Supplementary Information

Allosteric stabilization of the amyloid-β peptide hairpin by the fluctuating N-terminal

Liang Xu[a], Ruth Nussinov[b,c] and Buyong Ma[b]*

[a]School of Chemistry, Dalian University of Technology, Dalian, China

[b] Basic Science Program, Leidos Biomedical Research, Inc. Cancer and Inflammation

Program, National Cancer Institute, Frederick, MD 21702, USA

[c]Sackler Inst. of Molecular Medicine Department of Human Genetics and Molecular

Medicine Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Methods

The initial β-hairpin structure of Aβ(1–40) contains two β-strands formed by five backbone hydrogen bonds between Val^{18}-Ile^{31}, Phe^{20}-Gly^{29}, and Glu^{22}-Asn^{27}. The double mutants L17C and L34C were made to build the model of Aβ(1–40)cc, with the corresponding disulfide bond being 2.08 Å. The structural models of Aβ(1–34)cc and Aβ(17–40)cc were generated by removing the last six and first sixteen residues of Aβ(1–40)cc respectively. The starting structures were represented by Amber ff99SB force field parameters,[1] and solvated in a cubic box filled with TIP3P water molecules.[2] The distance between any protein atom and the boundary of box is at least 15 Å. One or three Na^+ ions were added to neutralize each system, resulting in 27934, 27912, and 17455 atoms for Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc systems, respectively.

In each REMD simulations,[3] 48 replicas were used and the temperature was exponentially distributed from 300 to 402 K (300.00, 301.93, 303.90, 305.85, 307.80, 309.77, 311.75, 313.74, 315.73, 317.74, 319.76, 321.79, 323.83, 325.88, 327.94, 330.00, 332.08, 334.17, 336.28, 338.39, 340.51, 342.65, 344.79, 346.96, 349.13, 351.31, 353.50, 355.70, 357.91, 360.14, 362.37, 364.62, 366.88, 369.15, 371.42, 373.71, 376.01, 378.33, 380.66, 383.00, 385.36, 387.72, 390.10, 392.49, 394.89, 397.30, 399.73, 402.17 K). The bond lengths in peptides and water molecules were constrained using the LINCS and SETTLE algorithms, respectively.[4] An integration step of 2 fs was applied, and long-range electrostatic interactions were treated using the PME method.[5] After 50000-steps minimization, each replica was heated to the desired temperature and equilibrated for 100 ps by MD simulation in the *NVT* ensemble. Protein and non-protein (water and ions) were separately coupled to an external heat bath with the coupling time constant of 0.1 ps using the velocity rescaling method.[6] The production run was carried out in the *NPT* ensemble, with the pressure kept at 1 bar with a coupling time constant of 2.0 ps using

the Parrinello-Rahman methd.[7] The production run of each replica was 200 ns, leading to 9.6-μs REMD simulations for each system. The exchange between two neighboring replicas was attempted every 5 ps, and the average acceptance ratio was about 23.2%, 24.2%, and 30.5% for system Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc, respectively. All REMD simulations were performed using Gromacs 4.5.3 software,[8] and the trajectory of 310 K (309.77 K) was collected for further analyses.

To examine the convergence of the REMD simulations, the cumulative average of β-sheet content was calculated (**Figure S1**),[9-12] which shows that each system reaches equilibrium after 100 ns. The average contents of β-sheet are 16.9%, 22.1%, and 11.1% for Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc, respectively. We further calculated the distribution of the radius of gyration (R_g), solvent accessible surface area (SASA), and secondary structures per residue at two different time intervals (100–150 ns and 150–200 ns) (**Figures S2 and S3**). The highly consistent distributions of R_g, SASA and secondary structures at different time intervals indicate the convergence of the present REMD simulations. To obtain representative conformations from each ensemble, the method as described by Daura et al.[13] was used to cluster structures based on a cutoff of 3.0 Å of the root-mean-square-deviation of Cα atoms.

The backbone conformational entropy loss from unfolded states to extended β-sheet conformation is estimated to be 4.78 cal/(mol·K·residue),[14] and the mean side-chain entropy loss is about 3.33 cal/(mol·K·residue).[15] On the other hand, a hydrogen bond in a β-sheet in water is about 1.58 kcal/mol.[16] Thus, the loss of entropy (including backbone and side-chain) due to formation of a β-sheet structure (2.51 kcal/mol·residue, T=310 K) cannot be totally compensated by the enthalpy of formation of a hydrogen bond (1.58 kcal/mol). If Aβ(17–34) constrained by the disulfide bond in the three peptides samples β-hairpin conformations with the same populations, we can estimate that the total entropy

loss of Aβ(17–34) is about 25.14 kcal/mol (consider 10 residues on average) for all peptides. The gain of enthalpy due to the formation of hydrogen bonds in the β-hairpin structures is about 15.80 kcal/mol. Consequently, if we assume that the potential energy of Aβ(17–34) relative to the unfolded state is the same for all peptides, the free energy change of Aβ(17–34) locked in a β-hairpin conformation is approximated to 9.34 kcal/mol.

Table S1. Retained water molecules within 10 Å of Aβ(17–34) fragment of Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc structures.

System	0–100 ns	100–200 ns	200–300 ns	300–400 ns
Aβ(1–40)cc	280.7 (14.0)	277.7 (13.5)	262.9 (14.4)	259.6 (12.7)
Aβ(1–34)cc	293.1 (16.7)	289.3 (12.8)	277.6 (14.8)	282.1 (14.9)
Aβ(17–40)cc	307.6 (17.3)	295.7 (15.6)	289.8 (15.2)	275.0 (15.9)
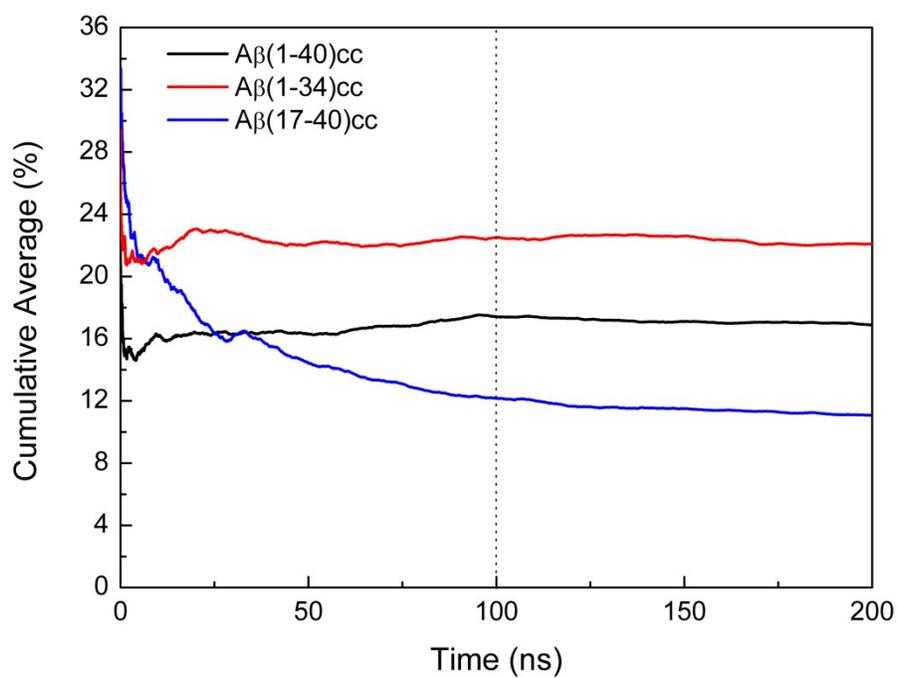
Figure S1. The cumulative average of β-sheet content of Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc throughout the 200-ns simulation.
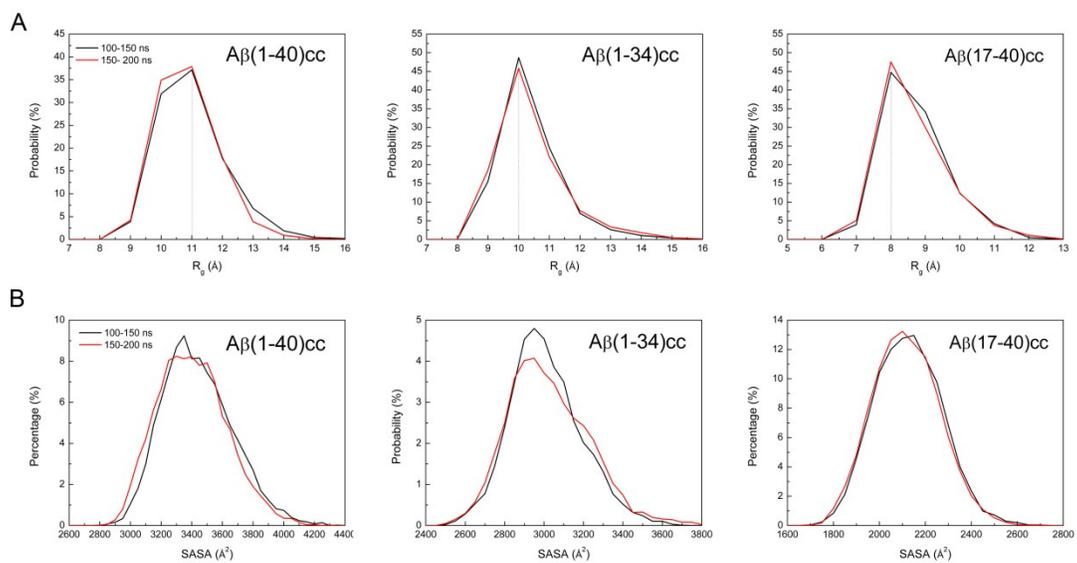
Figure S2. The distributions of radius of gyration R_g (A) and solvent accessible surface area SASA (B) at time intervals of 100-150 ns and 150-200 ns.
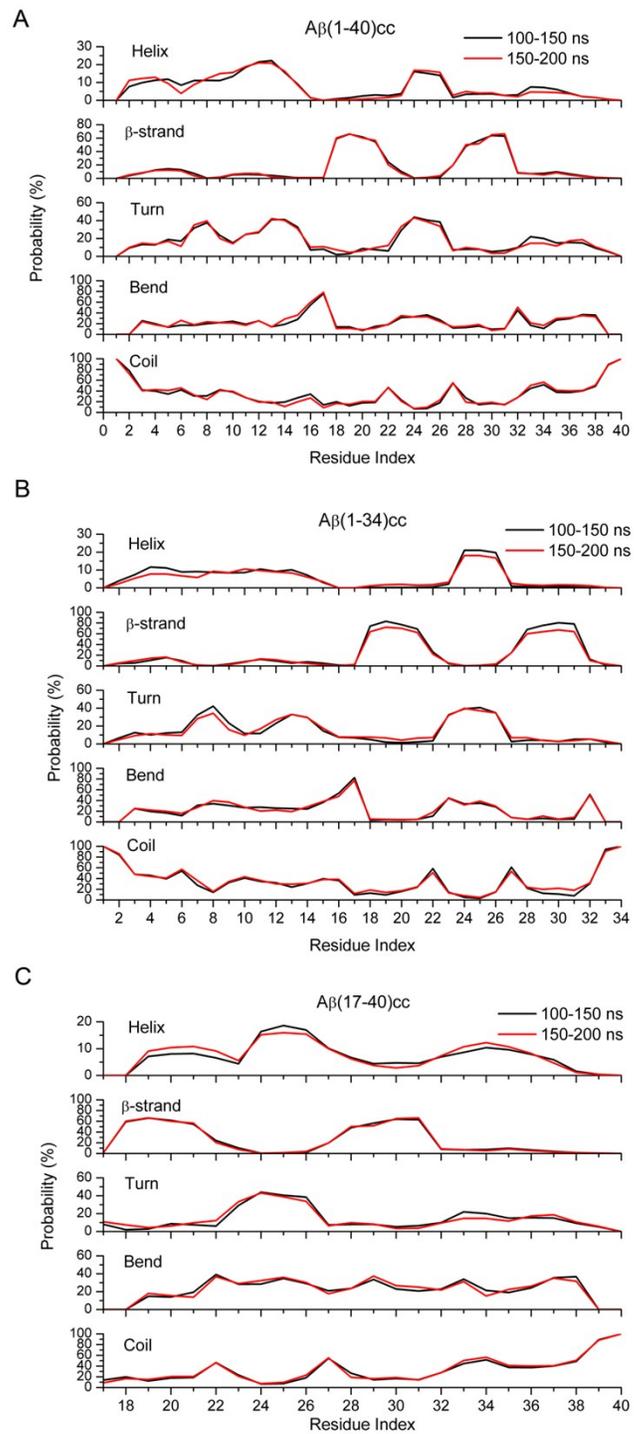
Figure S3. The per residue distribution of secondary structures at time intervals of 100-150 ns and 150-200 ns.
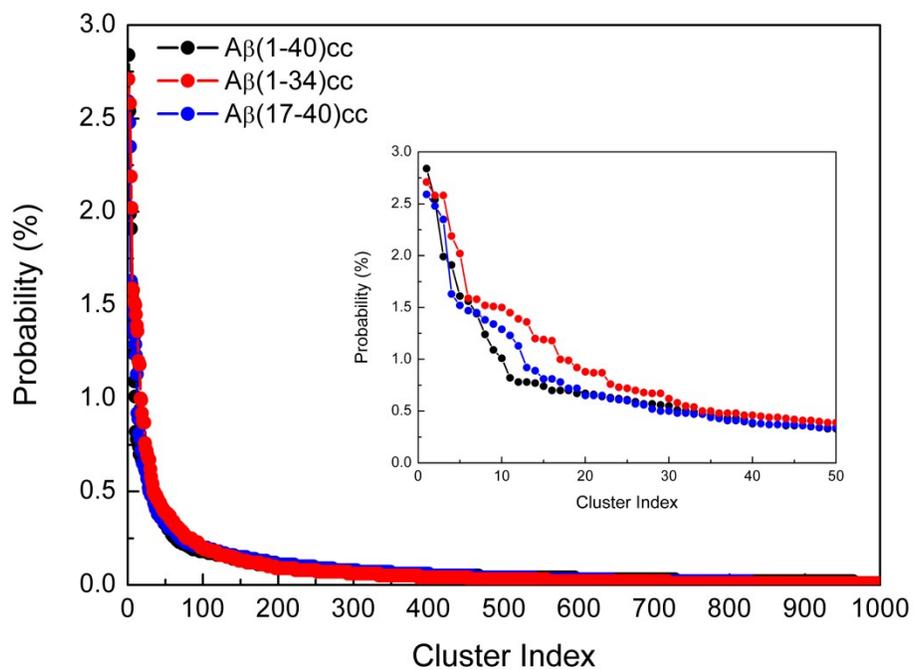
Figure S4. Cluster distribution of Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc. The distribution of the first 50 clusters is also shown in Inset.
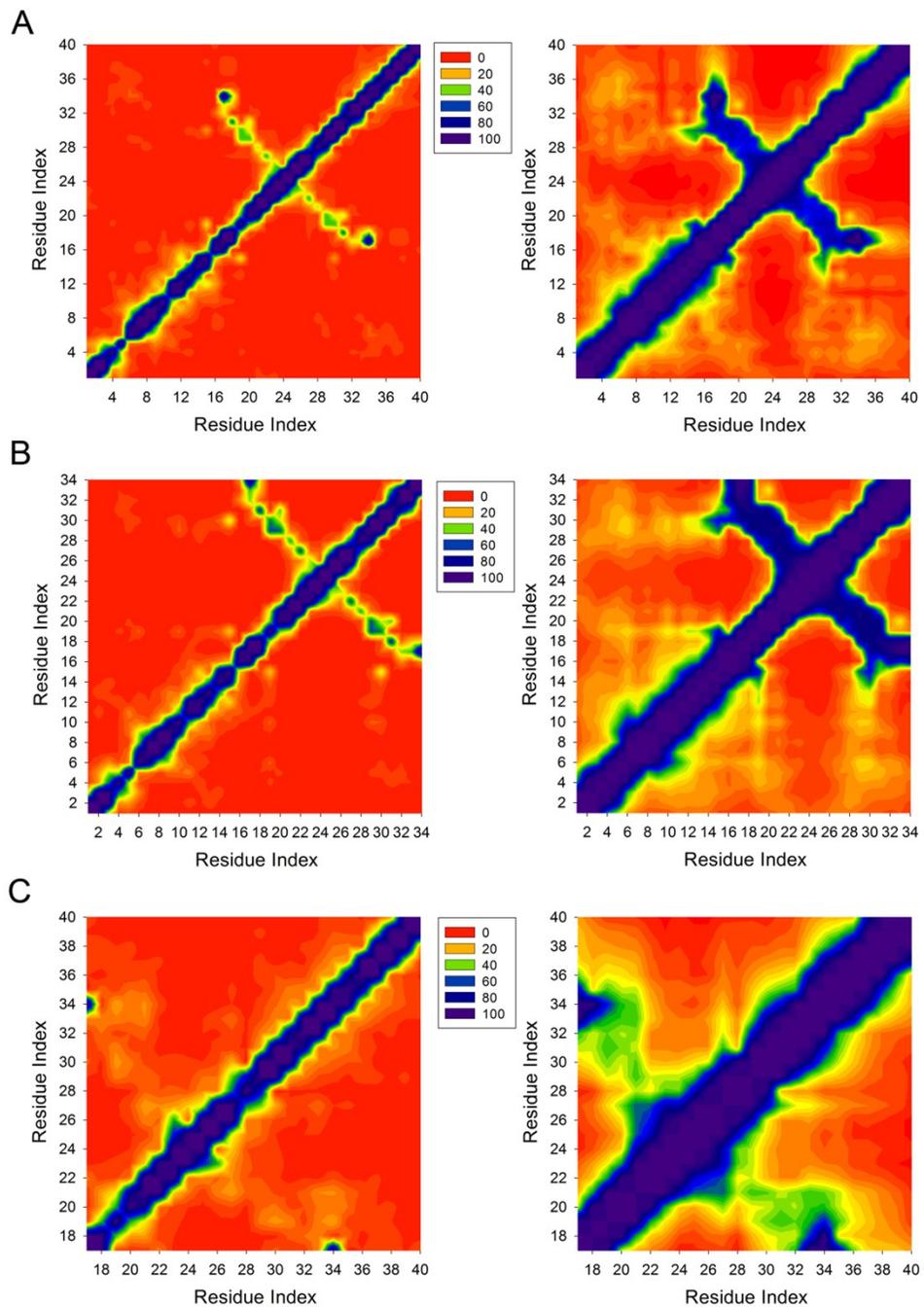
Figure S5. Intra-molecular contact maps. Contact frequencies (%) between amino acids in Aβ(1–40)cc (A), Aβ(1–34)cc (B), and Aβ(17–40)cc (C). A contact occurs if the center of mass of each residue is within 6 Å (left column) or 10 Å (right column) of the center of mass of another residue.
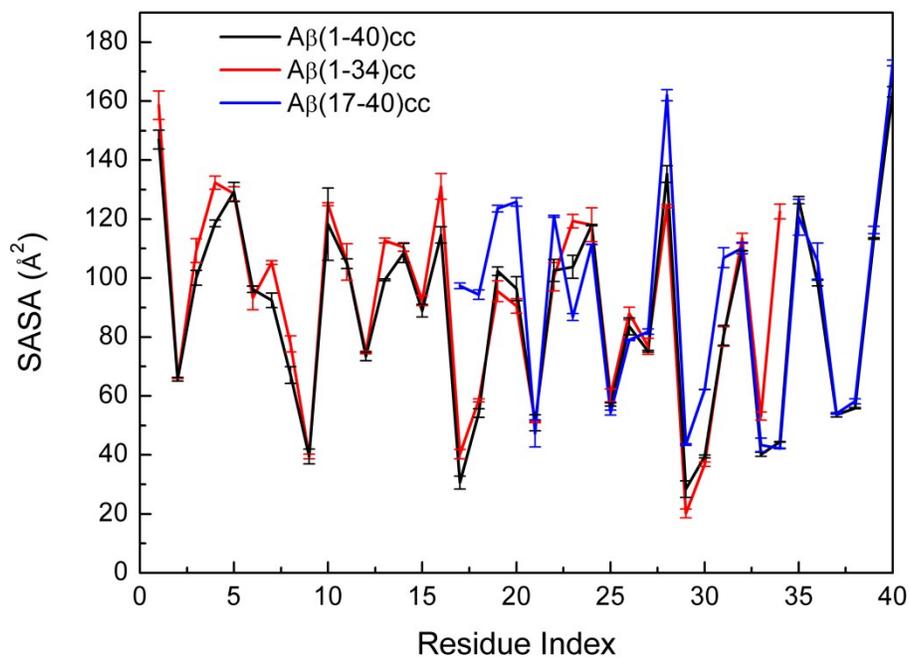
Figure S6. The solvent accessible surface area (SASA) calculated for each residue of Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc. The standard deviations were obtained by averaging the results of 100–150 ns and 150–200 ns.
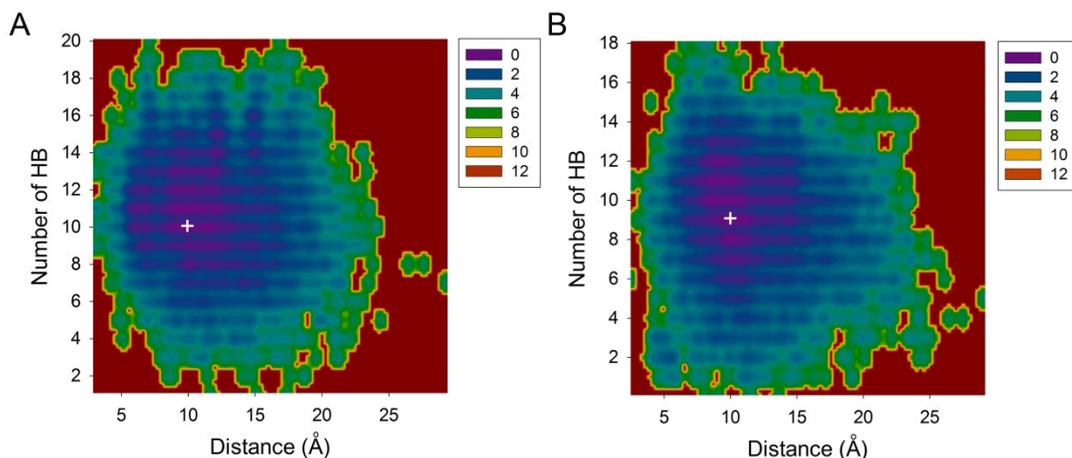
Figure S7. Free energy surface of Aβ(1–40)cc (A) and Aβ(1–34)cc (B) at 310 K in terms of the distance between the center of mass of Aβ(1–16) fragment and the center of mass of Aβ(17–34) fragment, and the number of backbone hydrogen bonds corresponding to the number of residues involving the formation of β-hairpin conformation. The free energy values (in kcal/mol) were obtained by $\Delta G = -k_B T (\ln P_i - \ln P_{max})$, where P_i and P_{max} are the probability distributions calculated for specific pairs of distance and number of hydrogen bonds. $\ln P_i - \ln P_{max}$ was used to ensure $\Delta G = 0$ for the lowest free energy point (white cross).
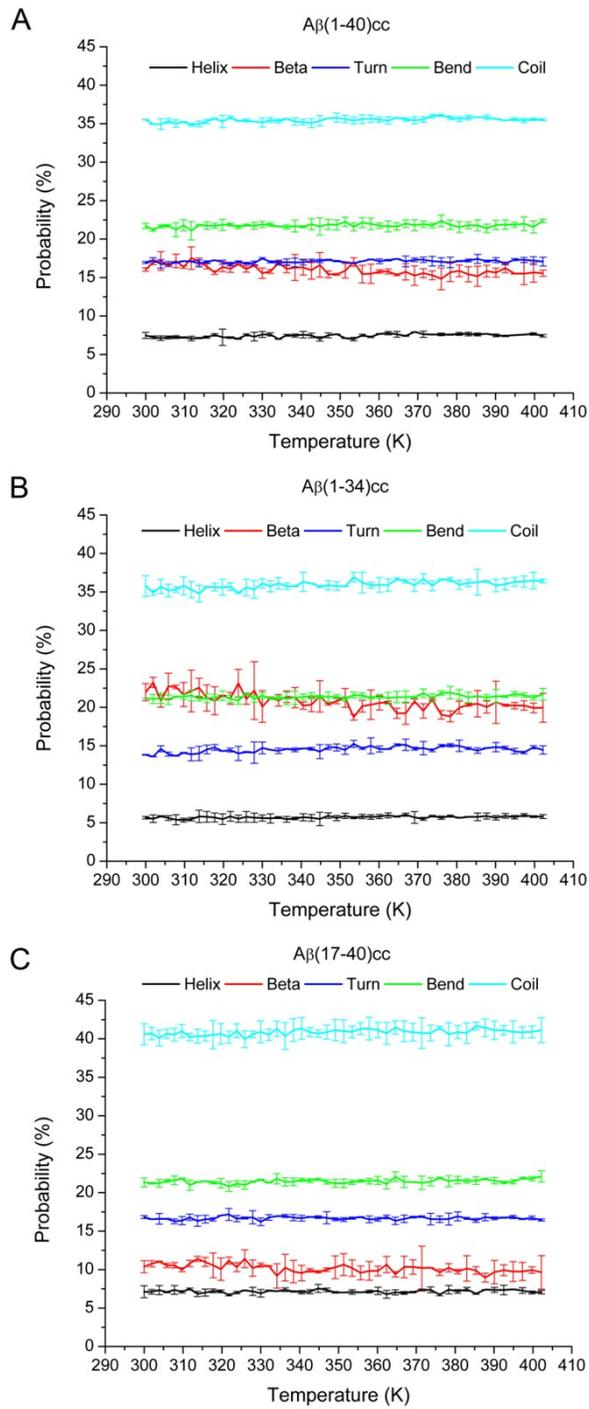
Figure S8. The average secondary structures of Aβ(1–40)cc (A), Aβ(1–34)cc (B), and Aβ(17–40)cc (C) in different temperatures.
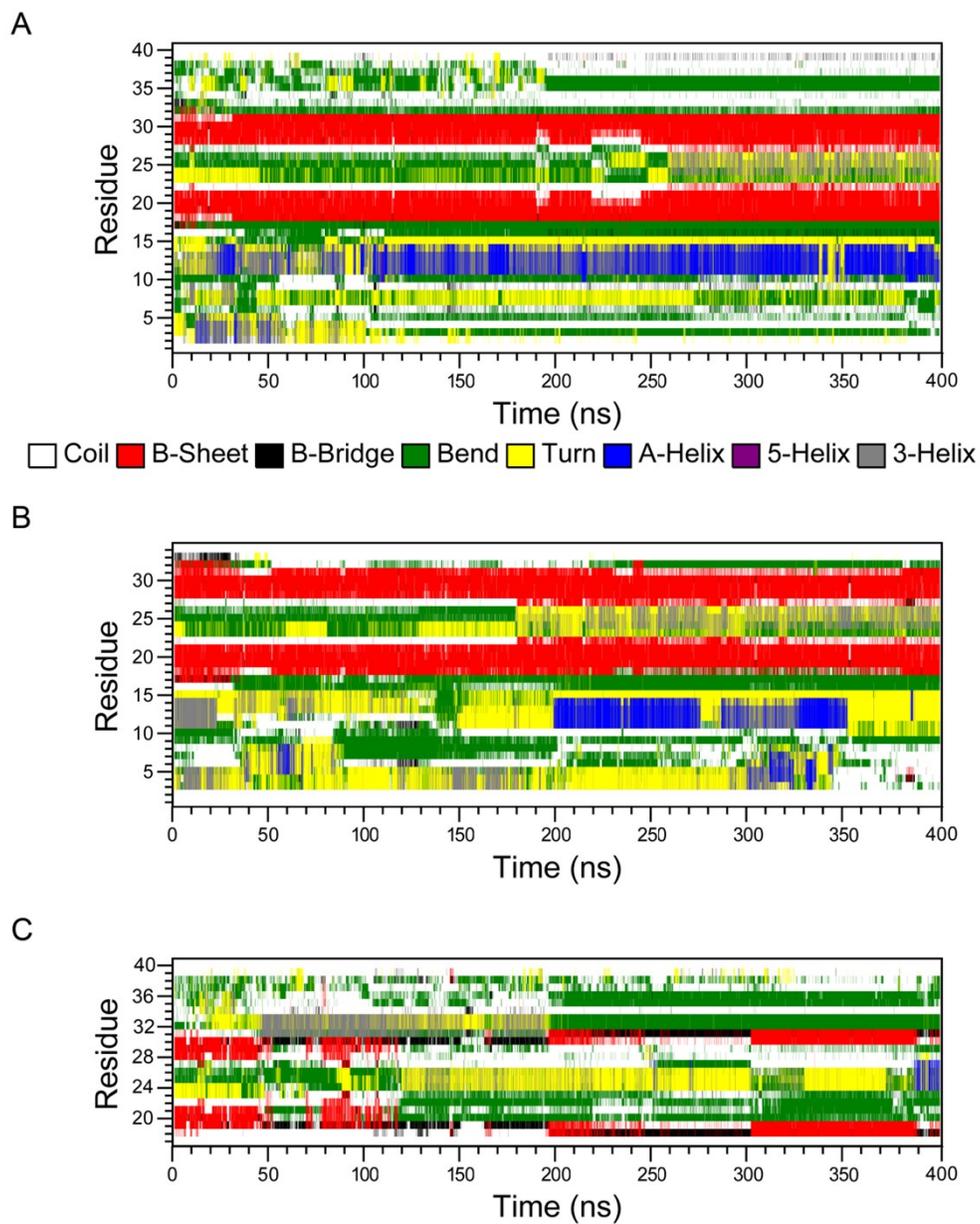
Figure S9. The secondary structures of Aβ(1–40)cc (A), Aβ(1–34)cc (B), and Aβ(17–40)cc (C) over the 400-ns conventional MD simulations at 310 K. All simulations were performed under the same condition as the replica in aqueous solution at 310 K, with each simulation being 400 ns.
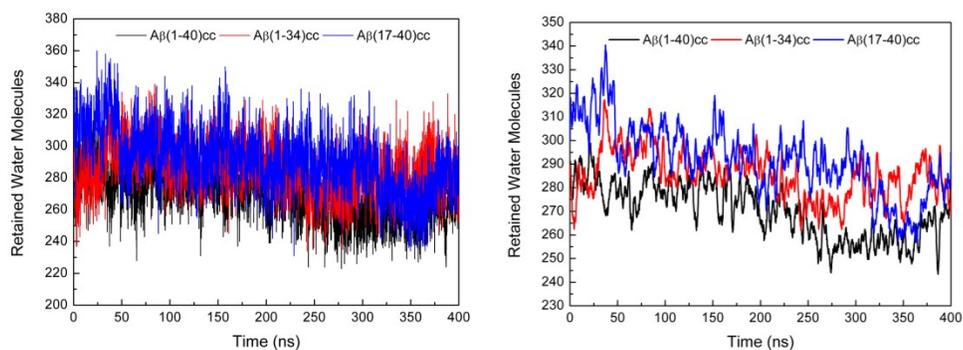
Figure S10. The number of the retained water molecules within 10 Å of Aβ(17–34) fragment of Aβ(1–40)cc, Aβ(1–34)cc, and Aβ(17–40)cc structures (A). The values in terms of adjacent average over 20 data points are shown in (B).

References

1. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Structure, Function, and Bioinformatics*, 2006, **65**, 712-725.
2. P. Mark and L. Nilsson, *The Journal of Physical Chemistry A*, 2001, **105**, 9954-9960.
3. Y. Sugita and Y. Okamoto, *Chemical Physics Letters*, 1999, **314**, 141-151.
4. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *Journal of Computational Chemistry*, 1997, **18**, 1463-1472.
5. T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1993, **98**, 10089.
6. G. Bussi, D. Donadio and M. Parrinello, *J Chem Phys*, 2007, **126**, 014101.
7. M. Parrinello, *Journal of Applied Physics*, 1981, **52**, 7182.
8. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, **4**, 435-447.
9. J. Khandogin and C. L. Brooks, *Proceedings of the National Academy of Sciences*, 2007, **104**, 16880-16885.
10. L. Xu, X. Wang and X. Wang, *ACS Chemical Neuroscience*, 2013, **4**, 1458-1468.
11. L. Xu, X. Wang, S. Shan and X. Wang, *Journal of Computational Chemistry*, 2013, **34**, 2524-2536.
12. L. Xu, K. Gao, C. Bao and X. Wang, *Journal of Computer-Aided Molecular Design*, 2012, **26**, 963-976.
13. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angewandte Chemie International Edition*, 1999, **38**, 236-240.
14. J. B. Thompson, H. G. Hansma, P. K. Hansma and K. W. Plaxco, *Journal of Molecular Biology*, 2002, **322**, 645-652.
15. A. J. Doig and M. J. E. Sternberg, *Protein Science*, 1995, **4**, 2247-2251.
16. S. Y. Sheu, D. Y. Yang, H. L. Selzle and E. W. Schlag, *Proceedings of the National Academy of Sciences*, 2003, **100**, 12683-12687.