Supplementary Information:

Disclosing the Mechanism of Spontaneous Aggregation and Template-Induced Misfolding of the Key Hexapeptide (PHF6) of Tau Protein Based on Molecular Dynamics Simulation

Hongli Liu[1], Haiyang Zhong[2], Xuewei, Liu[2], Shuangyan Zhou[1,3], Shuoyan Tan[1], Huanxiang Liu[1*] Xiaojun Yao[2,4]

[1] School of Pharmacy, Lanzhou University, Lanzhou 730000, China

[2] State Key Laboratory of Applied Organic Chemistry and Department of Chemistry, Lanzhou University, Lanzhou 730000, China

[3]Chongqing Key Laboratory on Big Data for Bio Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[4] State Key Laboratory of Quality Research in Chinese Medicine, Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Taipa, Macau 999078, China

* Corresponding author

Tel.: +86-931-891-5686

Fax: +86-931-891-5686

E-mail address: hxliu@lzu.edu.cn

Table S1. Details of simulation systems.

Systems	Simulation time (μs)	Total number of atoms	Box size (Å)
16 chains[a]	1	54733	83 × 85 × 93
PHF6_L[b]	1	61352	92 × 87 × 104
PHF6_R[c]	1	71219	94 × 93 × 96

[a]Spontaneous aggregation system, the system including sixteen PHF6 monomers.

[b]PHF6_L means the PHF6 monomer on the left of the template.

[c]PHF6_R means the PHF6 monomer on the right of the template.

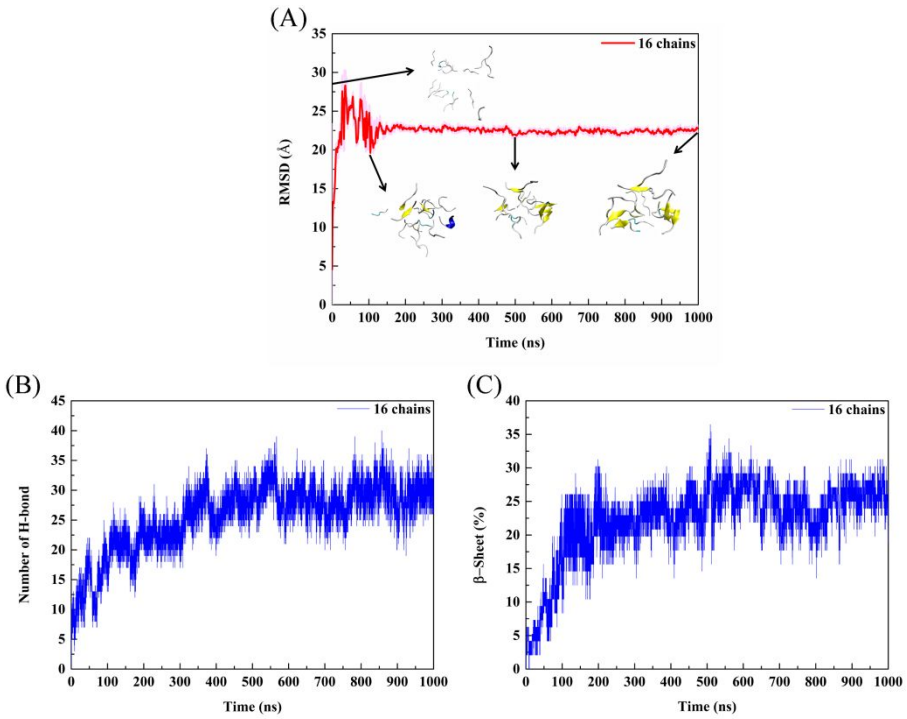

Figure S1.Verification of convergence of 16 chains system. (A) Time evolution of the RMSD of protein backbone atoms of 16 chains system. Representative structures were extracted from 0 ns, 100 ns, 500 ns, and 1 μs and shown in panel A. (B) and (C) are the number of backbone hydrogen bonds and the β-sheet content of 16 chains, respectively.

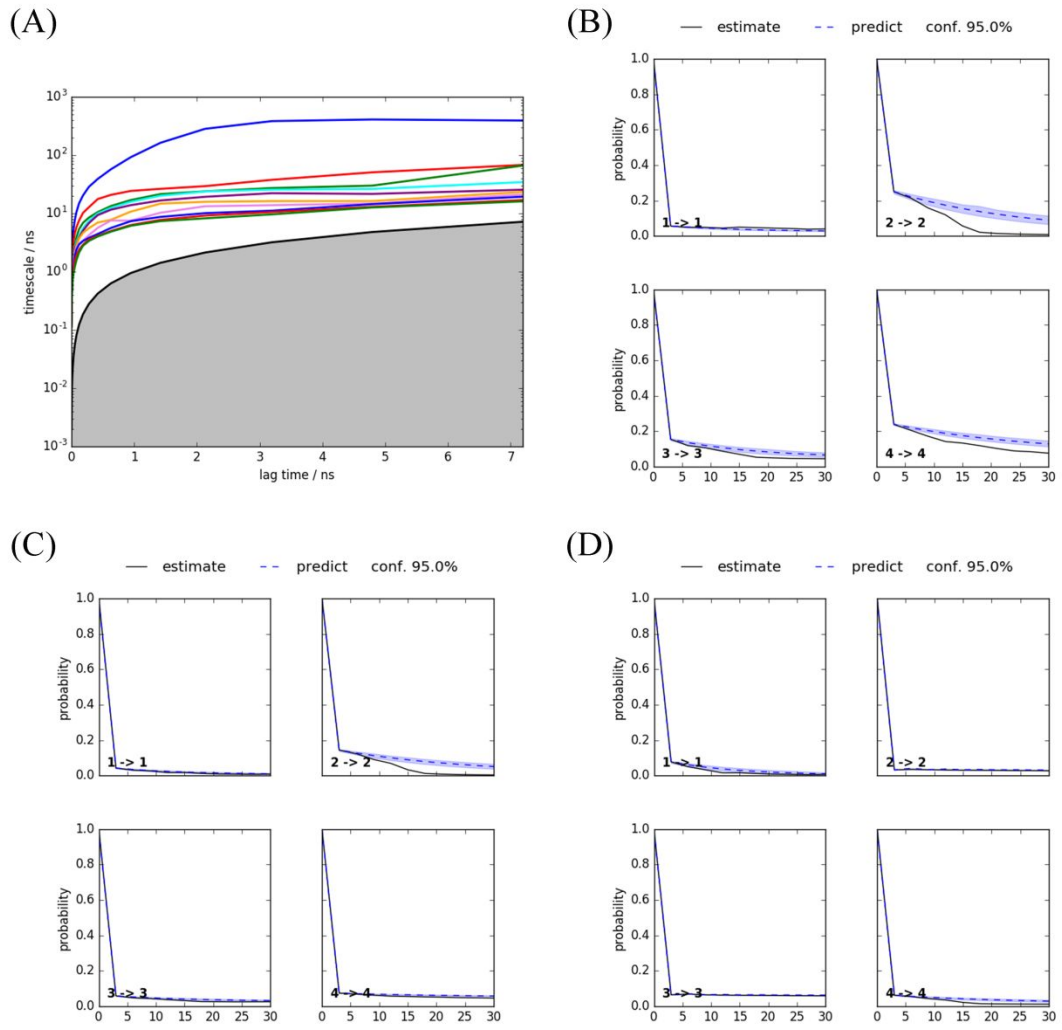

Figure S2. Validation of the Markov state model. (A) Implied relaxation timescales as a function of the lag time τ indicates that a lag time τ = 3 ns is appropriate. (B)-(D) are the Chapman-Kolmogorov test curves obtained by lumping 300 microstate into 10, 12, and 15 macrostates, respectively.

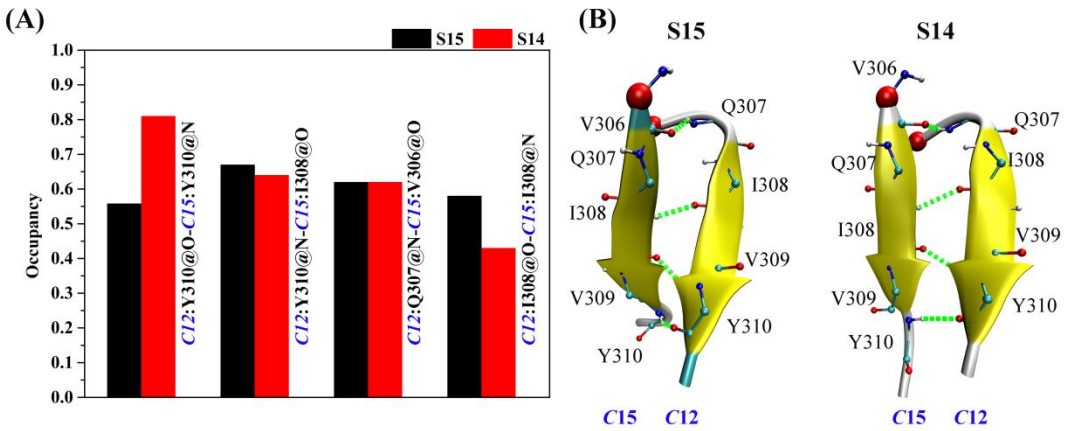

Figure S3. Detailed interactions between *C*12 and *C*15 of S15 and S14. (A) Hydrogen bond occupancy between the two peptide chains in S15 and S14. (B) Representative structures of S15 and S14, green dashed lines represent backbone hydrogen bonds.

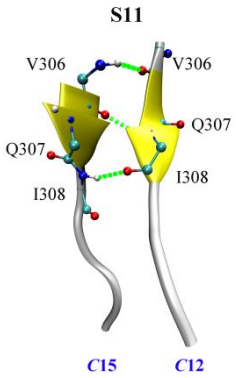

Figure S4. Representative structure of S11, green dashed lines represent backbone hydrogen bonds.

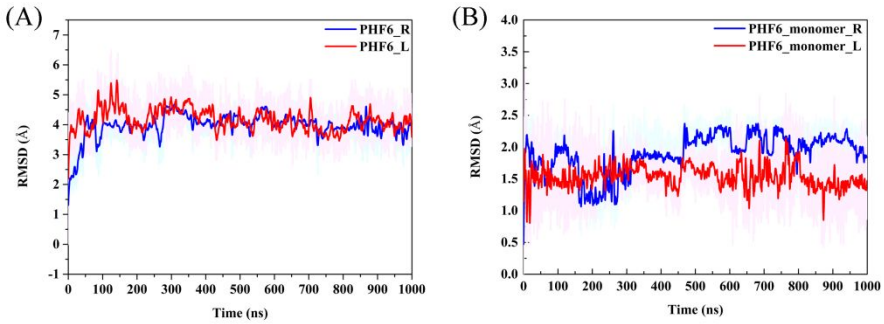

Figure S5. Verification of convergence of template-induced systems. (A) and (B) are time evolution of the RMSDs of protein backbone atoms of template-PHF6 complex and PHF6 monomer, respectively.

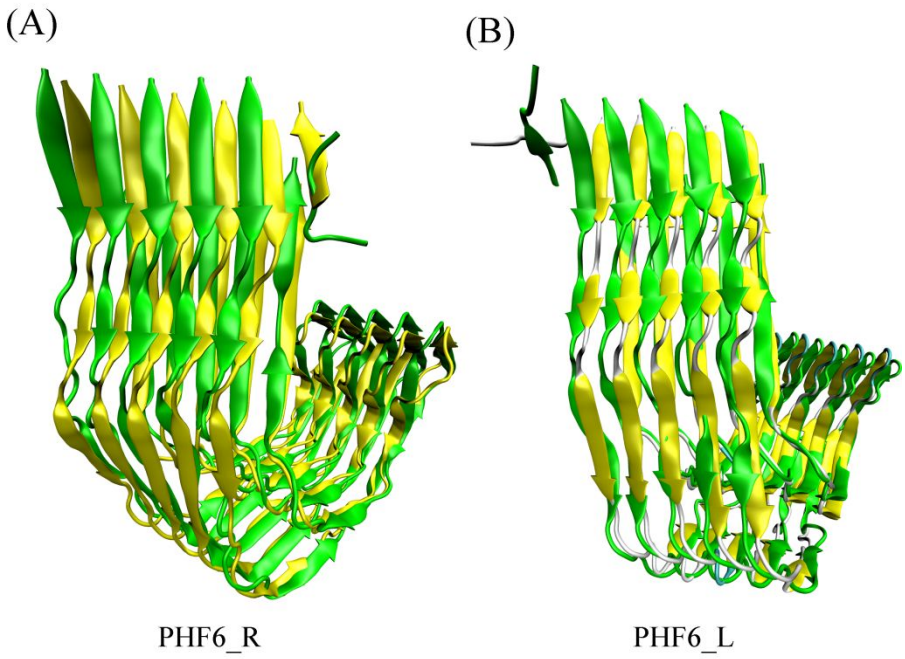

Figure S6. (A) and (B) are the superposition of the first two representative structures from the cluster analysis of the PHF6_R and PHF6_L systems, respectively. The green and yellow cartoon structures represent the first and second cluster structures, respectively.