

Supporting material for: Investigating how
peptide length and a pathogenic mutation modify
the structural ensemble of amyloid beta monomer

Yu-Shan Lin
Department of Chemistry,
Stanford University, Stanford, CA

Gregory R. Bowman
Department of Chemistry,
Stanford University, Stanford, CA
Departments of Chemistry and Molecular & Cell Biology
University of California, Berkeley, CA

Kyle A. Beauchamp
Biophysics Program
Stanford University, Stanford, CA

Vijay S. Pande ¹
Department of Chemistry and Biophysics Program,
Stanford University, Stanford, CA

¹Department of Chemistry and Biophysics Program, Stanford University, 318
Campus Drive West, Stanford University, CA 94305, U.S.A., Tel.: (650)723-3660,
Fax: (650)725-0259

Detail of MD simulations

The A β structure was obtained from PDB 1IYT (A β ₄₂ in 80% HFIP and 20% H₂O at low pH) (1). The peptide protonation state was then modified to mimic that of pH 7 (positively charged N-terminus, lysines and arginine, negatively charged C-terminus, aspartates and glutamates, neutral histidines). The sequence was modified to generate the initial structure for A β ₄₀ and A β ₄₂-E22K. These modifications enabled us to investigate the effect of peptide length and mutations on the monomer structural ensemble. We note that pH has significant effects on A β aggregation (8, 9) and the protonation states of A β peptides may be coupled to their conformational dynamics, which can be captured using constant pH molecular dynamics methods (2, 10). However, the pKa shifts of A β are beyond the scope of this study and will be investigated in the future.

The force field used was AMBER ff99sb and tip3p water (3, 11). This force field has been shown to reproduce reasonably well the experimental *J*-coupling values and residual dipolar coupling data for A β ₄₂ (4). For each system, energy minimization was done to remove strain (all bonds with hydrogen atoms constrained) and fifteen 10-ns long NVT simulations at 500K were performed (all bonds with hydrogen atoms constrained); timestep of 0.5 fs used), starting with different velocities, resulting in fifteen different peptide configurations (12, 13). Each peptide was then solvated with tip3p water and counter ions (Na⁺) in a 5nm \times 5nm \times 5nm box. We note that only neutralizing counter ions were used in our simulations and salt ions can have a significant effect on A β (14), which, however, is beyond the scope of this study. We also note that a fully extended A β would demand a box size much greater than 5nm but the use of a large box is extremely computationally intensive for explicit solvent simulations, a weakness implicit solvent simulations are free from.

After energy minimization to remove strain for each simulation box, a 100-ps long NVT simulation at 300K was performed (all bonds with hydrogen atoms constrained; timestep of 2 fs used) followed by a 100-ps long NPT simulation at 300K and 1 atm (all bonds with hydrogen atoms constrained; timestep of 2 fs used). Particle mesh Ewald was used for calculating long-range electrostatic interactions and a switching function with the switching starting at 9 Å and a cutoff at 12 Å was used for van der Waals interactions.

During the first-stage sampling, from each of the fifteen equilibrated configurations, 1000 trajectories (NPT simulations at 300K and 1 atm; all bonds with hydrogen atoms constrained; timestep of 2 fs used) were simulated with different initial velocities via the Folding@home distributed computing plat-

form for each system. After simulating for an aggregate time of $\sim 50 \mu\text{s}$, 1000 configurations were selected using the k-center algorithm as the starting configurations of the second-stage sampling.

From each of the 1000 configurations, 20 trajectories (NPT simulations at 300K and 1 atm; all bonds with hydrogen atoms constrained and a timestep of 2 fs) were simulated with different initial velocities via the Folding@home distributed computing platform for each system. The trajectories collected from this second-stage sampling were then used for final structural analysis for $A\beta$ peptides. Configurations along each trajectory were output every 100 ps.

To avoid using data with possible unphysical interactions between the peptide and its periodic images, once the minimum distance between the peptide and its images was less than 1 nm for consecutively 10 outputs (*i.e.* 1 ns continuously), the following data in that trajectory were discarded. After the clean-up, only trajectories longer than 20 ns were used for analysis. The total number of trajectories used for analysis was $\sim 8,000$ and the average simulation length of each trajectory is ~ 30 ns. The total accumulative simulation time for each system was $>200 \mu\text{s}$.

Check for convergence behavior

Since the fifteen initial configurations were generated by heating up a rather α helical configuration (1) at 500 K for 10 ns, there may have still existed some residual α -helix or structures that are biased toward gas-phase configurations. We thus included the results of secondary structure analysis for the initial configurations for $A\beta_{42}$ in Fig. S1, for comparison to Figure 3 reported in the main text. The secondary structures for each representative conformation were assigned using the DSSP program (5), based on which each residue can adopt one of eight secondary structures: α -helix (H), 3_{10} -helix (G), π -helix (I), isolated bridge (B), extended β (E), bend (S), turn (T), or loop/irregular (\sim). There was some secondary structure in the initial configurations – for example, prominent 3_{10} -helix propensity in the region of residues 12–16 and residues 30–35, significant α -helix propensity in the region of residues 12–20 and significant π -helix propensity in the region of residues 6–9.

In the top left panel of Fig. S2, we calculated and plotted the residue-dependent α -helix (H) propensity using data collected from different time segments during the first-stage sampling of $A\beta_{42}$. We observed that the initial α -helix propensity for $A\beta_{42}$ seemed to be equilibrated quickly after the peptide was placed in water.

In the top right panel of Fig. S2, we plotted how the residue-dependent propensity of α -helix evolved during the second-stage sampling. We saw that by the end of the second-stage sampling the α -helix propensity for A β_{42} seemed to have reached reasonable convergence. Although some changes of the α -helix propensity can still be seen for residues 16–18, the values for most of the residues appear to be converging well, at least on our sub-millisecond sampling time scale. To confirm the converging behavior, in the bottom left panel of Fig. S2, we plotted the time dependence of α -helix propensity for residues 5, 17 and 33, selected from the areas that had the most changes in the top right panel. Each of these time evolution could be fitted by an exponential decay (fitted curves shown as dashed lines in the Figure), especially for residues 5 and 33, demonstrating the converging behavior. With the current aggregate simulation time ($\sim 200 \mu\text{s}$), we are reasonably confident of the convergence behavior of the α -helix propensity, and note that the convergence of the extended β propensity might be more challenging and demand a much longer sampling time (15, 16).

In the bottom right panel of Fig. S2, we calculated and plotted the ${}^3J_{H_N H_\alpha}$ J -couplings with a Karplus equation (17) using data collected from different time segments during the second-stage sampling of A β_{42} . We observed that the calculated J -couplings, which depend only on the ϕ angles, converge faster than the secondary structure. When computed using the equilibrium populations from the second-stage sampling (see the section “Markov state model analysis” in the main text), the calculated J -couplings have a RMSD of 0.84 Hz and a correlation coefficient of 0.47, compared to experiment (6).

In Fig. S3, we show how the calculated structural properties of A β_{42} can depend on the number of independent trajectories in the simulations. Qualitative differences were observed when only 10 independent trajectories were used to compute the residue-dependent α -helix and extended β propensities. We also note that the α -helix propensity converges faster than the extended β propensity.

Difference contact map for A β_{42} -E22K compared to the wild-type A β_{42}

In Fig. S8 we show the difference contact map for A β_{42} -E22K, compared to the wild-type A β_{42} . A β_{42} -E22K has increased contacts between region I (\sim residues 1–15) and region II (\sim residues 16–23), and most particularly between region I and residues 19–23. This can be attributed to the elimination of the repulsion between E22 and the negatively charged side chains in

region I in $A\beta_{42}$ by the E22K mutation. The helix formed in the region of residues 20–24 also decreases contacts between residues 12–20 and residues 25–33. This stretch of helix may function as a spacer between these two regions.

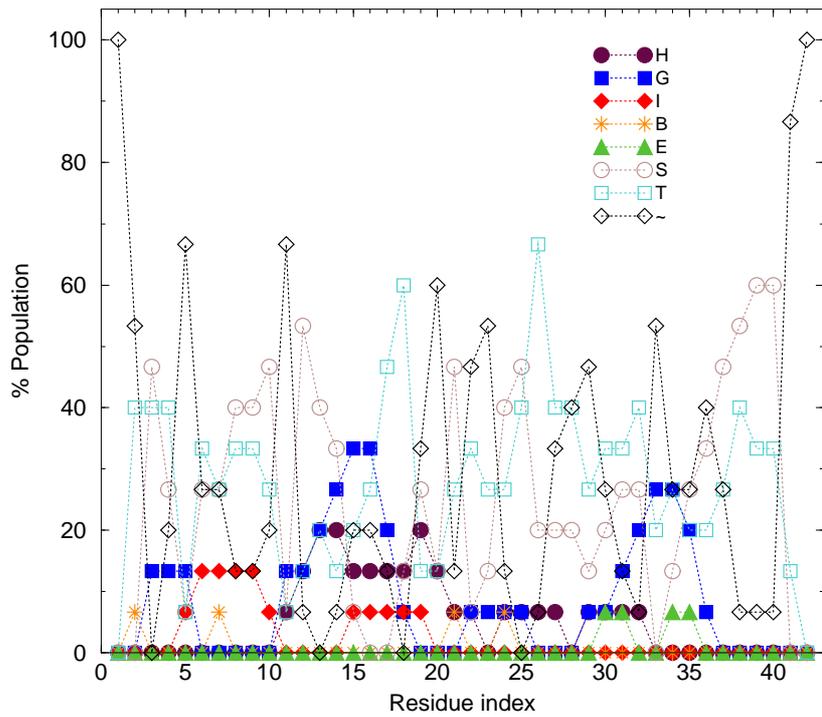


Figure S1: Averaged %population of the secondary structures based on the DSSP analysis for the fifteen initial configurations for $A\beta_{42}$. The eight DSSP assignments are α -helix (H), 3_{10} -helix (G), π -helix (I), isolated bridge (B), extended β (E), bend (S), turn (T), or loop/irregular (\sim).

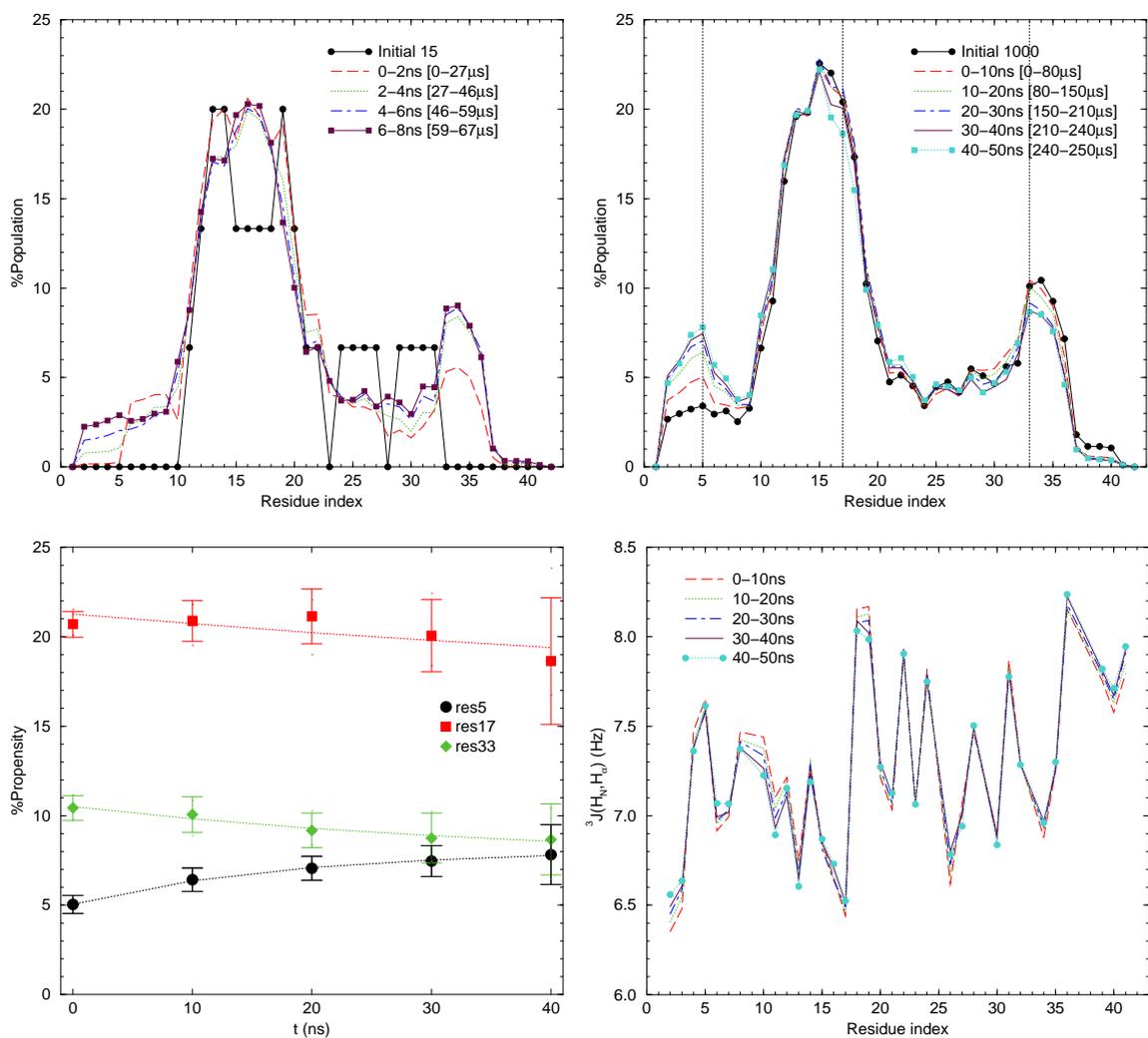


Figure S2: Top left : Time dependence of the residue-dependent α -helix propensity of $A\beta_{42}$ during the first-stage sampling. Given in the brackets in the legends are the aggregate simulation times. Top right : Time dependence of the residue-dependent α -helix propensity of $A\beta_{42}$ during the second-stage sampling. Bottom left : Time dependence of α -helix propensity for residues 5, 17 and 33 of $A\beta_{42}$ during the second-stage sampling. Bottom right : Time dependence of the residue-dependent $^3J_{H_N H_\alpha}$ J -couplings of $A\beta_{42}$ during the second-stage sampling.

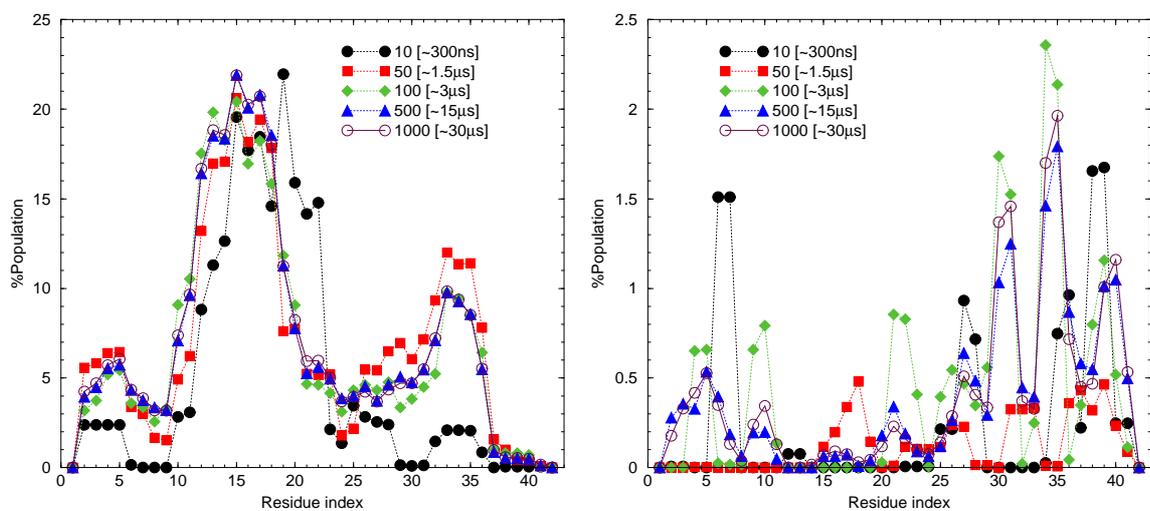


Figure S3: Left : Dependence of the residue-dependent α -helix propensity of $A\beta_{42}$ on the number of independent trajectories during the second-stage sampling. Given in the brackets in the legends are the aggregate simulation times. Right : Dependence of the residue-dependent extended β propensity of $A\beta_{42}$ on the number of independent trajectories during the second-stage sampling.

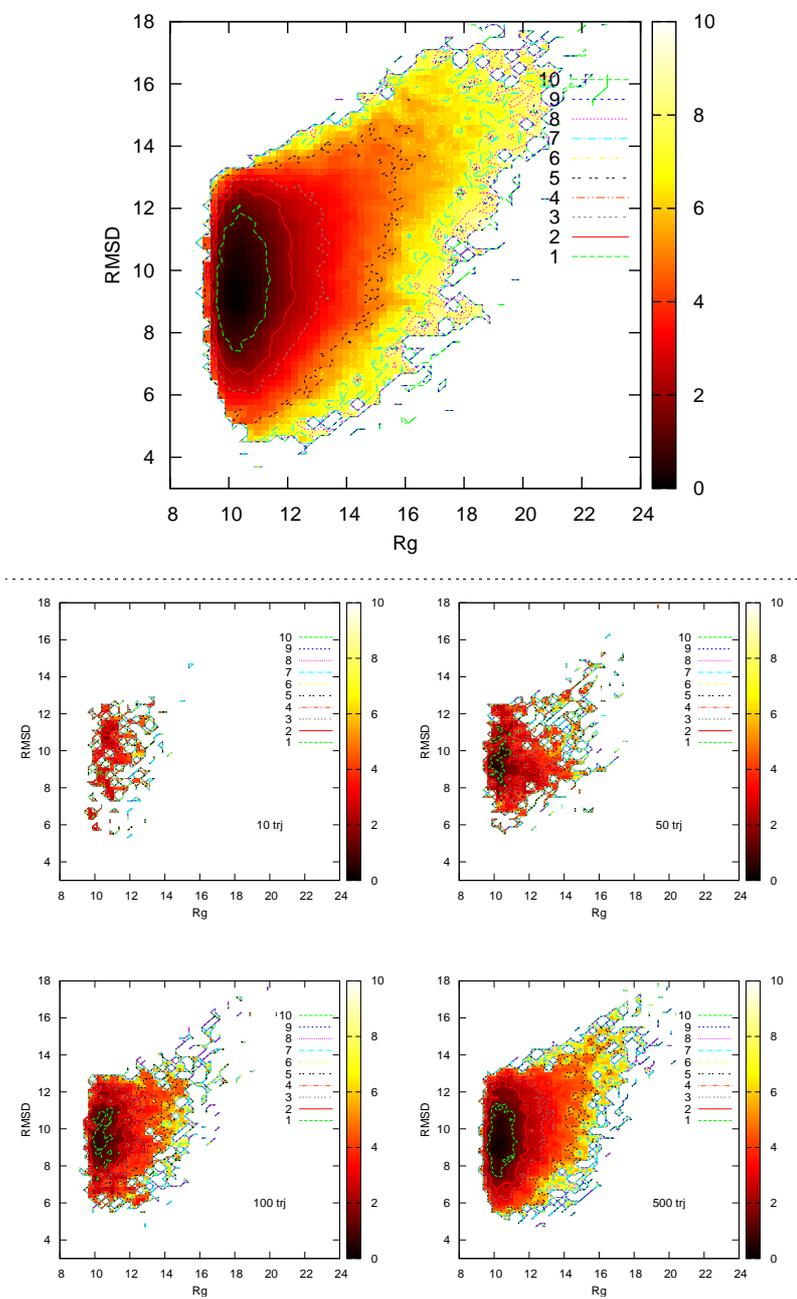


Figure S4: Top : Free energy landscape (in kT) of $A\beta_{42}$ as a function of radius of gyration (in Å) and RMSD (in Å) from the fibril structure (7) (backbone of residues 17-42). Bottom : Dependence of the analysis results on the number of independent trajectories during the second-stage sampling. Upper left, upper right, lower left and lower right are the results from 10, 50, 100 and 500 independent trajectories, respectively.

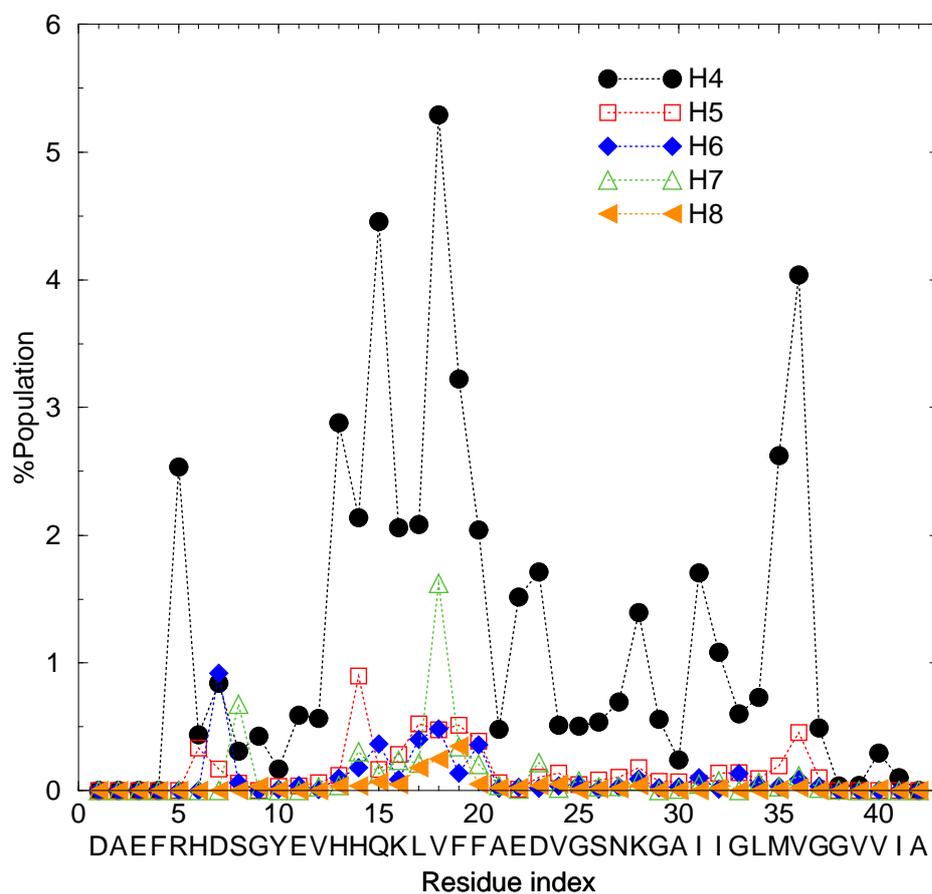


Figure S5: Length analysis of the α -helix segments in $A\beta_{42}$. Each data point in H_n represents the tendency to form an α -helix segment with a specific size n , ending at the residue. For example, $\sim 5.5\%$ of the conformations in the $A\beta_{42}$ ensemble have an α -helix segment of size four, involving residues 15–18. In most conformations only short α -helix segments involving four residues (H4) were formed.

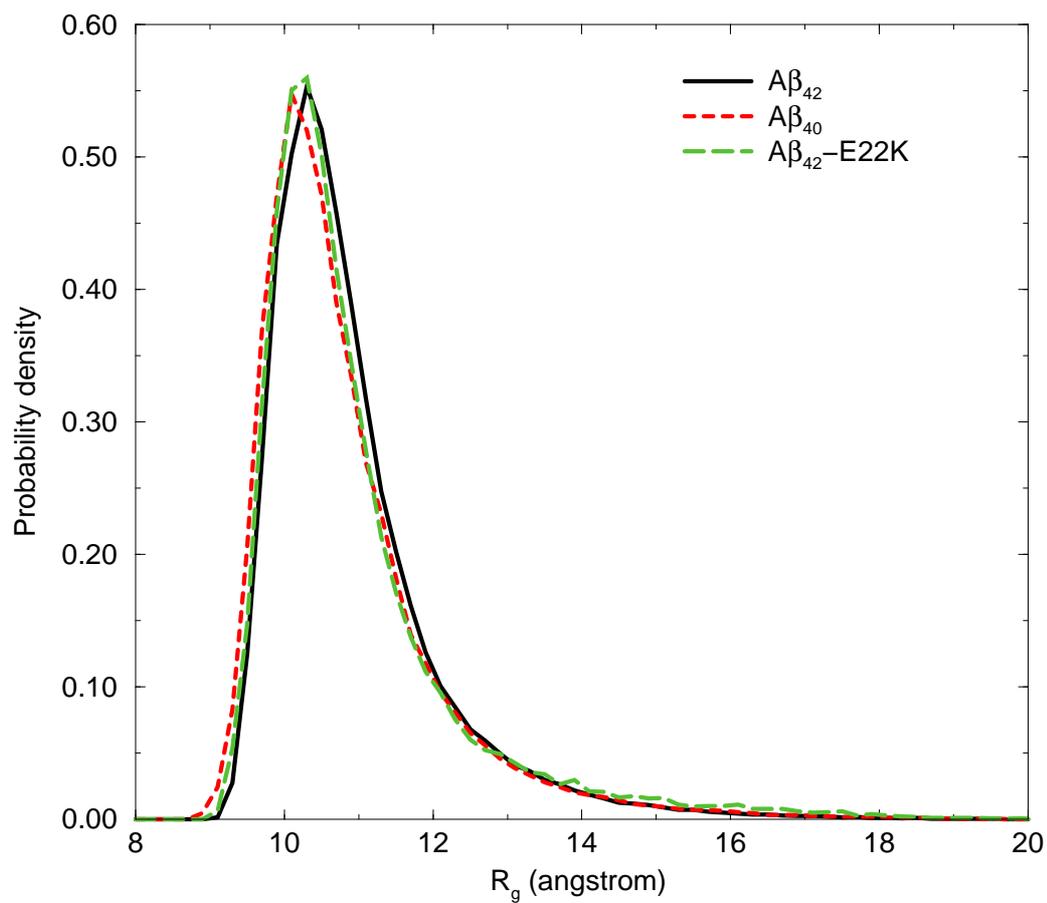


Figure S6: Distribution of the radii of gyration of Aβ₄₂, Aβ₄₀ and Aβ₄₂-E22K.

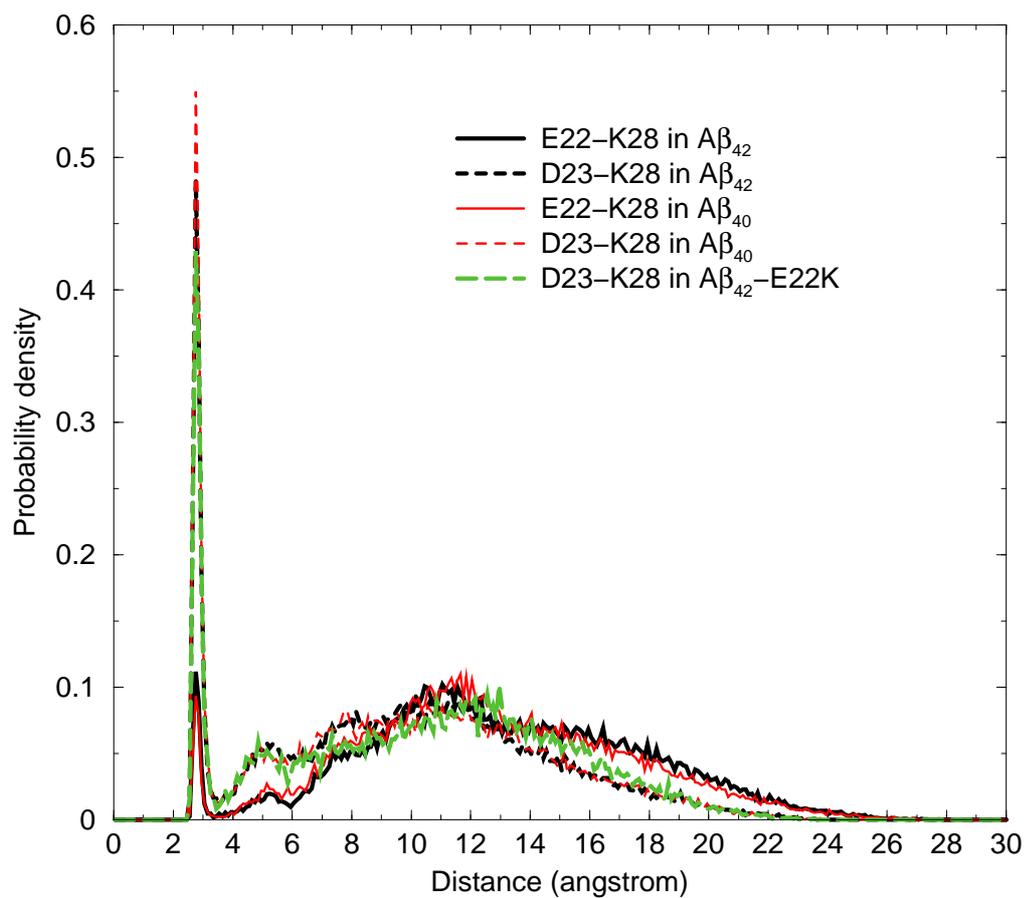


Figure S7: Distance distribution of the E22-K28 and D23-K28 salt bridges. The distance is defined as the minimum separation between the carboxylate oxygen atom of E22/D23 and the amino nitrogen atom of K28.

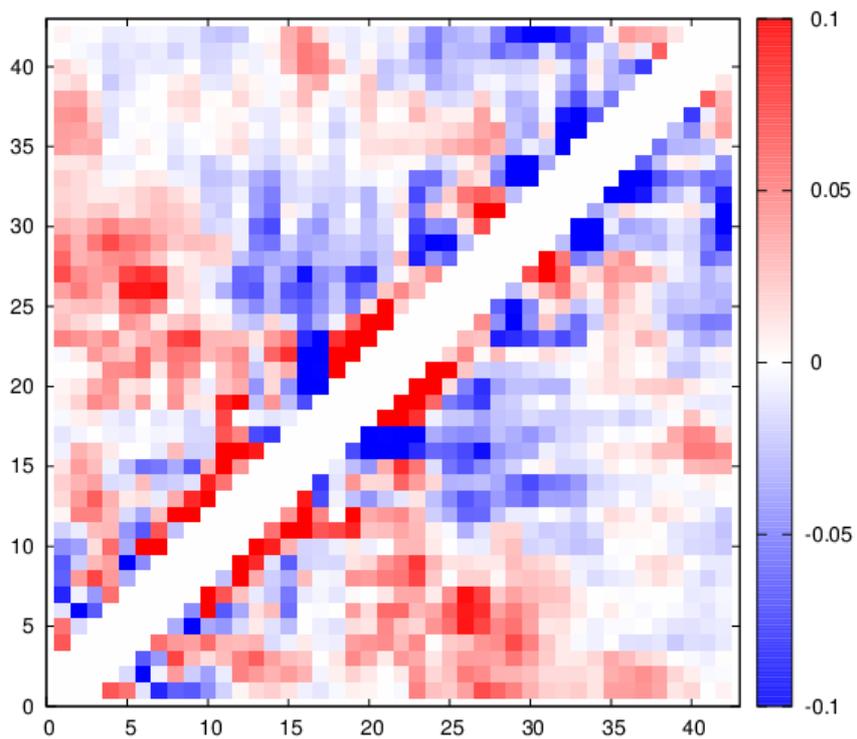


Figure S8: Difference contact map for $A\beta_{42}$ -E22K from $A\beta_{42}$. A positive value indicates more contacts while negative indicates fewer contacts in $A\beta_{42}$ E22K, compared to the wild-type $A\beta_{42}$.

Supporting References

1. Crescenzi, O., S. Tomaselli, R. Guerrini, S. Salvadori, A. M. D'Ursi, P. A. Temussi, and D. Picone, 2002. Solution structure of the Alzheimer amyloid β -peptide (1-42) in an apolar microenvironment. *Eur. J. Biochem.* 269:5642–5648.
2. Khandogin, J., and C. L. Brooks, III., 2007. Linking folding with aggregation in Alzheimer's β -amyloid peptides. *Proc. Natl. Acad. Sci. USA* 104:16880–16885.
3. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65:712–725.
4. Sgourakis, N. G., M. Merced-Serrano, C. Boutsidis, P. Drineas, Z. Du, C. Wang, and A. E. Garcia, 2011. Atomic-level characterization of the ensemble of the A β (1-42) monomer in water using unbiased molecular dynamics simulations and spectral algorithms. *J. Mol. Biol.* 405:570–583.
5. Kabsch, W., and C. Sander, 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.
6. Sgourakis, N. G., Y. Yan, S. A. McCallum, C. Wang, and A. E. Garcia, 2007. The Alzheimer's peptides A β 40 and 42 adopt distinct conformations in water: A combined MD/NMR study. *J. Mol. Biol.* 368:1448–1457.
7. Lührs, T., C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, and R. Riek, 2005. 3D structure of Alzheimer's amyloid- β (1-42) fibrils. *Proc. Natl. Acad. Sci. USA* 102:17342–17347.
8. Wood, S. J., B. Maleeff, T. Hart, and R. Wetzel, 1996. Physical, morphological and functional differences between pH 5.8 and 7.4 aggregates of the Alzheimer's amyloid peptide A β . *J. Mol. Biol.* 256:870–877.
9. McLaurin, J., D.-S. Yang, C. M. Yip, and P. E. Frase, 2000. Review: Modulating factors in amyloid- β fibril formation. *J. Struct. Biol.* 130:259–270.
10. Khandogin, J., and C. L. Brooks, III., 2005. Constant pH molecular dynamics with proton tautomerism. *Biophys. J.* 89:141–157.

11. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
12. Berendsen, H. J. C., D. van der Spoel, and R. van Drunen, 1995. GRO-MACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.
13. Lindahl, E., B. Hess, and D. van der Spoel, 2001. Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7:306–317.
14. Klement, K., K. Wieligmann, J. Meinhardt, P. Hortschansky, W. Richter, and M. Fändrich, 2007. Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer’s A β (1-40) amyloid fibrils. *J. Mol. Biol.* 373:1321–1333.
15. Kubelka, J., J. Hofrichter, and W. Eaton, 2004. The protein folding “speed limit”. *Curr. Opin. Struct. Biol.* 14:76–88.
16. Ghosh, K., S. B. Ozkan, and K. A. Dill, 2007. The ultimate speed limit to protein folding is conformational searching. *J. Am. Chem. Soc.* 129:11920–11927.
17. Vuister, G. W., and A. Bax, 1993. Quantitative J correlation: a new approach for measuring homonuclear three-bond $J(H^N H^\alpha)$ coupling constants in ^{15}N -enriched proteins. *J. Am. Chem. Soc.* 115:7772–7777.