 (red dots) and A $\beta$ 42 (black dots) oligomers. The grey curve indicates a random angular distribution. (e and f) Average branch counts  $b(n, l)$  in A $\beta$ 40 (e) and A $\beta$ 42 (f) oligomers as a function of  $l$  and  $n$ . Circles and triangles denote the first and, if any, second peaks in the length distribution  $P_n(l)$  for a given  $n$ . Errors reported hereafter were estimated by dividing the two trajectories ( $t \in [5 \mu\text{s}, 15 \mu\text{s}]$ ) into four blocks for each case unless stated otherwise.

the diameter value ( $\sim 2.7$  nm) reported in a recent cryo-ET study of large A $\beta$ 42 oligomers.<sup>28</sup>

We also examined A $\beta$ O backbone geometries through conformational analysis of the longest shortest paths. These paths can be simplified as a string of beads with each bead representing the center of mass of each monomer. The A $\beta$ O backbone conformations are thus defined by all three-bead angles  $\theta$  and four-bead dihedral angles  $\phi$  along the longest shortest paths. As shown in Fig. 1c, both A $\beta$ 40 and A $\beta$ 42 oligomers avoided a sharp  $\theta$  angle ( $< 60^\circ$ ) in their backbone parts, probably due to the exclusion of self-volume. The dihedral angle  $\phi$  distributions (Fig. 1d) were flat with a 60% : 40% ratio of *anti*-conformations ( $|\phi| > 90^\circ$ ) to *syn*-conformations ( $|\phi| < 90^\circ$ ). Thus, the A $\beta$ O backbone conformations observed here largely resemble those of a freely rotating chain that has a large rotational isomerism and high conformational flexibility. This may explain why A $\beta$ Os have been reported to be highly curvilinear and irregular.<sup>28</sup> Despite the flexibility of both types of A $\beta$ Os, we found that A $\beta$ 42 oligomers were more extended than A $\beta$ 40 oligomers, as indicated by a larger Flory's characteristic ratio  $C_\infty$  of A $\beta$ 42 oligomers ( $3.01 \pm 0.21$  versus  $1.95 \pm 0.23$ ) (see the



ESI and Fig. S5b†). For a freely rotating chain, its characteristic ratio can be related to the local angle  $\theta$  by  $C_\infty = (1 - \cos \theta)/(1 + \cos \theta)$ . Based on this relationship, we predicted that the  $C_\infty$  of A $\beta$ 42 oligomers is indeed larger than that of A $\beta$ 40 (3.00 *versus* 2.28), since on average A $\beta$ 42 oligomers have a wider  $\theta$  angle ( $\sim 120^\circ$  *versus*  $\sim 113^\circ$ ). This prediction is largely in accord with our  $C_\infty$  calculation.

Inspecting representative oligomeric structures revealed further A $\beta$ O branching structure, especially in A $\beta$ O with large size ( $n > \sim 10$ ) (Fig. 2). Intriguingly, branched morphologies of large A $\beta$ O have recently been reported in several studies using advanced imaging techniques such as high-resolution AFM, solution TEM, and cryo-ET.<sup>20,25,28</sup> We evaluated the degree of branching in A $\beta$ O by counting the number of branches ( $b$ ) present in these oligomers. This quantity was calculated by finding disconnected components after the entire backbone part was removed from the topological networks.

Fig. 1e and f show plots of the average number of branches  $\bar{b}(n, l)$  present in A $\beta$ O as a function of oligomer size  $n$  and backbone length  $l$ . As expected, A $\beta$ O with larger size and shorter backbones are more likely to have branches. With a given oligomer size and a given backbone length, A $\beta$ 40 oligomers contain more branches than A $\beta$ 42 oligomers on average

(Fig. 1e and f). This difference in branch structure is further exemplified by comparing representative A $\beta$ 40 and A $\beta$ 42 oligomer structures with select  $n$  and  $l$  values (Fig. 2a and c *versus* 2b and d). On average, there are  $0.6 \pm 0.1$  and  $0.4 \pm 0.1$  branches in mid-sized A $\beta$ O ( $n = 11-15$ ) formed by A $\beta$ 40 and A $\beta$ 42, respectively. In larger A $\beta$ O ( $n = 16-22$ ), A $\beta$ 40 oligomers developed twice as many branches as A $\beta$ 42 oligomers ( $1.4 \pm 0.1$  *versus*  $0.7 \pm 0.2$  branches). Thus, our results suggest that A $\beta$ 40 has a greater tendency to form branched structures than A $\beta$ 42 when assembled into large oligomers.

### Observed A $\beta$ oligomer mass distributions and cross collision-section areas agree with previous experimental results

To assess the relevance of our A $\beta$ O observations to those reported experimentally, we first calculated the mass distribution of A $\beta$ O that have been extensively characterized in experiments. Nevertheless, our simulations were conducted at 3.9 mM, a concentration much higher than normal experimental condition for characterizing the A $\beta$ O mass distribution. Following Kindt *et al.*,<sup>64,65</sup> the oligomer particle densities  $\rho_n$  derived from the simulations were used to parameterize a model of non-interacting aggregates with which one can obtain the A $\beta$ O partition functions  $q_n^0$  (Fig. S6a†) and thereby determine their mass distributions at any given peptide concentration. We paid attention to the oligomer mass distribution at two particular concentrations, a lower one at 30  $\mu$ M and a higher one at 200  $\mu$ M. At both concentrations, the mass distribution of A $\beta$ O has been probed experimentally.<sup>57,66</sup> At 30  $\mu$ M, A $\beta$ 40 and A $\beta$ 42 monomers, dimers, and trimers were predicted to be the major species, and their probabilities were close to what has been reported at the same concentration in ion-mobility mass-spectroscopy (IM-MS) experiments (Table 1).<sup>57</sup> At 200  $\mu$ M, oligomers up to 15-mer were predicted to exist for both alloforms with a nonnegligible probability ( $>0.01$ ), which agrees with another IM-MS experiment conducted at the same peptide concentration.<sup>66</sup> At the two concentrations, A $\beta$ 42 11-mers to 15-mers were two to four times more probable than those of A $\beta$ 40, also in accord with previous experimental observation.<sup>67</sup> For A $\beta$ O of even larger sizes ( $n \geq 15$ ), there was a linear correlation between  $\ln q_0(n)$  and  $n$  (Fig. S6b†), indicating the existence of a critical oligomer concentration (COC) for early metastable A $\beta$ O.<sup>24</sup>

Of note, the COCs of A $\beta$ 40 and A $\beta$ 42 were predicted based on our statistical mechanical model to be about 500  $\mu$ M and 300  $\mu$ M, respectively. We also calculated the COC using reported average monomer density from previous large scale simulations (*cf.* Fig. 1 in ref. 34 and Fig. 3a in ref. 36). Our COCs are comparable to the ones ( $\text{COC}_{40} = \sim 370$   $\mu$ M,  $\text{COC}_{42} = \sim 210$   $\mu$ M) derived based on the study of Urbanc *et al.*<sup>34</sup> using CG models but larger than those ( $\text{COC}_{40} = \sim 54$   $\mu$ M,  $\text{COC}_{42} = \sim 46$   $\mu$ M) based on the study of Barz *et al.*<sup>36</sup> using atomistic peptide models in implicit solvent. Several experimental studies using different characterization techniques reported 18–50  $\mu$ M COCs for A $\beta$ 40.<sup>68–70</sup> Hence, the COC of A $\beta$ 40 can be reproduced with the atomic simulations in implicit solvent but appeared somewhat overestimated in the simulations by Urbanc *et al.* and by

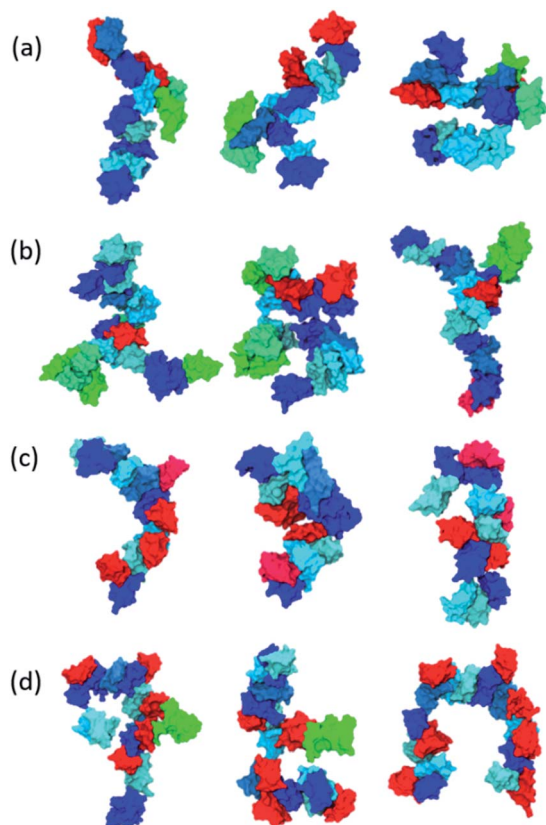


Fig. 2 Representative structures of A $\beta$ 40 (a and b) and A $\beta$ 42 (c and d) oligomers with  $n = 13$  and  $l = 9$  (a and c) or with  $n = 19$  and  $l = 12$  (b and d). All peptides are shown in surface representations. The peptides on the longest shortest paths are shown in blue with varying darkness and the other peptides in the backbones are shown in red. The peptides in the branches are shown in green.



Table 1 A $\beta$ O mass distributions reported experimentally and calculated in this study

<i>n</i>	A $\beta$ 40			A $\beta$ 42		
	30 $\mu$ M <sup>a</sup>	200 $\mu$ M	Expt <sup>b</sup>	30 $\mu$ M	200 $\mu$ M	Expt <sup>b</sup>
1	0.71	0.11	~0.61	0.60	$8.2 \times 10^{-2}$	~0.59
2	0.25	0.29	~0.19	0.36	0.37	~0.22
3	$3.4 \times 10^{-2}$	0.30	~0.11	$2.9 \times 10^{-2}$	0.22	~0.11
4	$1.9 \times 10^{-3}$	0.13	$\sim 4.0 \times 10^{-2}$	$1.9 \times 10^{-3}$	0.11	$\sim 4.2 \times 10^{-2}$
5–10	$1.3 \times 10^{-4}$	0.17	—	$1.7 \times 10^{-4}$	0.14	—
11–15	$6.9 \times 10^{-11}$	$1.1 \times 10^{-2}$	—	$2.2 \times 10^{-10}$	$6.3 \times 10^{-2}$	—
>15	$5.8 \times 10^{-14}$	$5.4 \times 10^{-4}$	—	$2.4 \times 10^{-14}$	$1.2 \times 10^{-2}$	—

<sup>a</sup> Monomer concentration at which the mass distributions were calculated. <sup>b</sup> Obtained from the IM-MS experiment carried out at 30  $\mu$ M peptide concentration.<sup>57</sup>

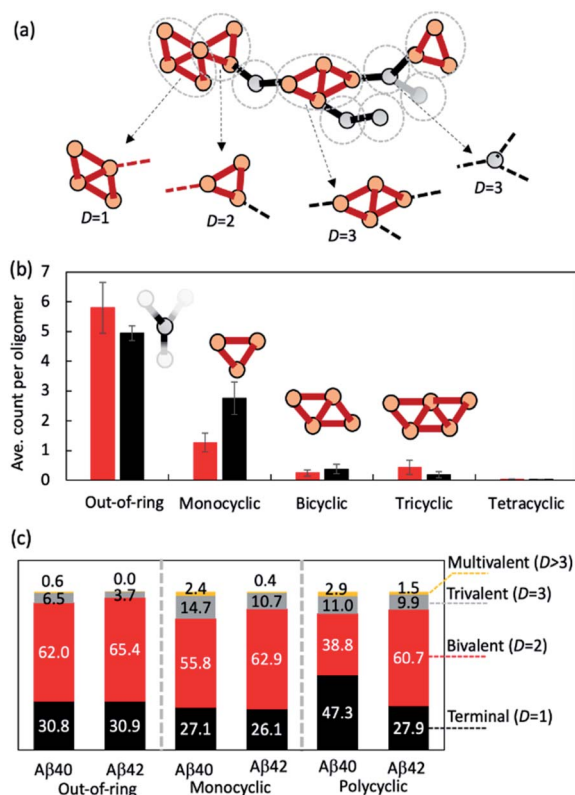


Fig. 3 Oligomeric topologies can be decomposed into linear and cyclic fragments organized in an elongated fashion. (a) Illustration of the topological decomposition procedure. Orange and grey circles denote nodes in rings or outside of rings, respectively. Red edges belong to rings and black ones do not. Grey dashed circles indicate how topologies were partitioned. (b) Average counts of each type of basic building fragments observed in A $\beta$ 40 (red bars) and A $\beta$ 42 (black bars) fragments (for  $n > 10$ ). (c) Percentage distributions of each type of basic structural fragment coordination number *D*.

us. The critical concentration of some prefibrillar aggregates of A $\beta$ 42 was determined in a more recent study to be 90 nM,<sup>71</sup> significantly lower than the COC measured for A $\beta$ 40 or derived based on any of the above simulations. As such, these A $\beta$ 42 aggregates were highly stable. The thermodynamic stability of protein aggregates is governed by their structures. It has been suggested<sup>24</sup> that for aggregation-prone proteins including A $\beta$ ,

$\beta$  sheet-rich aggregates such as fibrils assembled through strong intermolecular hydrogen bonding (HB) are normally more stable and have a lower critical concentration than early aggregates induced by amphiphilic character of polypeptide chains.<sup>72</sup> As the stability of the A $\beta$ 42 prefibrillar aggregates probed in that study even rivals that of A $\beta$ 40/42 fibrillar aggregates whose critical concentrations are 100–200 nM,<sup>73,74</sup> these aggregates may already develop a considerable  $\beta$ -sheet character, stabilized largely by intermolecular HBs. Conversely, the A $\beta$  oligomers observed here, which have not undergone conformational conversion yet, exhibit low  $\beta$ -sheet contents (14–15%), and their incorporation of each new monomer brings about only  $\sim 2.5$  additional interchain HBs. This structural difference between the aggregates probed in the experiment study and our simulations may explain the large difference between the measured and simulated critical concentrations of A $\beta$ 42.

We also examined the collision cross-section (CCS) values of low order oligomers, a quantity related to the overall A $\beta$ O shapes that has been carefully characterized in IM-MS experiments.<sup>57,66,67</sup> As described in Methods, the oligomer structures taken from solution simulations were relaxed in gas phase and their protonation states were properly adjusted to be consistent with experimental conditions reported by Bernstein *et al.*<sup>67</sup> The average CCS values agreed reasonably well with the experimental data, with a better agreement observed for A $\beta$ 40 than for A $\beta$ 42 (Table S2†). Compared to the data reported by Berstein *et al.*, our A $\beta$ 42 tetramer CCS value was underestimated by  $\sim 200$   $\text{\AA}^2$ , but agreed better with the CCS value reported more recently by Zhang *et al.* (2100  $\text{\AA}^2$  versus 2172  $\text{\AA}^2$ ).<sup>75</sup> Despite the deviation from these experimental data, our calculations revealed that the CCS values of A $\beta$ 42 dimers to hexamers are always greater than those of A $\beta$ 40, suggesting that low order A $\beta$ 42 oligomers tend to be more “open” than those of A $\beta$ 40. This trend is consistent with our characteristic ratio calculations, as well as previous experiments and atomistic simulations.<sup>36,67</sup>

### Early A $\beta$ oligomer internal topological networks consist of a mixture of line and small cyclic fragments organized in favor of elongated morphologies

Having validated our observations of A $\beta$ O, we proceeded to analyze molecular arrangements within the oligomers. In the



graph network representation framework each network represents a possible arrangement of monomers within an oligomer. The molecular arrangement pattern could be identified from the most representative networks. Using a network clustering algorithm that considers particle labelling degeneracy (see Methods),<sup>44,45,63</sup> we examined the nonisomorphic arrangements of A $\beta$ O<sub>n</sub>s sampled in our simulations (Fig. S7†).

There are two theoretically possible trimer arrangements, a linear arrangement and a triangular one, and we observed both with similar chance (Fig. S7†). We also sampled all six and 18 possible arrangements of monomers in tetramers and pentamers, respectively, but only three tetrameric arrangements and five pentameric arrangements were populated with a combined probability >0.9, indicating that only certain types of molecular arrangements are preferred in these oligomers. For larger A $\beta$ O<sub>n</sub>s, the number of preferred arrangements grows exponentially with oligomer size, but these arrangements account for only a small fraction (e.g.,  $5-6 \times 10^{-2}$  for  $n = 6$  and  $2-4 \times 10^{-6}$  for  $n = 10$ ) of all those theoretically allowed (Fig. S8†),<sup>44</sup> indicating that large oligomers formed by A $\beta$ 40 and A $\beta$ 42 can only access a tiny fraction of the theoretical topological space.

Despite their topological heterogeneity, our inspection of representative topological networks indicated that A $\beta$ O<sub>n</sub>s exhibit topological networks that appear to be made of structural elements including line segments and triangles (Fig. S7†). The triangular rings are either separated in the networks or share common nodes or edges. Bigger rings containing more than three nodes were also observed, but their probability of occurrence decays fast as ring size grows (Fig. S9†).

We then investigated how the line segments and triangle rings are organized to form elongated oligomeric morphologies. To this end, we developed an algorithm that decomposes the topological networks into basic building fragments. This algorithm is described in the Methods and illustrated in Fig. 3a. It identifies three building fragment types, namely nodes that do not belong to any ring structures (thus termed ‘out-of-ring nodes’), monocyclic fragments that do not share any edge with other rings, and fused polycyclic fragments whose cyclic members have common edges with one another.

Fig. 3b shows the average count of each type of building fragment present in large A $\beta$ 40 or A $\beta$ 42 oligomers ( $n > 10$ ). For both alloforms, the out-of-ring nodes are the most frequently observed fragment types (with 5–6 such nodes per oligomer). The chance of observing a monocyclic fragment is 1/3–1/2 of that of out-of-ring nodes. The probability of observing fused polycyclic fragments is further reduced. There is a noticeable difference in fragment compositions of A $\beta$ 40 and A $\beta$ 42 oligomers. There are ~20% more out-of-ring nodes in A $\beta$ 40 oligomers than in A $\beta$ 42 oligomers, but A $\beta$ 42 oligomers harbor twice as many monocyclic fragments as A $\beta$ 40 oligomers. This suggests that although both A $\beta$ 40 and A $\beta$ 42 can form oligomers with similar elongated shapes, the underlying pattern of molecular arrangement is not the same.

Because topological features of supramolecular assemblies are linked closely with the coordination behaviors of underlying building blocks,<sup>46,47,49</sup> we examined the coordination properties of the building fragments identified above. A key coordination

property is the coordination number  $D$ , defined here as the number of linkages between a fragment and the remaining part of the network (see Methods). Building fragments with a single linkage ( $D = 1$ ) form network termini. Those with a bivalent coordination ( $D = 2$ ) can only elongate the networks and, therefore, are necessary for elongated topology formation. Fragments with a trivalent coordination ( $D = 3$ ) will cause the network to bifurcate, and those with higher coordination numbers ( $D > 3$ ) could form even more complex network structures. Therefore, an elongated network must be composed predominantly of monovalent and bivalent fragments.

Fig. 3c compares the coordination propensities of the three building fragment types. In general, all three types are most likely coordinated with two neighbours and have a rather low chance (<~15%) of being coordinated with more than two. Out-of-ring nodes are less likely to interact with more than two neighbours compared with the other two types. The prevalence of bivalent fragments in the A $\beta$ O<sub>n</sub> topological networks explains why A $\beta$ O<sub>n</sub>s normally prefer elongated topologies. Of note, A $\beta$ 40 always has a greater tendency to form fragments with large coordination numbers ( $D \geq 3$ ) compared with A $\beta$ 42 (Fig. 3c), which corroborates our finding that A $\beta$ 40 oligomers are more likely to branch than A $\beta$ 42 oligomers.

### Greater A $\beta$ 40 oligomer branching propensity is attributed to the ability of A $\beta$ 40 to form strong C-terminal hydrophobic interactions at branching interfaces

To further gain insights into A $\beta$ O<sub>n</sub> branching propensity, we focused our analysis on where these oligomers branched. Because a branch is an extension in different directions from an oligomer backbone part, the nature of branch and backbone interfaces is key to understanding A $\beta$ O<sub>n</sub> branching propensity. We thus analyzed branching interfaces in the topological networks. Our results showed that these interfaces were formed mainly through interactions between a single node from the branch and an out-of-ring node from the topological network backbone or a backbone node that belongs only to a single ring (Fig. 4a). The former type of interface is more favored by A $\beta$ 40 oligomers (~40%) and the latter one is more favored by A $\beta$ 42 oligomers (~50%). In addition, about 20% of branching interfaces in A $\beta$ 42 oligomers consist of a backbone node and two branch nodes in a triangular arrangement.

To understand why A $\beta$ 40 oligomers exhibit a greater branching propensity, we analyzed the inter-chain residual contacts formed at branching interfaces. The resulting contact probability maps (Fig. 4b and c) showed that at A $\beta$ 40 oligomer branching points, the backbone peptide mainly used its L34–G37 region to contact the I31–L34 region of its neighbor. The probabilities of these contacts were mostly >0.2. In contrast, interfaces at A $\beta$ 42 oligomer branching points were formed between the G37–V40 and V12–H14 regions of the peptides from the backbone and the branch, respectively, with a lower probability (0.1–0.15). This result suggests that the greater tendency of A $\beta$ 40 oligomers to branch is likely attributed to the ability of A $\beta$ 40 to form stronger C-terminal hydrophobic interactions at branching interfaces.



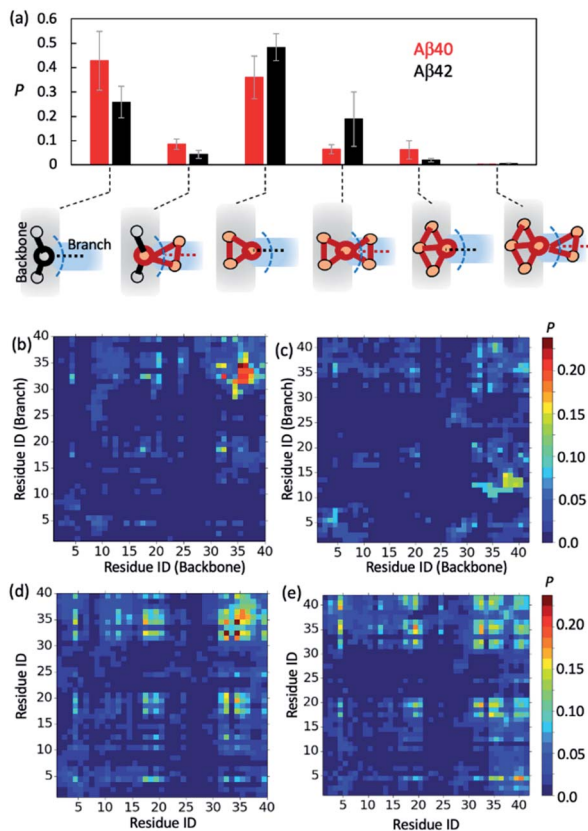


Fig. 4 A $\beta$ 40 and A $\beta$ 42 oligomers exhibit different branching interfaces. (a) Probabilities of different branching interface types identified in large oligomers ( $n > 10$ ) of A $\beta$ 40 (red bars) and A $\beta$ 42 (black bars). Blue dashed curves indicate the dividing surfaces between backbones and branches. (b and c) Probability maps of residual contacts formed between the backbones and branches at A $\beta$ 40 (b) and A $\beta$ 42 (c) oligomer branching interfaces. (d and e) Average probabilities of residual contacts between peptide chains observed in A $\beta$ 40 (d) and A $\beta$ 42 (e) oligomers.

Regardless of their locations in oligomers, both A $\beta$ 40 and A $\beta$ 42 can form extensive contacts with their neighbors (Fig. 4d and e), mainly using their central region L17–F21 and C-terminal region I31–V39/I41. Previous nuclear magnetic resonance (NMR) studies as well as simulations also revealed that the similar regions of both A $\beta$  alloforms participated in the initial docking of monomers to early A $\beta$  aggregates.<sup>41,76–78</sup> A different docking region, namely F19–N27, was recently reported by Brender *et al.*,<sup>79</sup> but it was found to recognize fibrils that are structurally different from A $\beta$ Os observed here. In addition, for both alloforms, the aromatic residues such as F4, F19 and F20 were involved heavily in hydrophobic interactions but direct intermolecular contacts between aromatic residues were infrequent (with an average chance of  $<0.15$ , Fig. S10<sup>†</sup>), suggesting that the aromatic  $\pi$ – $\pi$  interactions might not be crucial to early oligomer formation, as opposed to their reported role in fibril formation.<sup>80</sup> Overall, the average intermolecular contact patterns of A $\beta$ 40 and A $\beta$ 42 are markedly different from the ones observed at branching interfaces. On average A $\beta$ 42 can form more interchain residual contacts than A $\beta$ 40 (24.8

contacts *versus* 23.2 contacts), largely due to additional interactions associated with V40–A42 of A $\beta$ 42 (Fig. S10<sup>†</sup>). This is opposite to the observed tendencies of the two alloforms to form interchain interactions at branching interfaces but in line with the general notion that A $\beta$ 42 has a greater ability to form intermolecular interactions due to its enhanced hydrophobicity by I41 and A42. Taken together, the above results indicate that the difference in branching contact patterns between A $\beta$ 40 and A $\beta$ 42 cannot be explained based only on their average contact propensity.

As a polypeptide segment involved too much with intramolecular contacts can be hindered from intermolecular contacts, we thus examined intramolecular tertiary contacts in A $\beta$ 40 and A $\beta$ 42 to seek clues to the different contact propensity at branching interfaces between the two alloforms. In particular, the tertiary contacts involving the C-terminal region of A $\beta$  were carefully analyzed since A $\beta$ 40, but not A $\beta$ 42, can use this region to form extensive hydrophobic interactions at branching interfaces. As shown in Fig. S11a and b,<sup>†</sup> A $\beta$ 42 is more likely to form a C-terminal reversed loop than A $\beta$ 40. This loop is centered at V36–G37 and promoted by intramolecular interactions involving V39–I41, a structural feature of A $\beta$ 42 monomers reported previously in NMR and MD studies.<sup>81</sup> At branching interfaces, A $\beta$ 42 was observed to lose more intramolecular contacts in its C-terminal region than A $\beta$ 40 ( $-2.3$  *versus*  $-0.9$  residual contacts) when it needs to use this region to form substantial intermolecular contacts ( $\geq 5.0$  residual contacts) (Fig. S11c and d<sup>†</sup>). In other words, A $\beta$ 42 needs to disrupt its C-terminal tertiary structure to a greater extent to form C-terminal intermolecular interactions at branching interfaces. Therefore, the stronger tendency of A $\beta$ 40 to form interchain C-terminal interactions at branching interfaces is a special coordination property of A $\beta$ 40, probably due to its lesser tendency to form loop or turn structures at its C-terminus.

### Oligomer elongation and branching are affected by their size and shape

Finally, we sought to understand how A $\beta$ Os elongated themselves or developed branches. We first monitored the association events that took place during the simulation period  $t \in [5 \mu\text{s}, 15 \mu\text{s}]$  (see Method). There were  $\sim 2.5 \times 10^4$  total association events each for A $\beta$ 40 and A $\beta$ 42, during most of which the resulting initial binding complexes quickly dissociated within  $\sim 1$ – $2$  ns and failed to generate metastable A $\beta$ Os. Only a fraction ( $\sim 0.04$ ) of the events led to the formation of firmly bound complexes. The initial binding was stabilized mainly by electrostatic interactions (0.9–1.0 salt bridges) involving A $\beta$ 's N-terminal regions and, to a lesser extent, by non-specific hydrophobic interactions (0.3–0.4 contacts). In most cases ( $>0.97$ ), there was only a single A $\beta$  from each colliding aggregate that participated in the binding interfaces. Hence, the ring structures in topological networks did not arise from the association events but instead were formed through the internal structural rearrangement of A $\beta$ Os.

We then monitored the change in the backbone length ( $l$ ) and the number of branches ( $b$ ) in topological networks after



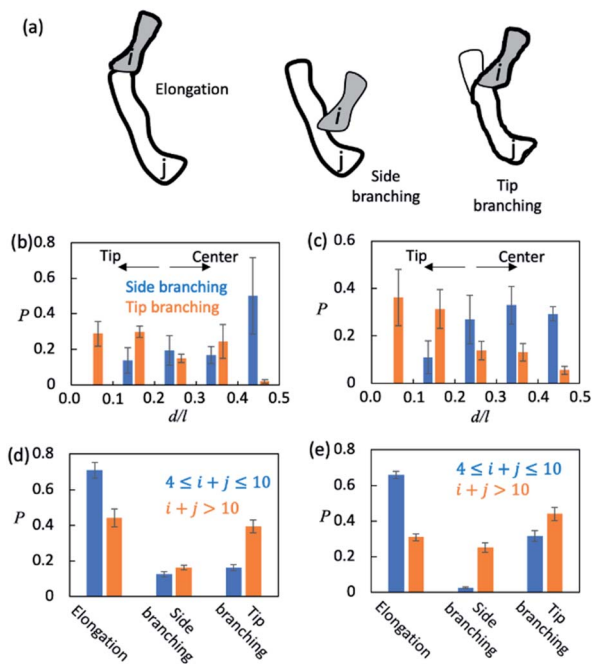


Fig. 5 A $\beta$ O elongation and lateral branching mechanisms. (a) Schematic representation of A $\beta$ O elongation and side/tip branching observed in the present study. The backbone parts are marked with thick outlines. (b and c) Positional preference of branching points observed during side and tip branching of A $\beta$ 40 (b) and A $\beta$ 42 (c) oligomers. The position of a branch point is measured as  $d/l$  where  $l$  is the length of the backbone and  $d$  is the length of the shortest path connecting the branch point and the closer end of the backbone. (d and e) Elongation and side/tip branching occurrence frequencies for small ( $4 \leq n \leq 10$ ) and large ( $n > 10$ ) oligomers formed by A $\beta$ 40 (d) and A $\beta$ 42 (e).

each association event (see the ESI $^\dagger$ ). An increase in  $l$  indicates elongation of oligomers and an increase in  $b$  indicates the formation of new side branches (Fig. 5a). Intriguingly, we also observed events in which both  $l$  and  $b$  increased. These events involved the attachment of a monomer/oligomer to a site near a second oligomer's backbone tips, unlike the side branching events in which the attachment point was close to the center of the second oligomer's backbone (Fig. 5b and c). As these events effectively caused the newly formed oligomers to bifurcate at its tip, they were henceforth termed tip branching.

We counted the elongation and side/tip branching occurrence frequencies. As shown in Fig. 5d and e, A $\beta$ O's with relatively small sizes ( $4 \leq n \leq 10$ ) were mainly formed *via* elongation at a chance of 60–70%. On the other hand, larger A $\beta$ O's ( $n > 10$ ) were formed more often *via* side or tip branching. This observation is expected since A $\beta$ O's exhibit a linear topology that can only elongate at its tips, but new branches can be developed from any other A $\beta$ O part. In addition, tip branching is the preferred branching mode for both A $\beta$ 40 and A $\beta$ 42 oligomers, especially when their sizes are large ( $n > 10$ ). We attributed this preference for tip branching mode to the curvilinear shapes of A $\beta$ O's whose center parts are likely less accessible, a disadvantage to side branching. Moreover, the ratios of tip branching events to side branching events were

$\sim 2.5$  and  $\sim 1.8$  for A $\beta$ 40 and A $\beta$ 42 oligomers, respectively, indicating that side branching is even less favored by A $\beta$ 40 oligomers. This finding is consistent with our characteristic ratio calculation which showed that A $\beta$ 40 oligomer shapes are more curvilinear and compact (Fig. S5b $^\dagger$ ). Taken together, our results suggest that both A $\beta$ O elongation and branching can be affected by oligomeric topologies and geometries.

## Discussion

In summary, we used computer simulations to reveal the anatomy of early A $\beta$  oligomer (A $\beta$ O) aggregates. These A $\beta$ O's had curvilinear morphologies, consistent with experimental reports.<sup>28</sup> Of particular interest is our observation that A $\beta$ O's could occasionally develop branched structures, with larger sized oligomers (*e.g.*,  $n > 15$ ) having a greater branching tendency. There have been several reports of oligomer branching for both A $\beta$ 40 and A $\beta$ 42 using imaging techniques such as AFM<sup>82</sup> and, more recently, liquid-state TEM,<sup>20</sup> as well as high-resolution cryo-ET.<sup>28</sup> Especially considering their curvilinear, irregular appearance and assembly thickness, the oligomeric species observed here most likely represent one type of early A $\beta$ O that was reported in the cryo-ET study.<sup>28</sup> Although previous atomistic and CG simulations of systems containing 20–30 full-length A $\beta$  chains permitted for the observation of large oligomers ( $n \approx 20$ ) with linear topologies,<sup>34,36</sup> our simulations of 100 chain-containing systems allowed us to observe A $\beta$ O branching. Lateral branching was also observed in recent studies of shorter peptide fragment (2–7 amino acids) self-assembly through simulations of hundreds of interacting peptides.<sup>83–85</sup> These studies, together with ours, highlight the need to include enough peptide chains in self-assembly simulations to observe events like oligomer branching, which takes place on a longer length scale than small oligomer formation.<sup>86,87</sup>

We have elucidated how A $\beta$  coordination properties dictate topological and geometric features of resulting A $\beta$ O's. We found that in A $\beta$ O's, A $\beta$  peptides with low coordination numbers tend to form A $\beta$ O's linear parts while those with high coordination numbers tend to form cyclic (mostly triangular ring) or fused polycyclic fragments, which on their own can be deemed as building blocks that were predominantly bivalent. A $\beta$ O's, being composed of primarily bivalent building blocks, thereby prefer elongated topologies. The adjacent A $\beta$  peptides along the elongated A $\beta$ O backbones exhibited no clear preference for any rotational isomeric arrangement, indicating that the A $\beta$ O's are expected to behave like a freely rotating chain that is characterized by its floppy appearance, consistent with what was observed experimentally. We also found that the more extended A $\beta$ 42 oligomer shapes (compared to A $\beta$ 40 oligomers), which has been reported experimentally and computationally,<sup>36,67</sup> could largely be attributed to a larger bond angle formed by every three consecutive monomers along the A $\beta$ 42 oligomer backbone.

Of particular interest is our finding that A $\beta$ 40 oligomers are more likely to branch than A $\beta$ 42 oligomers. The growth of branches from linear aggregates is one of the important ways of cross-linking linear aggregates, which leads to the formation of fibrous network structures required for gelation.<sup>86,87</sup> Although





the formation of gels by A $\beta$  has never been reported, both AFM and liquid-state TEM experiments have confirmed that A $\beta$ 40, but not A $\beta$ 42, could form network-like early aggregates.<sup>20,25</sup> These findings corroborate our observation that A $\beta$ 40 oligomers have greater branching propensity. The network-like aggregates appeared kinetically stable<sup>20,25</sup> and may be off the fibril pathway,<sup>56,88</sup> which could explain in part why A $\beta$ 40 fibrilization kinetics are slower than A $\beta$ 42.<sup>12</sup> Furthermore, we ascribe the greater branching ability of A $\beta$ 40 peptides to the formation of strong hydrophobic interactions between the C-terminal parts of the peptides at branching interfaces (Fig. 4b). Conversely, A $\beta$ 42 does not exhibit such strong interchain interactions at these interfaces (Fig. 4c) even though in oligomeric states, on average, A $\beta$ 42 can form stronger interactions with its neighbors than A $\beta$ 40 (Fig. 4d and e). It seems that C-terminal turn structure formation in A $\beta$ 42, but not in A $\beta$ 40 (Fig. S11<sup>†</sup>), prevented strong contacts formed between the C-terminal parts of the peptides at early stage, a consequence of competition between intramolecular and intermolecular interactions. These findings raise the possibility of redirecting early assembly processes either away from or toward the formation of branched or network-like aggregates by targeting those specific C-terminal interactions that are essential for oligomer branching. This prediction awaits further experimental and computational assessment.

## Materials and methods

### Model setup and simulation details

The PACE force field (version 1.4, <https://github.com/hanlabpkusz/hanlab/tree/master/PACE%20for%20GROMACS>) was used for all simulations.<sup>39</sup> It couples a united-atom peptide model with explicit coarse-grained solvent. The details of PACE and its parameterization can be found in ref. 38 and 39 PACE, as used here, has proven to be accurate in reproducing experimental observations of A $\beta$  aggregation, including A $\beta$  monomer <sup>3</sup>*J*-coupling constants,<sup>42</sup> A $\beta$ 42 dimer rupture force and distance patterns,<sup>40</sup> A $\beta$ 17–42 fibril growth thermodynamics and kinetics,<sup>42</sup> and the affinity of A $\beta$ 40 fibrils for peptides.<sup>89,90</sup> For both A $\beta$ 40 and A $\beta$ 42, 100 peptide chains were randomly dispersed and orientated in a 35 × 35 × 35 nm<sup>3</sup> box containing CG water particles and neutralized and buffered with a 0.15 M NaCl solution. Initial chain conformations were randomly taken from 700 ns replica of A $\beta$ 40/42 monomer exchange molecular dynamic (REMD) simulations. Of note, for each alloform, two independent self-assembly simulations were performed, each starting with a distinct set of conformations randomly selected from the pool of disordered or partially folded ones (Fig. S3<sup>†</sup>). As shown in Fig. S12,<sup>†</sup> for both alloforms, the two simulations yielded very similar results, all consistent with the main findings of the present study including the linear growth of A $\beta$ O<sub>s</sub> with their size, a more extended shape of large A $\beta$ O<sub>42</sub>, and a greater branching propensity of large A $\beta$ O<sub>40</sub>, indicating that our findings are independent of initial monomer conformations used.

There were 24 replicas at temperatures of 310–450 K in each REMD simulation. Exchange was attempted every 2 ps and

accepted with a probability  $\sim 30\%$ . All simulation systems were first energy minimized for 5000 steps, followed by 10 000 steps of pre-equilibrium simulations with heavy atoms being restrained before production runs. Self-assembly simulations were conducted at 1 atm and 310 K. Simulation timestep was set to 4.5 fs, a typical value used in PACE simulations.<sup>38,39</sup> GROMACS software (version 5) was used to conduct the simulations mentioned above.

To calculate the CCS values of a given oligomeric structure, we first estimated the p*K*<sub>a</sub> values of all titratable groups with PROPKA.<sup>91,92</sup> Those groups with the highest p*K*<sub>a</sub> were protonated such that the oligomer charge states were adjusted to those reported experimentally.<sup>67</sup> The structure was energy-minimized in gas phase with PACE for 5000 steps each, and its CCS value was calculated using the trajectory method<sup>93,94</sup> implemented in the IMPACT software.<sup>95</sup>

### Graph network analysis of topological networks

The topology of an oligomer of size *n* can be described by a graph  $G(V, E)$ , where *V* are nodes, each representing one monomer, and *E* are edges of connecting nodes. Two nodes are connected by an edge if not less than 10 pairs of atoms from the corresponding monomers are closer to each other than 0.45 nm. The longest shortest path of *G* was detected by first identifying the shortest paths between any pair of nodes *via* the breadth-first search and then choosing the longest one. All ring substructures in *G* were identified through a depth-first search excluding those containing shortcuts, followed by the removal of redundancies using a bookkeeping list.

To decompose a topological graph *G*, we first constructed an auxiliary graph *G'* whose nodes denote rings found in *G* and connectivity indicates whether rings share edges. Each disconnected component of *G'* corresponds to a monocyclic or polycyclic fragment. Any ring fragment sharing at least one edge with the longest shortest path was considered to belong to the backbone part of an oligomer. The coordination number of a building fragment was the summation of the coordination numbers of all its nodes. Of note, a direct coordination with another ring fragment is only counted as a single coordination (red dashed lines in Fig. 2a) as ring structures are closed.

The similarity between two topological networks *G*<sub>A</sub> and *G*<sub>B</sub> was measured with a score *Q* expressed as follows

$$Q = \frac{\|M_A \cap M_B(t)\|}{\|M_A \cup M_B(t)\|} \quad (1)$$

where the adjacency matrices *M*<sub>A</sub> and *M*<sub>B</sub> describe the connectivity of *G*<sub>A</sub> and *G*<sub>B</sub>, respectively. To consider permutation degeneracy, we attempted to randomly swap two rows and two columns of *M*<sub>B</sub>(*t*) each time, generating a new *M*<sub>B</sub> for a better *Q*. The calculation was stopped when *Q* no longer increased for 30 × *n* attempts in a row. Only a value of *Q* = 1 indicates that *G*<sub>A</sub> and *G*<sub>B</sub> are isomorphic and correspond to the same topology.

### Mass distribution calculation

Considering a system of *N* peptide chains in a volume *V* that can form various clusters, there will be different partitions of the



system into clusters, each of which must satisfy  $\sum_{n=1}^N nN_n = N$ ,

where  $N_n$  is the count of clusters of size  $n$ . Assuming that the system is ideal, the total partition function of the system can be written as<sup>64</sup>

$$Q_{\text{tot}}(V, T) = \sum_{\left\{ N_n \mid \sum_{n=1}^N nN_n = N \right\}} \left( \prod_{n=1}^N \frac{1}{N_n!} \left( \frac{V}{V_0} q_n^0(V_0, T) \right)^{N_n} \right), \quad (2)$$

where  $q_n^0$  is the partition function of a single cluster of size  $n$  at a reference state whose particle density  $\rho^0$  is  $1/V_0$  with  $V_0$  equal to the simulation volume ( $35^3 \text{ nm}^3$ ). From eqn (2), one can derive the expression of the average counts  $N_n$  of clusters of any size and fit these average counts to those from simulations to obtain  $q_n^0$ . The resulting  $q_n^0$  values (Fig. S6a†) permitted a best fit to the simulation cluster counts with a square sum of deviation  $< 10^{-6}$ .

Now let  $\rho_n$  be the equilibrium density of  $n$ -mers in unit  $\rho^0$ . For a system with infinite peptide chains,  $\rho_n = \frac{q_n^0}{(q_1^0)^n} \rho_1^n = K_n \rho_1^n$ , where  $K_n$  is the equilibrium constant of the formation of  $n$ -mers. Thus, the mass percentage of  $n$ -mers can be expressed as a function of monomer concentration  $\rho_1$ , *i.e.*,

$$P_m(n) = \frac{nK_n \rho_1^n}{\sum_{n'=1}^{\infty} n' K_{n'} \rho_1^{n'}} \quad (3)$$

Of note, the estimation of  $K_n$  with  $q_n^0$  is not reliable for large  $n$  due to insufficient sampling. Instead, they were extrapolated according to an apparent linear relationship between  $\ln q_n^0$  and  $n$  for  $n \geq 20$  (Fig. S6b†).

### Detection of association and dissociation events

The association events and dissociation events that took place between  $t$  and  $t + dt$  were identified with an algorithm illustrated in Fig. S13.† We first determined the oligomerization status of the systems at two time points based on peptide chain contact status. Two unbound peptide chains were thought to form a contact if they have no less than 10 pairs of atoms within 0.45 nm and, thereafter, they remain in contact unless the minimum atomic distance between them is beyond 0.45 nm. This approach avoids considering oligomer reorganization as dissociation.<sup>96</sup> We constructed a graph whose nodes represent oligomers from either frame  $t$  or frame  $t + dt$ , and two nodes from different frames can be connected if they share the same peptide chains. Each node from frame  $t$  with more than one linkage indicates a dissociation event and each node from  $t + dt$  with more than one linkage indicates an association event. Starting from an association event, we constructed an event tree of four subsequent association/dissociation events. If at any point of the event tree the dissociated oligomers were identical to one of the original

two, the original association event was thought to be nonreactive.

## Author contributions

M. Y. performed the computational experiments. M. Y., X. T. and W. H. analyzed the data. W. H. designed and supervised the research. M. Y. and W. H. wrote the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank financial supports from the National Science Foundation of China (grant no. 21977011) and the Shenzhen Fundamental Research Program (grant no. GXWD20201231165807007-20200827170132001).

## Notes and references

- J. Hardy and D. J. Selkoe, *Science*, 2002, **297**, 353–356.
- E. Y. Hayden and D. B. Teplow, *Alzheimer's Res. Ther.*, 2013, **5**, 60.
- D. J. Selkoe and J. Hardy, *EMBO Mol. Med.*, 2016, **8**, 595–608.
- I. Benilova, E. Karran and B. De Strooper, *Nat. Neurosci.*, 2012, **15**, 349–357.
- E. N. Cline, M. A. Bicca, K. L. Viola and W. L. Klein, *J. Alzheimer's Dis.*, 2018, **64**, S567–s610.
- P. T. Velasco, M. C. Heffern, A. Sebollela, I. A. Popova, P. N. Lacor, K. B. Lee, X. Sun, B. N. Tian, K. L. Viola, A. L. Eckermann, T. J. Meade and W. L. Klein, *ACS Chem. Neurosci.*, 2012, **3**, 972–981.
- P. Liu, M. N. Reed, L. A. Kotilinek, M. K. Grant, C. L. Forster, W. Qiang, S. L. Shapiro, J. H. Reichl, A. C. Chiang, J. L. Jankowsky, C. M. Wilmot, J. P. Cleary, K. R. Zahs and K. H. Ashe, *Cell Rep.*, 2015, **11**, 1760–1771.
- T. Watanabe-Nakayama, K. Ono, M. Itami, R. Takahashi, D. B. Teplow and M. Yamada, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 5835–5840.
- M. I. Ivanova, Y. Lin, Y. H. Lee, J. Zheng and A. Ramamoorthy, *Biophys. Chem.*, 2021, **269**, 106507.
- S. Sharma, P. Modi, G. Sharma and S. Deep, *Biophys. Chem.*, 2021, **278**, 106665.
- G. J. Morgan, *Biophys. Chem.*, 2022, **281**, 106711.
- R. Jakob-Roetne and H. Jacobsen, *Angew. Chem., Int. Ed.*, 2009, **48**, 3030–3059.
- D. C. Bode, M. D. Baker and J. H. Viles, *J. Biol. Chem.*, 2017, **292**, 1404–1413.
- P. H. Nguyen, A. Ramamoorthy, B. R. Sahoo, J. Zheng, P. Faller, J. E. Straub, L. Dominguez, J. E. Shea, N. V. Dokholyan, A. De Simone, B. Ma, R. Nussinov, S. Najafi, S. T. Ngo, A. Loquet, M. Chiricotto, P. Ganguly, J. McCarty, M. S. Li, C. Hall, Y. Wang, Y. Miller, S. Melchionna, B. Habenstein, S. Timr, J. Chen, B. Hnath, B. Strodel, R. Kaye, S. Lesné, G. Wei, F. Sterpone,



- A. J. Doig and P. Derreumaux, *Chem. Rev.*, 2021, **121**, 2545–2647.
- 15 L. Feng, H. Watanabe, P. Molino, G. G. Wallace, S. L. Phung, T. Uchihashi and M. J. Higgins, *J. Mol. Biol.*, 2019, **431**, 2687–2699.
- 16 S. Banerjee, Z. Sun, E. Y. Hayden, D. B. Teplow and Y. L. Lyubchenko, *ACS Nano*, 2017, **11**, 12202–12209.
- 17 E. E. Cawood, T. K. Karamanos, A. J. Wilson and S. E. Radford, *Biophys. Chem.*, 2021, **268**, 106505.
- 18 B. R. Sahoo, S. J. Cox and A. Ramamoorthy, *Chem. Commun.*, 2020, **56**, 4627–4639.
- 19 I. A. Mastrangelo, M. Ahmed, T. Sato, W. Liu, C. Wang, P. Hough and S. O. Smith, *J. Mol. Biol.*, 2006, **358**, 106–119.
- 20 K. Nakajima, T. Yamazaki, Y. Kimura, M. So, Y. Goto and H. Ogi, *J. Phys. Chem. Lett.*, 2020, **11**, 6176–6184.
- 21 B. A. Chromy, R. J. Nowak, M. P. Lambert, K. L. Viola, L. Chang, P. T. Velasco, B. W. Jones, S. J. Fernandez, P. N. Lacor, P. Horowitz, C. E. Finch, G. A. Krafft and W. L. Klein, *Biochemistry*, 2003, **42**, 12749–12760.
- 22 M. P. Lambert, A. K. Barlow, B. A. Chromy, C. Edwards, R. Freed, M. Liosatos, T. E. Morgan, I. Rozovsky, B. Trommer, K. L. Viola, P. Wals, C. Zhang, C. E. Finch, G. A. Krafft and W. L. Klein, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6448–6453.
- 23 D. M. Walsh, D. M. Hartley, Y. Kusumoto, Y. Fezoui, M. M. Condrón, A. Lomakin, G. B. Benedek, D. J. Selkoe and D. B. Teplow, *J. Biol. Chem.*, 1999, **274**, 25945–25952.
- 24 F. Hasecke, T. Miti, C. Perez, J. Barton, D. Schölzel, L. Gremer, C. S. R. Grüning, G. Matthews, G. Meisl, T. P. J. Knowles, D. Willbold, P. Neudecker, H. Heise, G. Ullah, W. Hoyer and M. Muschol, *Chem. Sci.*, 2018, **9**, 5937–5948.
- 25 N. J. Economou, M. J. Giammona, T. D. Do, X. Zheng, D. B. Teplow, S. K. Buratto and M. T. Bowers, *J. Am. Chem. Soc.*, 2016, **138**, 1772–1775.
- 26 W. Xi and U. H. E. Hansmann, *Sci. Rep.*, 2017, **7**, 6588.
- 27 W. Xi, D. N. Dean, K. A. Stockmal, S. E. Morgan, U. H. E. Hansmann and V. Rangachari, *J. Chem. Phys.*, 2019, **150**, 075101.
- 28 Y. Tian, R. Liang, A. Kumar, P. Szwedziak and J. H. Viles, *Chem. Sci.*, 2021, **12**, 6896–6907.
- 29 L. Nagel-Steger, M. C. Owen and B. Strodel, *ChemBioChem*, 2016, **17**, 657–676.
- 30 I. M. Ilie and A. Cafflich, *Chem. Rev.*, 2019, **119**, 6956–6993.
- 31 S. J. Bunce, Y. Wang, K. L. Stewart, A. E. Ashcroft, S. E. Radford, C. K. Hall and A. J. Wilson, *Sci. Adv.*, 2019, **5**, eaav8216.
- 32 W. Zheng, M. Y. Tsai and P. G. Wolynes, *J. Am. Chem. Soc.*, 2017, **139**, 16666–16676.
- 33 B. Strodel, *Curr. Opin. Struct. Biol.*, 2021, **67**, 145–152.
- 34 B. Urbanc, M. Betnel, L. Cruz, G. Bitan and D. B. Teplow, *J. Am. Chem. Soc.*, 2010, **132**, 4266–4280.
- 35 S. Zhang, D. M. Fox and B. Urbanc, *J. Phys. Chem. B*, 2017, **121**, 5523–5535.
- 36 B. Barz, Q. Liao and B. Strodel, *J. Am. Chem. Soc.*, 2018, **140**, 319–327.
- 37 B. Barz, O. O. Olubiyi and B. Strodel, *Chem. Commun.*, 2014, **50**, 5373–5375.
- 38 W. Han and K. Schulten, *J. Chem. Theory Comput.*, 2012, **8**, 4413–4424.
- 39 W. Han and K. Schulten, *J. Phys. Chem. B*, 2013, **117**, 13367–13377.
- 40 Y. Zhang, M. Hashemi, Z. Lv and Y. L. Lyubchenko, *Nanoscale*, 2016, **8**, 18928–18937.
- 41 Y. Cao, X. Jiang and W. Han, *J. Chem. Theory Comput.*, 2017, **13**, 5731–5744.
- 42 W. Han and K. Schulten, *J. Am. Chem. Soc.*, 2014, **136**, 12450–12460.
- 43 J. Zivkovic and B. Tadic, *Nanoscale Syst.: Math. Model., Theory Appl.*, 2013, **2**, 30–48.
- 44 N. Arkus, V. N. Manoharan and M. P. Brenner, *SIAM J. Discrete Math.*, 2011, **25**, 1860–1901.
- 45 N. Arkus, V. N. Manoharan and M. P. Brenner, *Phys. Rev. Lett.*, 2009, **103**, 118303.
- 46 S. Hormoz and M. P. Brenner, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 5193–5198.
- 47 Z. Zeravcic, V. N. Manoharan and M. P. Brenner, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15918–15923.
- 48 W. M. Jacobs, A. Reinhardt and D. Frenkel, *J. Chem. Phys.*, 2015, **142**, 021101.
- 49 W. M. Jacobs, A. Reinhardt and D. Frenkel, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 6313–6318.
- 50 A. Hirata, L. J. Kang, T. Fujita, B. Klumov, K. Matsue, M. Kotani, A. R. Yavari and M. W. Chen, *Science*, 2013, **341**, 376–379.
- 51 M. D. Kirkitadze, M. M. Condrón and D. B. Teplow, *J. Mol. Biol.*, 2001, **312**, 1103–1119.
- 52 P. Das, B. Murray and G. Belfort, *Biophys. J.*, 2015, **108**, 738–747.
- 53 P. H. Nguyen and P. Derreumaux, *Biophys. Chem.*, 2020, **264**, 106421.
- 54 K. A. Ball, A. H. Phillips, D. E. Wemmer and T. Head-Gordon, *Biophys. J.*, 2013, **104**, 2714–2724.
- 55 S. Vivekanandan, J. R. Brender, S. Y. Lee and A. Ramamoorthy, *Biochem. Biophys. Res. Commun.*, 2011, **411**, 312–316.
- 56 Y. Lin, B. R. Sahoo, D. Ozawa, M. Kinoshita, J. Kang, M. H. Lim, M. Okumura, Y. H. Huh, E. Moon, J. H. Jang, H. J. Lee, K. Y. Ryu, S. Ham, H. S. Won, K. S. Ryu, T. Sugiki, J. K. Bang, H. S. Hoe, T. Fujiwara, A. Ramamoorthy and Y. H. Lee, *ACS Nano*, 2019, **13**, 8766–8783.
- 57 R. Pujol-Pina, S. Vilapriñyó-Pascual, R. Mazzucato, A. Arcella, M. Vilaseca, M. Orozco and N. Carulla, *Sci. Rep.*, 2015, **5**, 14809.
- 58 S. A. Kotler, J. R. Brender, S. Vivekanandan, Y. Suzuki, K. Yamamoto, M. Monette, J. Krishnamoorthy, P. Walsh, M. Cauble, M. M. Holl, E. N. Marsh and A. Ramamoorthy, *Sci. Rep.*, 2015, **5**, 11811.
- 59 F. Meng, M. M. J. Bellaiche, J. Y. Kim, G. H. Zerze, R. B. Best and H. S. Chung, *Biophys. J.*, 2018, **114**, 870–884.
- 60 V. A. Voelz, G. R. Bowman, K. Beauchamp and V. S. Pande, *J. Am. Chem. Soc.*, 2010, **132**, 1526–1528.



- 61 H. Yu, W. Han, W. Ma and K. Schulten, *J. Chem. Phys.*, 2015, **143**, 243142.
- 62 M. Nick, Y. Wu, N. W. Schmidt, S. B. Prusiner, J. Stöhr and W. F. DeGrado, *Biopolymers*, 2018, **109**, e23096.
- 63 R. S. Hoy, J. Harwayne-Gidansky and C. S. O'Hern, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 051403.
- 64 J. T. Kindt, *J. Chem. Theory Comput.*, 2013, **9**, 147–152.
- 65 L. A. Patel and J. T. Kindt, *J. Chem. Theory Comput.*, 2017, **13**, 1023–1033.
- 66 M. Kloniecki, A. Jabłonowska, J. Poznański, J. Langridge, C. Hughes, I. Campuzano, K. Giles and M. Dadlez, *J. Mol. Biol.*, 2011, **407**, 110–124.
- 67 S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wytenbach, M. M. Condrón, G. Bitan, D. B. Teplow, J. E. Shea, B. T. Ruotolo, C. V. Robinson and M. T. Bowers, *Nat. Chem.*, 2009, **1**, 326–331.
- 68 L. O. Tjernberg, A. Pramanik, S. Björling, P. Thyberg, J. Thyberg, C. Nordstedt, K. D. Berndt, L. Terenius and R. Rigler, *Chem. Biol.*, 1999, **6**, 53–62.
- 69 W. Yong, A. Lomakin, M. D. Kirkitadze, D. B. Teplow, S. H. Chen and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 150–154.
- 70 R. Sabaté and J. Estelrich, *J. Phys. Chem. B*, 2005, **109**, 11027–11032.
- 71 M. Novo, S. Freire and W. Al-Soufi, *Sci. Rep.*, 2018, **8**, 1783.
- 72 B. Soreghan, J. Kosmoski and C. Glabe, *J. Biol. Chem.*, 1994, **269**, 28551–28554.
- 73 K. Brännström, A. Öhman, L. Nilsson, M. Pihl, L. Sandblad and A. Olofsson, *J. Am. Chem. Soc.*, 2014, **136**, 10956–10964.
- 74 E. Hellstrand, B. Boland, D. M. Walsh and S. Linse, *ACS Chem. Neurosci.*, 2010, **1**, 13–18.
- 75 X. Zheng, D. Liu, R. Roychaudhuri, D. B. Teplow and M. T. Bowers, *ACS Chem. Neurosci.*, 2015, **6**, 1732–1740.
- 76 N. L. Fawzi, J. Ying, R. Ghirlando, D. A. Torchia and G. M. Clore, *Nature*, 2011, **480**, 268–272.
- 77 N. L. Fawzi, J. Ying, D. A. Torchia and G. M. Clore, *J. Am. Chem. Soc.*, 2010, **132**, 9948–9951.
- 78 P. H. Nguyen, M. S. Li, G. Stock, J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 111–116.
- 79 J. R. Brender, A. Ghosh, S. A. Kotler, J. Krishnamoorthy, S. Bera, V. Morris, T. B. Sil, K. Garai, B. Reif, A. Bhunia and A. Ramamoorthy, *Chem. Commun.*, 2019, **55**, 4483–4486.
- 80 E. Gazit, *FASEB J.*, 2002, **16**, 77–83.
- 81 N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang and A. E. Garcia, *J. Mol. Biol.*, 2007, **368**, 1448–1457.
- 82 T. Kowalewski and D. M. Holtzman, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3688–3693.
- 83 P. Chakraborty, Y. Tang, T. Yamamoto, Y. Yao, T. Guterman, S. Zilberzwige-Tal, N. Adadi, W. Ji, T. Dvir, A. Ramamoorthy, G. Wei and E. Gazit, *Adv. Mater.*, 2020, **32**, 1906043.
- 84 M. Chiricotto, S. Melchionna, P. Derreumaux and F. Sterpone, *J. Phys. Chem. Lett.*, 2019, **10**, 1594–1599.
- 85 P. Chakraborty, Y. Tang, T. Guterman, Z. A. Arnon, Y. Yao, G. Wei and E. Gazit, *Angew. Chem., Int. Ed.*, 2020, **59**, 23731–23739.
- 86 J.-L. Li and X.-Y. Liu, *Adv. Funct. Mater.*, 2010, **20**, 3196–3216.
- 87 E. R. Draper and D. J. Adams, *Chem*, 2017, **3**, 390–410.
- 88 M. So, D. Hall and Y. Goto, *Curr. Opin. Struct. Biol.*, 2016, **36**, 32–39.
- 89 X. Jiang, Y. Cao and W. Han, *ACS Chem. Neurosci.*, 2018, **9**, 935–944.
- 90 Y. Jiang, X. Jiang, X. Shi, F. Yang, Y. Cao, X. Qin, Z. Hou, M. Xie, N. Liu, Q. Fang, F. Yin, W. Han and Z. Li, *iScience*, 2019, **17**, 87–100.
- 91 M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
- 92 C. R. Søndergaard, M. H. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284–2295.
- 93 M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *J. Phys. Chem.*, 1996, **100**, 16082–16086.
- 94 A. A. Shvartsburg and M. F. Jarrold, *Chem. Phys. Lett.*, 1996, **261**, 86–91.
- 95 E. G. Marklund, M. T. Degiacomi, C. V. Robinson, A. J. Baldwin and J. L. P. Benesch, *Structure*, 2015, **23**, 791–799.
- 96 M. Carballo-Pacheco, A. E. Ismail and B. Strodel, *J. Chem. Theory Comput.*, 2018, **14**, 6063–6075.

