

Supporting Material

This material includes additional methods and 16 supplemental figures

Force field selection

As reported previously, the four widely used force fields including AMBER, CHARMM, OPLS, and GROMOS performs well in general, but they and their different versions display different secondary structure propensities ¹⁻³. For example, GROMOS96 favors β -sheet structures for two peptide monomers ¹ and the A β (16-22) dimers ²; CHARMM27 cannot fold the all- β protein WW domain due to overstabilization of α -helical structures ⁴; CHARMM22* was found not to be able to stabilize the native state of Engrailed homeodomain ⁵; AMBER ff99SB generates significantly lower α -helix content of poly-Ala-based peptides than experiments ⁶, and it also understabilizes the stability of β -hairpin peptide Trpzip2 ⁷; OPLS can not fold the fast-folding protein villin in 10- μ s MD simulations ³. Therefore, a suitable force field should be selected for the simulated

peptide system before MD simulations are launched. Indeed, we made some tests using three different force fields including CHARMM22*, AMBER99SB-ILDN and GROMOS 53A6, starting from preformed 3+3 and 4+4 parallel bilayer β -sheet. We found that the parallel bilayer β -sheet collapsed and most of the chains become coil conformations in the two independent MD runs using CHARMM22* and AMBER99SB-ILDN (data not shown), while GROMOS 53A6 results in diverse β -sheet-rich hexamers and stable 4+4 parallel bilayer β -sheet (Figs 7, 9).

The impact of force fields on the stability of the preformed bilayer β -sheet

To choose a proper force field, we performed six independent MD simulations using three different force fields including CHARMM22*, AMBER99SB-ILDN and GROMOS 53A6, starting from preformed 3+3 and 4+4 parallel bilayer β -sheet. We found that GROMOS 53A6 results in diverse β -sheet-rich hexamers and stable 4+4 parallel bilayer β -sheet (Figs. S12, S15). However, the parallel bilayer β -sheet collapsed and most of the chains become coil conformations in all of the MD runs using CHARMM22* and AMBER99SB-ILDN (Figs. S10, S11, S13, S14). Therefore, at this stage we cannot conclude whether CHARMM22* and AMBER99SB-ILDN force fields reflect more in vitro reality since all force fields have been parametrized for folded proteins. These MD simulation results, together with the experimental observation of high fibril formation propensity of GVGIAQ peptide, indicate that GROMOS 53A6 is a suitable force field for the GVGIAQ fragment of SOD1 protein.

The importance of I149 in maintaining hexamer stability

I149T mutation has been observed in full length SOD-1, and I149T mutant fragment GVTGIAQ is able to form amyloid. We checked the hexamer stabilities of two mutants (I149P and I149A) to illustrate the importance of I149 to maintain fibril structures. The I149P parallel bilayer β -sheet collapsed completely and most of the chains randomized in the 100-ns and 200-ns MD simulations (Fig. 7b, d and Fig. S6). We then performed two 200-ns MD simulations of I149A to examine the effect of other amino acid mutation on the structure and aggregation of GVGIAQ peptides (Fig. S9). The largest C α -RMSD values of 1.01 nm (I149A_1) and 1.25nm (I149A_2) combined with the average C α -RMSD values of 0.67 nm and 0.87 nm in two MD runs indicate that the preformed I149A mutant is not stable (Fig. S9a, c). Transitions from the bilayer β -sheet structure to a

collapsed coil-rich state are observed in Fig. S9b, d, indicating that substitution of isoleucine 149 by alanine results in disordered structures.

A total of 3 figures (Figs. S1, S2, S3) show the initial random states and the convergence check for the REMD runs of the wild type and I149P systems at 310 K, including the probability of end-to-end distance and the number of hydrogen bonds, secondary structure contents (including all structures, coil, β -sheet, β -bridge, and bend) of each residue and time evolution of the first replica exchange in both systems.

Figure S4 shows the hydrophobic and hydrophilic average solvent accessible surface areas (SASA) of each residue and side-chain to get insight into the relative distribution of hydrophobic side chains on the inner and outer sides. As seen from Fig. S4a, the hydrophobic values of residues G147, G150, I151 and Q153 increase from 1.18, 0.59, 3.06 and 1.69 nm² to 1.27, 0.64, 3.79 and 1.72 nm², while V148 and A152 decrease from 3.22, 1.75 nm² to 2.93, 1.39 nm². The hydrophobic value of isoleucine is much higher than that of proline. We also found in Fig. S4c that the value of hydrophobic SASA in I149P hexamers is larger than that in wild type. When focusing on hydrophilic SASA of each residue (Fig. S4b), we found that the hydrophilic values of most residues decrease after substitution excluding G147 (from 0.99nm² to 1.49 nm²) and I(P)149. We know that the hydrophilic index of isoleucine is larger than that of proline, from 4.5 to 1.6. Naturally, the hydrophilic SASA in wild type is larger than that in I149P hexamers (Fig. S4d). These results indicate that the SASA changes after mutation and more hydrophobic side chains are located on the exterior in I149P hexamers relative to the wild type.

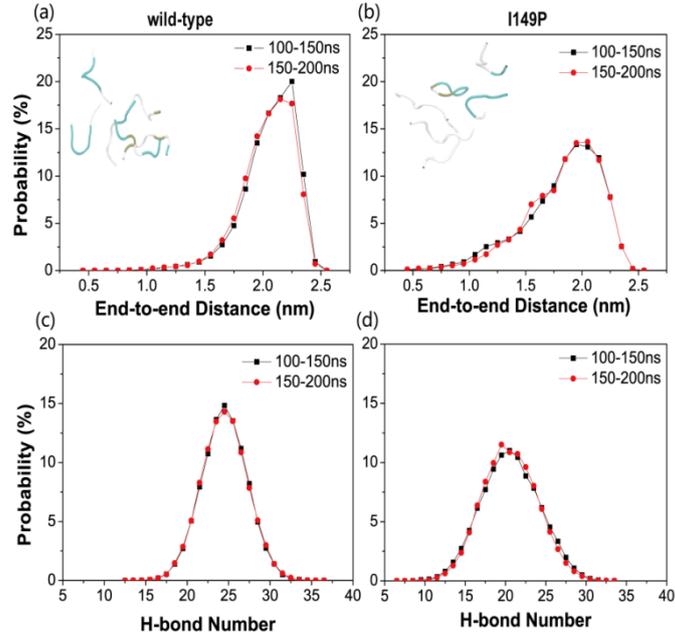


Figure S1. The initial random states and convergence check for the REMD runs of wild type and I149P systems at 310 K. Initial state and probability of end-to-end distance (the Ca-Ca distance between G147 and Q153) for the REMD run of wild type (a) and I149P (b) systems. Probability of number of hydrogen bonds in wild type (c) and I149P hexamers (d) within the time intervals of 100-150ns and 150-200ns.

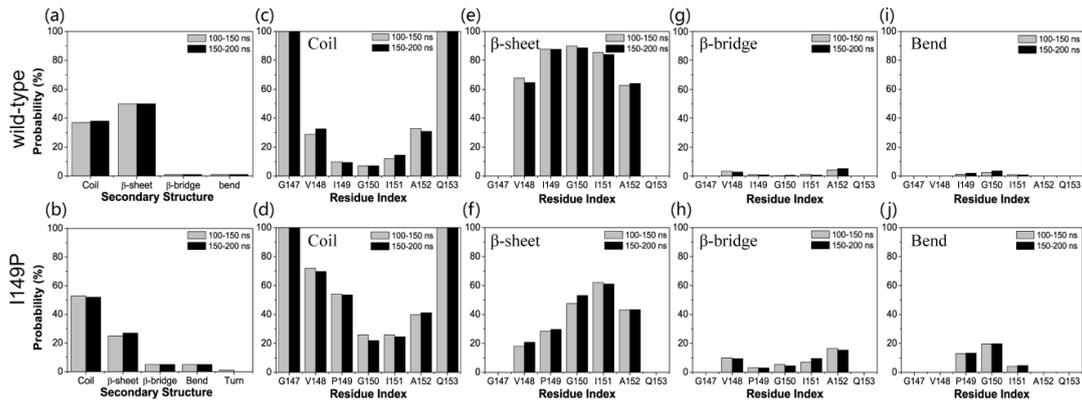


Figure S2. The calculated secondary structure probability of each residue in the REMD runs within the time intervals of 100-150 ns and 150-200 ns for wild type and I149P systems at 310 K for (a, b) all structures, (c, d) coil, (e, f) β -sheet, (g, h) β -bridge and (i, j) bend.

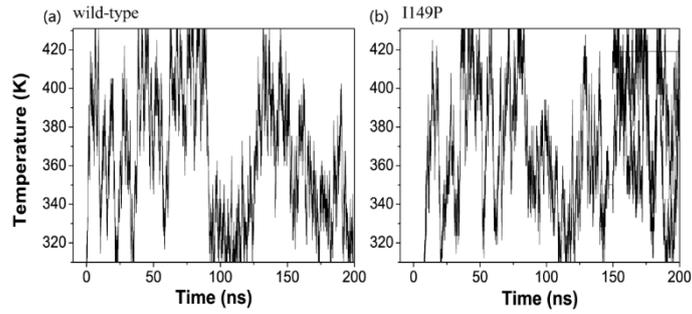


Figure S3. Time evolution of temperature swapping for the first replica at 310 K in wild type (a) and I149P hexamers (b).

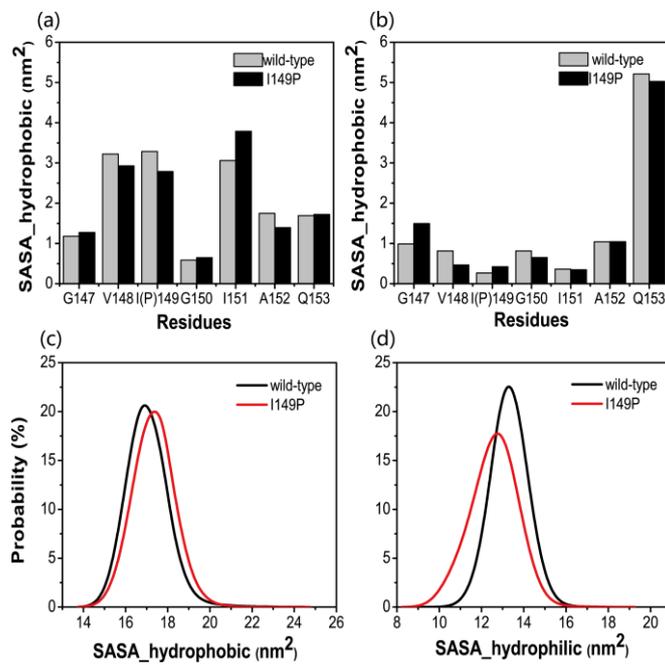


Figure S4. The probability distribution of solvent accessible surface area (SASA) of hydrophobic and hydrophilic for the residues (a, b) and side-chains (c, d) in wild type and I149P hexamers at 310K.

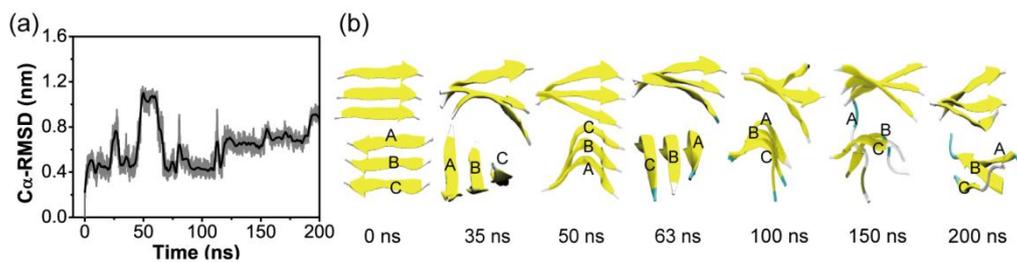


Figure S5. Analysis of the structural stability of 3+3 fibril-like parallel bilayer β -sheet of GVIGIAQ hexamer. The parallel bilayer β -sheet is not stable. Chain A in the initial parallel bilayer β -sheet dissociates from chain B at around $t=150$ ns and forms an antiparallel β -sheet with chain B, leading to the formation of a mixed parallel-antiparallel bilayer β -sheet. (a) Time evolution of Ca-RMSD with respect to the 3+3 parallel bilayer β -sheet at $t = 0$ ns and (b) representative snapshots at seven different time points.

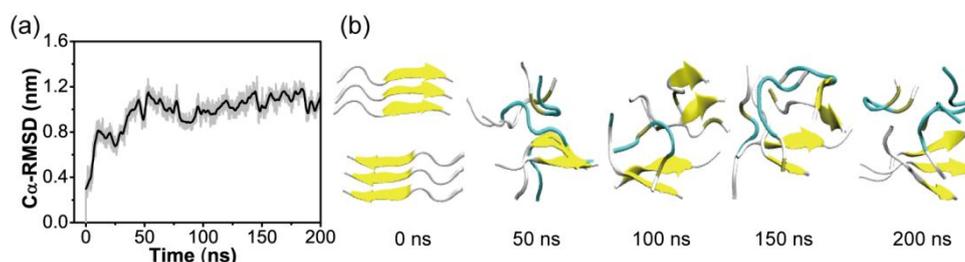


Figure S6. Analysis of the structural stability of 3+3 bilayer β -sheet for I149P mutant hexamers. The Ca-RMSD value increases quickly in the beginning of the simulation and reaches 1.0 nm within the first 50 ns. This large Ca-RMSD value, together with the disordered snapshots in Figure S6b, indicate that the I149P mutation disrupts the parallel 3+3 bilayer β -sheet. (a) Time evolution of Ca-RMSD with respect to the 3+3 bilayer β -sheet at $t = 0$ ns and (b) representative snapshots at five different time points.

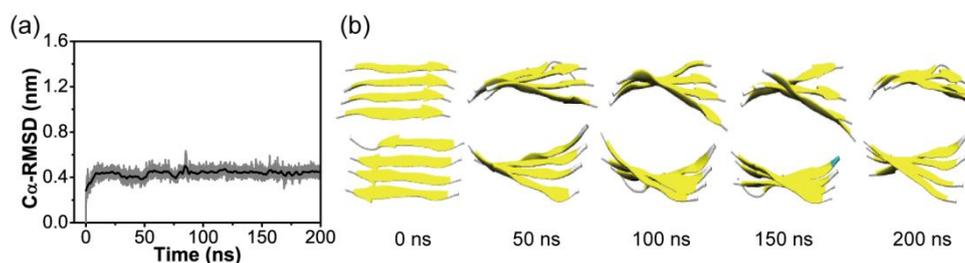


Figure S7. Analysis of the structural stability of the 4+4 fibril-like parallel bilayer β -sheet of GVIGIAQ peptide. The value of Ca-RMSD stays at 0.4 nm during almost all of the simulation period, which is quite similar to the result in Fig. 9. The 4+4 parallel bilayer β -sheet exist in the whole 200-ns simulation, with a twisting of β -sheet in two layers. (a) Time evolution of

Ca-RMSD with respect to the 4+4 bilayer β -sheet at $t = 0$ ns and (b) representative snapshots at five different time points.

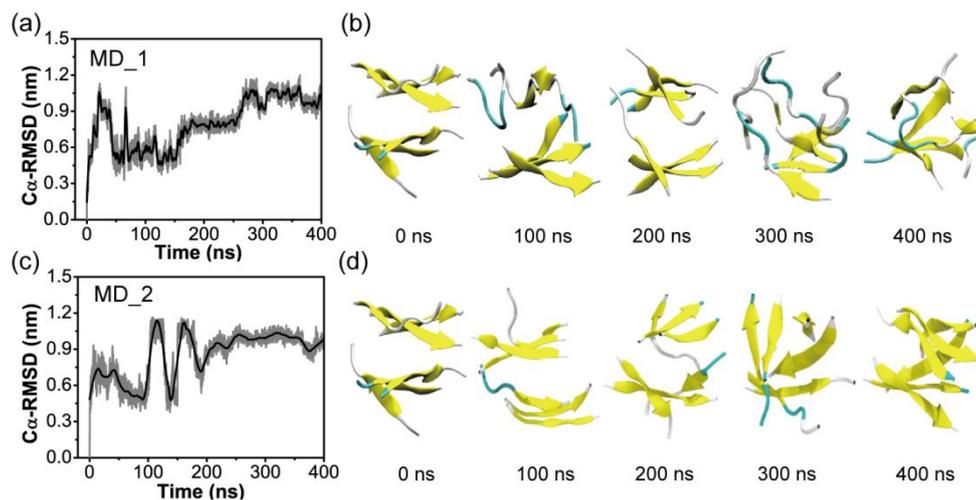


Figure S8. The time evolution of Ca-RMSD with respect to the mixed 3+3 parallel-antiparallel bilayer β -sheet at $t = 200$ ns in Fig. 7(c) in the two MD trajectories (a, c) and their representative snapshots at 5 different time points (b, d). The transition from the mixed parallel-antiparallel bilayer β -sheet to the fibril-like parallel bilayer β -sheet was not observed, indicating a higher energy barrier between the two states.

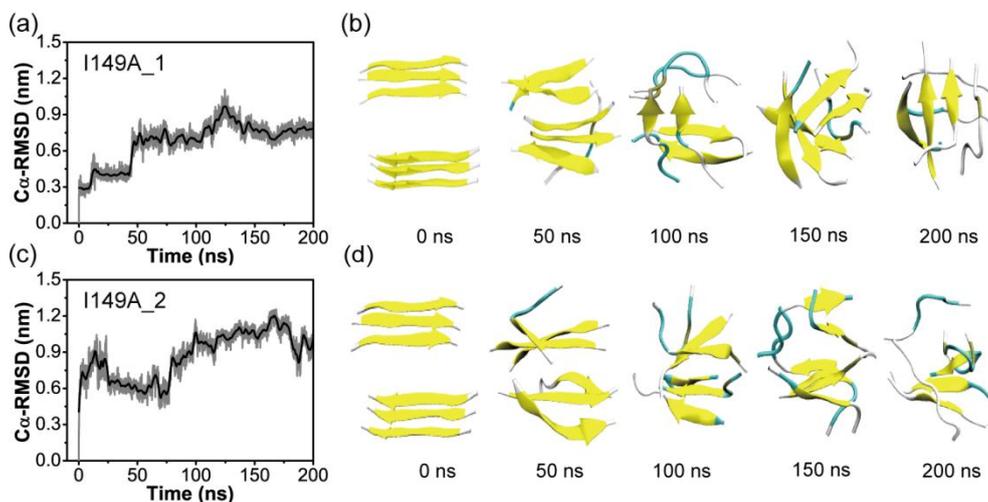


Figure S9. Results of the structural stability of the 3+3 fibril-like parallel bilayer β -sheet of I149A mutant hexamers. The largest C α -RMSD values of 1.01 nm (I149A_1) and 1.25nm (I149A_2) combined with the average C α -RMSD values of 0.67 nm and 0.87 nm in two MD runs indicate that the preformed I149A mutant is not stable. (a, c) Time evolution of C α -RMSD with respect to the 3+3 bilayer β -sheet at t = 0 ns and (b,d) their representative snapshots at five different time points.

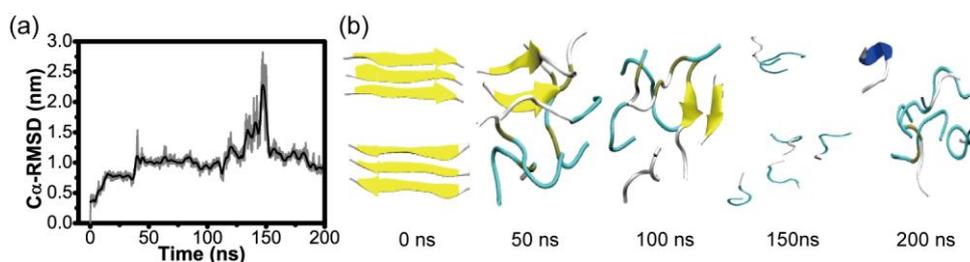


Figure S10. The time evolution of C α -RMSD with respect to the initial structure of 3+3 bilayer beta-sheet using CHARMM22* force field (a) and representative snapshots at 5 different time points (b).

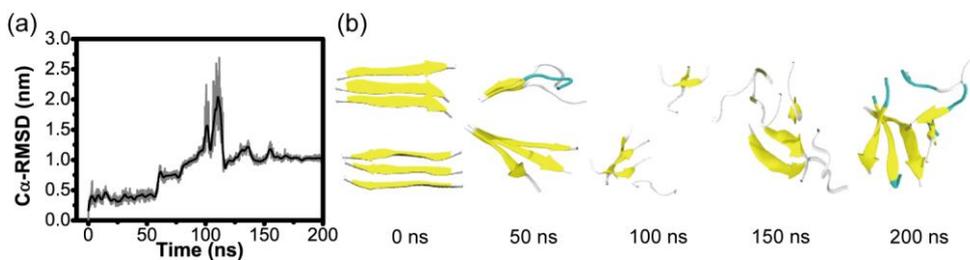


Figure S11. The time evolution of C α -RMSD with respect to the initial structure of 3+3 bilayer beta-sheet using AMBER99SB-ILDN force field (a) and representative snapshots at 5 different time points (b).

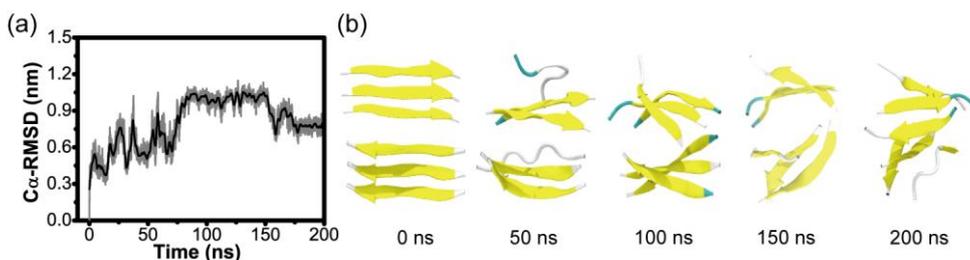


Figure S12. The time evolution of Ca-RMSD with respect to the initial structure of 3+3 bilayer beta-sheet using GROMOS-53A6 force field (a) and representative snapshots at 5 different time points (b).

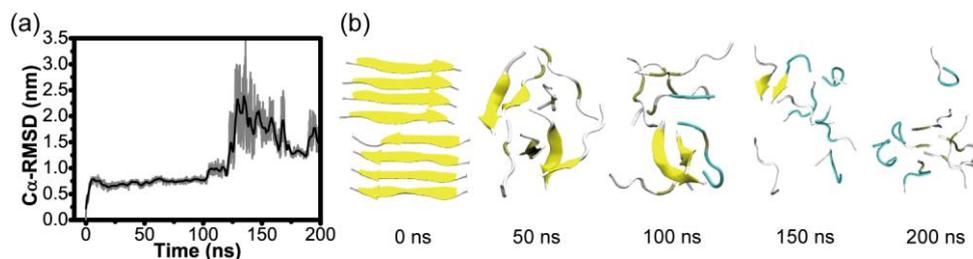


Figure S13. The time evolution of Ca-RMSD with respect to the initial structure of 4+4 bilayer beta-sheet using CHARMM22* force field (a) and representative snapshots at 5 different time points (b).

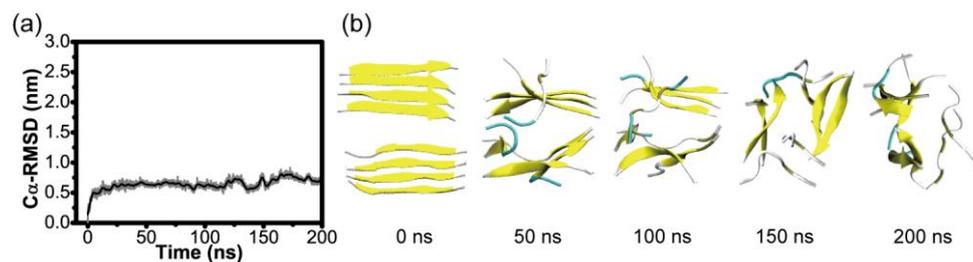


Figure S14. The time evolution of Ca-RMSD with respect to the initial structure of 4+4 bilayer beta-sheet using AMBER99SB-ILDN force field (a) and representative snapshots at 5 different time points (b).

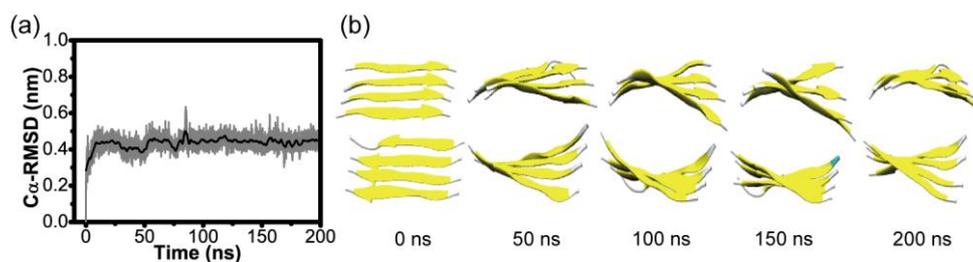


Figure S15. The time evolution of Ca-RMSD with respect to the initial structure of 4+4 bilayer beta-sheet using GROMOS-53A6 force field (a) and representative snapshots at 5 different time

points (b).

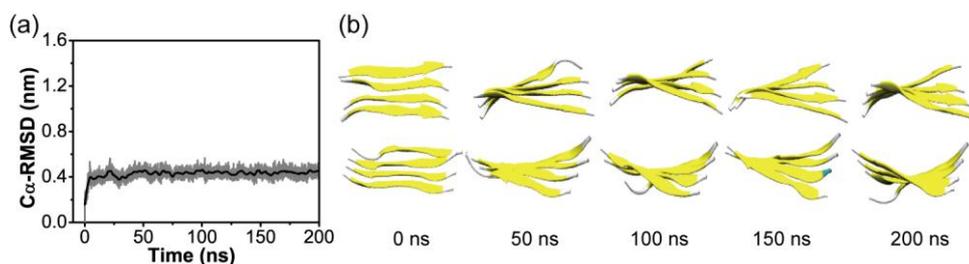


Figure S16. Analysis of a MD simulation starting from the 4+4 bilayer β -sheet with 0.1 M NaCl included. The Ca-RMSD value and the representative structures are quite similar to those in the MD simulations without NaCl (Figs. 9, S7), showing that the physiological salt concentration has little effect on the 4+4 parallel bilayer β -sheet. (a) The time evolution of Ca-RMSD with respect to the initial 4+4 bilayer β -sheet, (b) representative snapshots at five different time points.

REFERENCES

1. Yoda, T., Sugita, Y., and Okamoto, Y. (2004) Comparisons of force fields for proteins by generalized-ensemble simulations, *Chemical Physics Letters* 386, 460-467.
2. Nguyen, P. H., Li, M. S., and Derreumaux, P. (2011) Effects of all-atom force fields on amyloid oligomerization: replica exchange molecular dynamics simulations of the Abeta(16-22) dimer and trimer, *Physical chemistry chemical physics : PCCP* 13, 9778-9788.
3. Lindorff-Larsen, K., Maragakis, P., Piana, S., Eastwood, M. P., Dror, R. O., and Shaw, D. E. (2012) Systematic validation of protein force fields against experimental data, *PloS one* 7, e32131.
4. Freddolino, P. L., Park, S., Roux, B., and Schulten, K. (2009) Force field bias in protein folding simulations, *Biophysical journal* 96, 3772-3780.
5. Lindorff-Larsen, K., Piana, S., Dror, R. O., and Shaw, D. E. (2011) How fast-folding proteins fold, *Science* 334, 517-520.
6. Best, R. B., and Hummer, G. (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides, *The journal of physical chemistry. B* 113, 9004-9015.
7. Nymeyer, H. (2009) Energy landscape of the trpzip2 peptide, *The journal of physical chemistry. B* 113, 8288-8295.