Supporting Information

Computational Study on Structure and Aggregation Pathway of Aβ_{42} Amyloid Protofibril

MinJun Lee†, Jeseong Yoon*,†, and Seokmin Shin*,†

†Department of Chemistry, Seoul National University, Seoul 151-747, Korea

Corresponding Authors: *yjs0@snu.ac.kr, *sshin@snu.ac.kr

Table of Contents

1. Structure and Dynamics of Triple-β Fibril Motif
   Figure S1-S2: Comparison of triple-β motif between experimental and simulation structure  S2
   Table S1: Sidechain interactions of triple-β motif structure from simulation  S4
   Figure S3: Backbone RMSD of four internal chains and two external chains of triple-β motif  S5
   Figure S4: Conformational dynamics of triple-β motif through 10 μs MD simulation  S6

2. Dynamics of N-terminus in Triple-β Fibril Motif
   Figure S5: Backbone Torsion Angle Dynamics of N-terminus in Triple-β Fibril Motif  S8

3. Steered MD Simulations of Triple-β Fibril Motif: Lock Phase Study
   Figure S6: Force extension curves of steered MD simulations  S9
   Table S2: φ(r) profiles from steered MD simulations  S10
   Figure S7: Second major rupture pattern  S11
   Figure S8: Trajectories of major and second major rupture patterns  S12

4. REMD Simulation of Monomeric Aβ_{42} + Triple-β Fibril Motif: Dock Phase Study
   Figure S9: Configurations of Aβ_{42} monomer in docking process by REMD simulation  S13
   Figure S10: Distance change between Aβ_{42} monomer and fibril edge  S14
   Figure S11: Ratio of radius of gyration for β_{1}, β_{2}, and β_{3}  S15
   Figure S12: Time evolutions of β-sheet content for β_{1}, β_{2}, and β_{3} and representative configurations  S16
   Figure S13: Time evolutions of β-sheet and α-helix content of whole sequence  S17
   Figure S14: Secondary structure of β_{2} in Aβ_{42} monomer and monomer-fibril distance  S18
   Figure S15: Backbone hydrogen bond formation of β_{3} between monomer and fibril  S19
1. Structure and Dynamics of Triple-β Fibril Motif

Comparison of triple-β motif between experimental and simulation structure

**Figure S1.** In experimental structure (PDB ID: 2MXU), residues 12 to 18, 24 to 33, and 36 to 40 form β-sheets (blue arrow). In simulations, residues 12 to 19, 22 to 33, and 39 to 41 form β-sheets (yellow ribbon).
Comparison of triple-β motif between experimental and simulation structure

Figure S2. Experimental (PDB ID: 2MXU) and simulation structures of the triple-β motif. Residues shown in orange color have β-sheet geometry in either one of experimental and simulation structure. Comparison of sidechain contact for these residues shows that simulation structure has more compact hydrophobic packing due to the presence of water. Additional sidechain contacts occur between residues 12-34, 14-32, and 32-39. (purple circles)
Sidechain interactions of triple-β motif structure from simulation

<table>
<thead>
<tr>
<th>Intra-Chain Sidechain Interactions</th>
<th>Inter-Chain Sidechain Interactions (inner chain–outer chain)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HP contact</strong> Val12–Leu34</td>
<td><strong>HP contact</strong> Val12–Val12</td>
</tr>
<tr>
<td><strong>HP core Ia</strong></td>
<td><strong>HP core Ia</strong></td>
</tr>
<tr>
<td>His14–Ile32</td>
<td>His14–Ile32</td>
</tr>
<tr>
<td>Leu17–Phe19</td>
<td>Leu17–Ile32</td>
</tr>
<tr>
<td>Phe19–Ile32</td>
<td>Leu17–Leu17</td>
</tr>
<tr>
<td><strong>HP core Ib</strong></td>
<td><strong>HP core Ib</strong></td>
</tr>
<tr>
<td>Phe19–Phe20</td>
<td>Phe19–Phe19</td>
</tr>
<tr>
<td>Phe19–Ala30</td>
<td>Phe19–Ala30</td>
</tr>
<tr>
<td>Phe20–Ala30</td>
<td>Phe20–Phe20</td>
</tr>
<tr>
<td><strong>HP contact</strong> Phe20–Val24</td>
<td><strong>HP contact</strong> Val24–Val24</td>
</tr>
<tr>
<td>Phe20–Asn27</td>
<td>Val24–Val24</td>
</tr>
<tr>
<td><strong>HP core II</strong></td>
<td><strong>HP core II</strong></td>
</tr>
<tr>
<td>Ile31–Val39</td>
<td>Val39–Ile31</td>
</tr>
<tr>
<td>Met35–Val39</td>
<td>Met35–Val39</td>
</tr>
<tr>
<td><strong>HP contact</strong> Val39–Ile41</td>
<td><strong>HP contact</strong></td>
</tr>
<tr>
<td>Lys28–Ile41</td>
<td>Ile41–Lys28</td>
</tr>
<tr>
<td>Salt bridge Lys28–C-ter</td>
<td>Salt bridge C-ter–Lys28</td>
</tr>
</tbody>
</table>

Table S1. Sidechain interactions in the triple-β motif. Interactions forming two hydrophobic cores – HP core I and HP core II – are designated by colors consistent with Fig. 2. HP core I has two sub-cores named as HP core Ia and HP core Ib in this table. HP contact means hydrophobic interactions not included in HP core I and II. Two polar interactions are hydrogen bonding between Asn27-Asn27 sidechains and salt-bridge between C-terminal carboxylic group and Lys28 sidechain. Ala30, Ile31, Ile32, and Met35 are pivotal residues in the two hydrophobic cores. Note that inter-chain interaction pairs are written in order of inner chain residue–outer chain residue. For example, Val39–Ile31 means that Val39 of the inner chain interacts with Ile31 of the outer chain.
Backbone RMSD of four internal chains and two external chains of triple-β motif

Figure S3. Backbone RMSD of the two external chains (red line) are larger than the four internal chains (blue line). Two external chains are relatively flexible compared to the four internal chains.
Conformational dynamics of triple-β motif through 10 µs MD simulation

Figure S4. (a) Conformational change of the four internal chains. Denoted numbers are chain numbers. Two external chains are not shown. Residues 19-42 form stable core of the structural motif for Aβ42 protofibril (gray ribbons). On the contrary, residues 11-18 form unstructured N-terminal tail (green ribbons). (b) Conformational change of the two external chains as represented by colored ribbons. Residues 21-29 and 38-42 are colored in blue and red, respectively and the rest are in yellow. Four internal chains are represented by the transparent gray ribbons. Residues 21-29 and residues 38-42 are relatively unstable compared to other residues in core (19-42). Residues 11-15 form unstructured N-terminal tail similarly as in the four internal chains. Detailed configurational changes are explained in the following section.
Detailed descriptions of conformational changes in Figure S4

Conformational change of the four internal chains [see Figure S4a]
- 140 ns: Highly ordered structure is maintained in both the core residues (19-42) and the N-terminal residues (11-18).
- 750 ns: Hydrophobic core structure around residues 19-20 and 32-34 becomes more compact by backbone contraction which accompanies closer hydrophobic contact between N-terminal tail residues (11-15) and residues 32-34. This conformational change is due to an optimization of hydrophobic interactions in aqueous condition.
- 1170 ns: The plane made of the flexible N-terminal tails becomes tilted with respect to the fibril axis.
- 2500 ns: Backbone hydrogen bonding of the chains 4 and 5 at N-terminus are partially broken.
- 6400 ns: N-terminal tails become more unstructured and the corresponding hydrophobic sidechain packing is weakened.

Conformational change of the two external chains [see Figure S4b]
- 140 ns: Residues 23-28 of the chain 1 and residues 21-29 of the chain 6 are detached from the fibril.
- 750 ns: Backbone interactions of the chain 6 disrupted at 140 ns are recovered. Instead, residues 38-42 of the chain 1 are detached from the chain 2.
- 1170 ns: Reformed backbone alignment of the chain 6 at 750 ns are disrupted again. As mentioned in the main text, the lack of stability of this region is due to the lack of sidechain interactions contributing to hydrophobic core formation and distorted backbone geometry not adequate for $\beta$-sheet formation. C-terminal stretch of the chain 1 becomes disordered while keeping close to the chain 2.
- 2500 ns: Unstable sequence of the chain 6 (~ residues 21-29) becomes more unstructured with additional disruption of backbone hydrogen bonding with adjacent chains. In addition, N-terminal tail stretch of chains 6 and 1 are separated from each neighboring chain.
- 6400 ns: As the unstable region mentioned above becomes fully unstructured, conformation transition to $\alpha$-helical structure is observed, which is obviously the process of maximizing the hydrophobic interactions when exposed to aqueous environment.

To summarize, the structure of the triple-$\beta$ fibril motif is highly stable inside the fibril. Aqueous environment causes additional optimization of the hydrophobic core structure which makes overall geometry more compact. After losing one of the two neighboring chains at the terminal position, unstable sequences (which are around N-terminal, C-terminal region, and ~residues 21-29) show disordered and flexible motion.
2. Dynamics of N-terminus in Triple-β Fibril Motif

Backbone Torsion Angle Dynamics of N-terminus in Triple-β Fibril Motif

**Figure S5.** Dynamics of Backbone Torsion Angles ($\phi$, $\psi$) of N-terminal residues in Triple-β Fibril Motif.
3. Steered MD Simulations of Triple-β Fibril Motif: Lock Phase Study

Force extension curves of steered MD simulations

Figure S6. Force vs time curves (left column) and force vs distance curves (right column). Each row has two representations of force curves for the same set of ten independent simulation trajectories with constant pulling rate. As pulling rate is decreased, deviation of peak heights among different trajectories is reduced.
\( \varphi(r) \) profiles from steered MD simulations

<table>
<thead>
<tr>
<th>( k_{\text{pulling}} )</th>
<th>0.002 nm/ps</th>
<th>0.001 nm/ps</th>
<th>0.0005 nm/ps</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>major</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>a</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>a'</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>b</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>~ major</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table S2. Statistics of the logistic function profiles. Major pattern shown in Fig. 5 appears in 16 trajectories of the total 30 trajectories. Especially, most of the trajectories with \( k_{\text{pulling}} = 0.0005 \) nm/ps (8 among 10 trajectories) show the major pattern. Second major rupture pattern denoted by ‘a’ in the table appears in 7 trajectories. The rest of the patterns show minor contribution and corresponding plots are not shown here.
Second major rupture pattern

Figure S7. Second major rupture pattern (pattern ‘a’ in Table S2) is observed in 7 trajectories. Collective rupture and sequential rupture in residues 11-34 are similar to the major pattern. On the other hand, interactions of residues 35-41 become disrupted much earlier compared to the major pattern. The shape of the plot in the second box (28O-29N ~ 4Cter-5Lys28) seems to be symmetric with respect to the 35O-36N, indicating that the disruption of HP core II occurs from the breaking of salt-bridge interaction between Lys28 and C-terminus. It drives the propagation of simultaneous breaking in backbone hydrogen bonds from C-terminus to Met35 and from Lys28 to Met35.
Trajectories of major and second major rupture patterns

Figure S8. (a) Configurations corresponding to the trajectory of the major rupture pattern. (b) Configurations corresponding to the trajectory of the second major rupture pattern. In the major rupture pattern, C-terminus is the last part to be disrupted. In the second major rupture pattern, C-terminus is disrupted earlier than the major pattern as shown in blue dashed circles.
4. REMD Simulation of Monomeric $\text{A}\beta_{42}$ + Triple-$\beta$ Fibril Motif: Dock Phase Study

Configurations of $\text{A}\beta_{42}$ monomer in docking process by REMD simulation

**Figure S9.** (a) Minimum distance between $\text{A}\beta_{42}$ monomer and protofibril motif. The isolated monomer comes into contact with protofibril motif at 10.3 ns (dashed line in red). After contact, the distance fluctuates around 4 Å. (b) Representative S-like conformations of monomeric chain in REMD trajectory. Designated numbers (blue) are the same as in Fig. 6. S-like conformations are more dominant in isolated monomeric state than after contact with protofibril motif. After contact, $\beta_2$ forms strong helix to provide stable interface with fibril edge which leads to possible formation of strong salt-bridge interaction. For this, $\beta_3$ can form $\beta$-sheet structure which initiates the lock process towards complete fibril structure.
Distance change between Aβ₄₂ monomer and fibril edge

Figure S10. (a) Center of mass distance between the outermost chain of protofibril motif and Aβ₄₂ monomer. (b) RMSD of the combined system (outermost chain in protofibril motif + monomeric Aβ₄₂) with respect to the triple-β configuration. Yellow curve shows the average plot of the original data (black curve).
Ratio of radius of gyration for $\beta_1$, $\beta_2$, and $\beta_3$

**Figure S11.** Ratio of radius of gyration for $\beta_1$, $\beta_2$, and $\beta_3$, $R_g/(R_g)_0$, where $R_g$ is the radius of gyration of the combined system (outermost chain in protofibril motif + monomeric A$\beta_{42}$) and $(R_g)_0$ is the value of $R_g$ when the two chains form the triple-$\beta$ configuration. Small value of ratio illustrates that monomeric A$\beta_{42}$ is in close contact to protofibril to form the triple-$\beta$ configuration. It is noted that the behavior of this value is the same as the behavior of the distance in Fig. 7.
Time evolutions of $\beta$-sheet content for $\beta_1$, $\beta_2$, and $\beta_3$ and representative configurations

**Figure S12.** (a) Time evolution of $\beta$-sheet content of each $\beta$ sequence. $\beta$-sheet character of $\beta_1$ and $\beta_2$ sequences is weakened as $A_{\beta 42}$ monomer forms a contact to protofibril. On the contrary, $\beta_3$ sequences shows increase of $\beta$-sheet content after contact between monomer and protofibril. (b) Additional conformation of simulation time range a in Fig. 7. $\beta_1$ and $\beta_2$ sequences have partially helical conformations. (c) Additional conformations of simulation time range b in Fig. 7. $\beta_1$ sequence shows structural change from partially helical to fully helical conformation. $\beta_2$ sequence maintains partially helical conformations.
Time evolutions of $\beta$-sheet and $\alpha$-helix content of whole sequence

**Figure S13.** (a) $\beta$-sheet content of $A\beta_{42}$ monomer for whole sequence. Upper graphs are raw data curves and lower graphs are smoothened curves. After contact between monomer and protofibril motif, $\beta$-sheet content is slightly increased due to strengthened $\beta$-sheet character in $\beta_3$ sequence even though $\beta$-sheet contents of $\beta_1$ and $\beta_2$ sequences decrease. (b) $\alpha$-helix content of $A\beta_{42}$ monomer for whole sequence. $\alpha$-helix content increases after contact with protofibril motif as a result of helix formation in $\beta_1$ and $\beta_2$ sequences which allows increased interactions with sidechains on fibril surface.
Secondary structure of $\beta_2$ in $\alpha$-42 monomer and monomer-fibril distance

Figure S14. (a) Secondary structure of $\beta_2$ sequence of the monomer. It is shown that K28, G29, and A30 have low $\alpha$-helix content compared to other residues which demonstrates partially unfolded geometry of $\beta_2$ sequence. (b) Representative structure for $\beta_2$ sequence of $\alpha$-42 monomer. Unfolded geometry of residues 28-30 (red ribbon) allows close contact to fibril edge compared to residues 22-27 (black ribbon) and 31-33 (green ribbon). (c) Minimum heavy atom distance between the outermost chain in protofibril motif and the center of mass of each regions in $\beta_2$. It is shown that the distance of residues 28-30 is around 4 Å whereas residues 22-27 and 31-33 show longer distances.
Backbone hydrogen bond formation of $\beta_3$ between monomer and fibril

Figure S15. (a) $\beta$-sheet backbone hydrogen bonding formation in the C-terminal tail of $\beta_3$ sequence between $\text{A}\beta_{42}$ monomer and fibril motif (b) N-O distances for two hydrogen bonds shown in (a) designated by dotted lines and denoted by HB1 and HB2. Formation of hydrogen bond can be identified by the N-O distance less than 4 Å. For HB2, N-O distance curve frequently visits the threshold distance throughout the simulation except for early stage. On the other hand, HB1 rarely visits 4 Å until 200 ns. From this, it is suggested that the development of $\beta$-sheet structure in $\beta_3$ sequence occurs in order of HB2 $\rightarrow$ HB1.