doi:10.1016/j.jmb.2006.10.011 J. Mol. Biol. (2007) 365, 535–550

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Determining the Critical Nucleus and Mechanism of Fibril Elongation of the Alzheimer's $A\beta_{1-40}$ Peptide

Nicolas Lux Fawzi[1], Yuka Okabe[2], Eng-Hui Yap[1] and Teresa Head-Gordon[1,2,3]*

[1]*UCSF/UCB Joint Graduate Group in Bioengineering Berkeley, CA 94720, USA*

[2]*Department of Bioengineering University of California Berkeley, CA 94720, USA*

[3]*Department of Chemistry Cambridge University Cambridge CB2 1EW, UK*

**Corresponding author*

We use a coarse-grained protein model to characterize the critical nucleus, structural stability, and fibril elongation propensity of $A\beta_{1-40}$ oligomers for the C_{2x} and C_{2z} quaternary forms proposed by solid-state NMR. By estimating equilibrium populations of structurally stable and unstable protofibrils, we determine the shift in the dominant population from free monomer to ordered fibril at a critical nucleus of ten chains for the C_{2x} and C_{2z} forms. We find that a minimum assembly of 16 monomer chains is necessary to mimic a mature fibril, and show that its structural stability correlates with a plateau in the hydrophobic residue density and a decrease in the likelihood of losing hydrophobic interactions by rotating the fibril subunits. While $A\beta_{1-40}$ protofibrils show similar structural stability for both C_{2x} and C_{2z} quaternary structures, we find that the fibril elongation propensity is greater for the C_{2z} form relative to the C_{2x} form. We attribute the increased propensity for elongation of the C_{2z} form as being due to a stagger in the interdigitation of the N-terminal and C-terminal β-strands, resulting in structural asymmetry in the presented fibril ends that decreases the amount of incorrect addition to the N terminus on one end. We show that because different combinations of stagger and quaternary structure affect the structural symmetry of the fibril end, we propose that differences in quaternary structures will affect directional growth patterns and possibly different morphologies in the mature fiber.

Keywords: aggregation; amyloid; Aβ peptide; molecular dynamics; Alzheimer's

Introduction

The aggregation of peptides or proteins into ordered amyloid fibril morphologies is associated with over 20 human diseases, including Alzheimer's disease, dialysis-related amyloidosis, and bovine spongiform encephalopathy.[1,2] The fibrils have a characteristic "cross-β" structure, where intermolecular β-sheets run along the long axis of the fibril, stabilizing the assemblies that can extend to micrometers in length.[2] Although early attention focused on the toxicity of the amyloid fibrils as the cause of disease, it is now hypothesized that oligomers formed during early aggregation are actually the major toxic species.[3,4] This shift underscores the need to develop an understanding of the entire aggregation process that ultimately leads to the specific structure of the final amyloid fibril.

Alzheimer's is a neurodegenerative disease linked to the aggregation and amyloid fibril formation of a set of short ~40 residue peptides, amyloid β ($A\beta_{1-39,1-40,1-42}$), created by proteolytic cleavage of the amyloid precursor protein (APP).[5] These fragments contain part of the C-terminal region of the APP protein, and are known to be highly prone to fibrilization *in vitro* and *in vivo*.[6–10] The structure of the monomeric peptide has no well-defined folded state, although tertiary structures that are dependent on solution conditions have been proposed from experimental and simulation work.[11–14] The backbone conformation can vary from α-helical structure in non-polar solutions as determined by solution NMR,[11,12] to disordered N-terminal and C-terminal tails with a consistent turn region, as determined from electrospray mass spectrometry and implicit solvent molecular dynamics.[13,14] The structure of the $A\beta_{21-30}$ sub-peptide, encompassing a proteolysis re-

Abbreviation used: APP, amyloid precursor protein.

E-mail address of the corresponding author: tlhead-gordon@lbl.gov

sistant region of the full-length sequence, has also been determined by NMR.[15]

At the other extreme, the complete Aβ$_{1-40}$ amyloid fibril state has been studied extensively by Tycko and co-workers, who have published a series of model structures based on constraints from solid-state NMR.[16–20] The proposed structure is shown in Figure 1 and is described as U-shaped monomers with two in-register parallel intermolecular β-sheet regions (N- and C-terminal β-sheets); the cross-section of the fibril is composed of two monomers with hydrophobic C-terminal regions in van der Waals contact. The original NMR data[19] supported two possible intra-fibril contact types (unflipped and flipped) for the C-terminal β-strand, and eventually the unflipped form was eliminated on the basis of tertiary side-chain–side-chain contacts.[16]

Furthermore, two quaternary structures denoted as C_{2x} and C_{2z} were proposed,[21] based on approximate $C2$ symmetry around the x axis (approximately orthogonal to the fibril axis and parallel with the β-strand directions) and $C2$ symmetry around the z axis (parallel with the fibril axis), respectively, and shown in Figure 1. Note that these are only pseudo-symmetry designators since there is imperfect matching of side-chain interdigitation in the C-terminal region on opposite subunits of the relevant protofibril symmetry axis in both cases. More complete NMR data revealed that only the C_{2z} quaternary structure was likely to be formed *in vitro* on the basis of specific 2D NMR cross-peaks that give tertiary contacts that are inconsistent with the C_{2x} quaternary form.[16] Most recently, however, a fibril made from shortened, mutated Aβ monomers covalently linked at the N termini created fibrils with a likely C_{2x} symmetry, indicating that the C_{2x} form may be found under certain conditions.[22]

Finally, the NMR data also support interdigitation of the N-terminal and C terminal β-strands to form side-chain contacts with a particular "stagger" of N-terminal and C-terminal hydrophobic contacts,[16] shown schematically in Figure 2. On the basis of the results of isotopic dilution studies, side-chain contacts are proposed between the C termini of monomer i with the N termini of monomers $i+1$ and $i+2$ (STAG (−2)) or between the N termini of monomer i with the C termini of monomers $i+1$ and $i+2$ (STAG (+2)).[16] In totality, the solid-state NMR work is a truly seminal contribution to the amyloid field, since these experimental models have provided well-defined structural constraints on the "folded state" of the Aβ$_{1-40}$ monomer in the context of the formed fibril.

Given the possible toxicity of the earlier protofibril states, the focus is now to understand how the Aβ monomers assemble into the highly ordered mesoscopic fibril, as proposed by the NMR experimental models. The mechanism of fibrillization of full-length Aβ peptides (Aβ$_{1-39,1-40,1-42}$) has been shown to follow an apparent nucleation-dependent polymerization,[9,10,17,22] whereby a small number of monomers associate through a free energy barrier corresponding to a critical nucleus size, beyond which initiates a gradient of favorable free energy or "down-hill" polymerization into a macroscopic fibril (Figure 3).[23] However, the structural characteristics and oligomer size of this ensemble of fibril nucleating species have yet to be determined, and the mechanism of monomer addition is unclear. This is due, in part, to the limited access of experimental characterization to this earliest aggregation stage, thus providing an opportunity for theoretical studies to bridge the experimental gap between the monomer and fibril endpoints and to develop testable hypotheses.

Many computational studies using coarse-grained as well as all-atom models have focused on the formation of the antiparallel β-sheet structure by sub-peptides of Aβ, particularly Aβ$_{16-22}$.[24–26] The antiparallel structure of these peptides, however, suggests that studies of the steps in fibril formation of this system will not lead to information regarding the nucleation and fibril forming properties of the in-register parallel structures formed by the full-length Aβ$_{1-40}$ and Aβ$_{1-42}$ peptides. More recent simulation work has therefore focused on the full-length Aβ peptides. Coarse-grained simulations of the Aβ$_{1-40}$ and Aβ$_{1-42}$ monomers and dimers reported by Stanley and co-workers have reproduced some of the properties of the disordered peptides in solution,[27,28] but underscore the computational and modeling difficulty of forming structures resembling fibrils. All-atom simulations conducted by Shea and co-workers give detailed insight into the monomer structure in dilute solution and *in vacuo*.[13,14] In a set of all-atom molecular dynamics simulations with explicit water representation, Hummer and co-workers demonstrate that with incomplete NMR data from Tycko and co-workers, a set of four related but distinct minimum fibril models consistent with the NMR restraints are all structurally stable for at least a few nanoseconds.[21] All-atom simulation of the full-length Aβ peptide aggregation is likely difficult due to the extremely long experimental timescales (hours to days, depending on conditions) and large system sizes (about ten peptides of 40 amino acid residues) necessary for fibril formation.

We have recently developed a new coarse-grained protein model that is a greatly enhanced version of coarse-grained models we have used in studying protein folding and non-disease protein aggregation.[29–34] The new model, described in Methods, has been validated on folding thermodynamics and kinetics for proteins L and G, and provides higher structural resolutions (~3.0 Å C^{α} RMSD) of the folded state relative to our old model, especially for descriptions of β-sheets (E.-H. Y., N.L.F. and T.H.-G., unpublished results). We use this new model for the first time to simulate Aβ$_{1-40}$ oligomerization in order to address three primary questions regarding the association of Aβ$_{1-40}$ peptides into fibrils.

First, what is the number of peptides involved in the critical nucleus for subsequent fibril elongation and does it differ among quaternary forms? Starting from a mature fibril structure composed of 40 chains,