Supplementary Material

1 SUPPLEMENTARY FIGURES

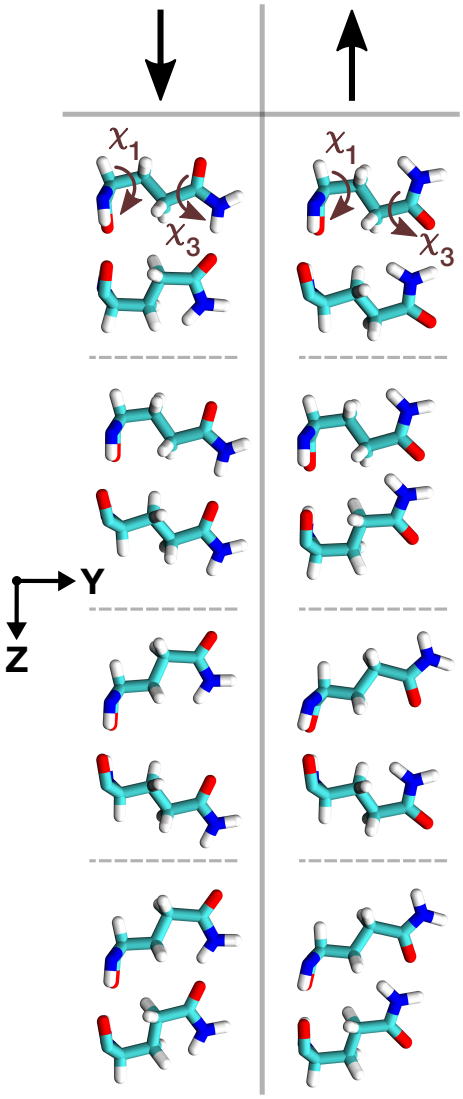

Figure S1. Schematic visualization of the eight (χ_1, χ_3) rotamer classes of a Gln–Gln pair, in which a side-chain–side-chain hydrogen bonding interaction could take place. The side-chains of the pair are located on neighbouring β-strands of an antiparallel β-sheet inside the polyQ amyloid core. Depending on the H-bond direction, the eight distinct structural classes are categorized into two forms: $\downarrow$ and $\uparrow$.

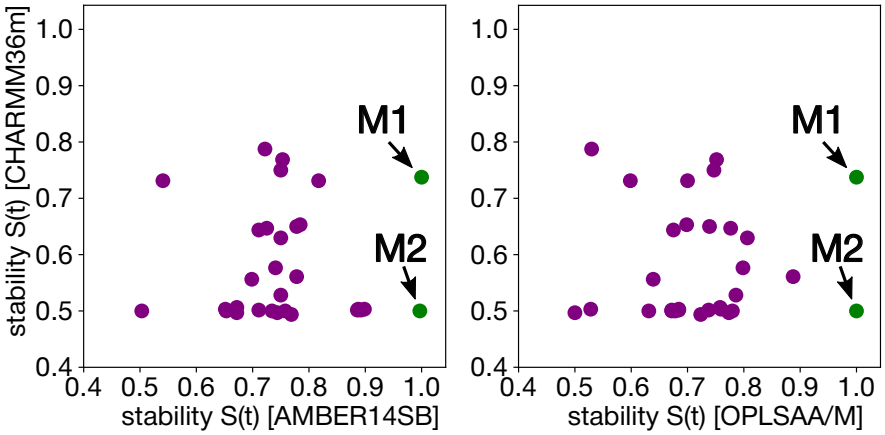

Figure S2. The stabilities $S(t)$ (see Eq. (1) in Methods of the main text) of the 30 candidate structures for the polyQ amyloid core examined with up-to-1-μs molecular dynamics simulations using three distinct force fields: CHARMM36m (Huang et al. (2017)), AMBER14SB (Maier et al. (2015)) and OPLSAA/M (Robertson et al. (2015)). Strikingly, both **M1** and **M2** models demonstrated complete stability when simulated with AMBER14SB and OPLSAA/M, while in CHARMM36m none of the 30 candidates did.

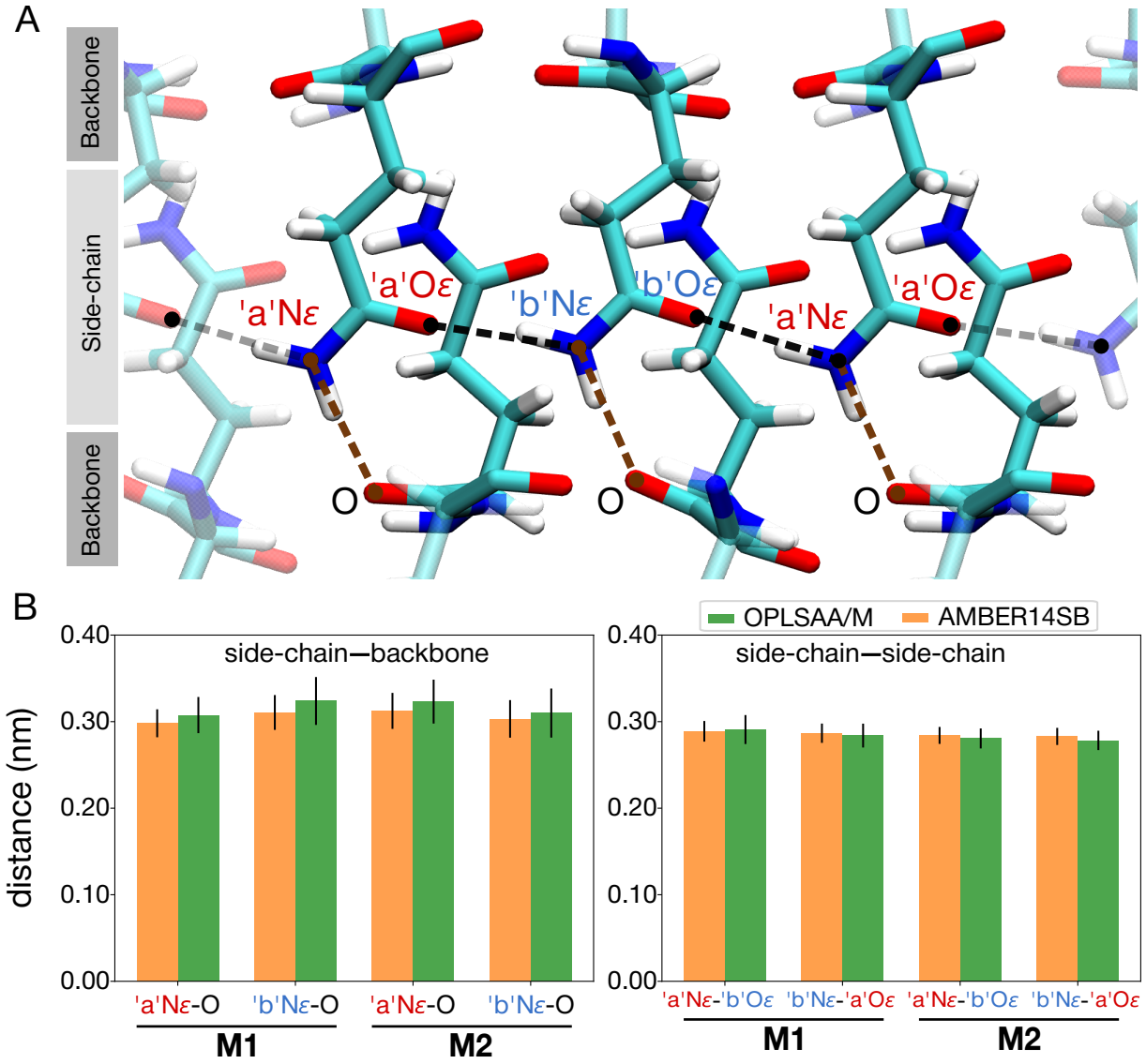

Figure S3. (A) Spatial arrangement of the side-chain and backbone atoms in the polyQ amyloid core model **M2**. Dashed lines represent the distances between the N_ϵ atoms of the "a" and "b" types of Gln residues and their nearest backbone O atoms (brown), as well as the N_ϵ–O_ϵ distances between side-chains (black). (B) Side-chain–backbone distances (left panel), i.e., the distances between the N_ϵ atoms of the "a" and "b" types of Gln residues and their closest backbone O atoms. Side-chain–side-chain distances (right panel), i.e., the N_ϵ–O_ϵ distances between the side chains, either from "a" to "b" or from "b" to "a". The calculations are presented for both the **M1** and **M2** models, using two force fields: OPLSAA/M (green) and AMBER14SB (orange); error bars show standard deviation.

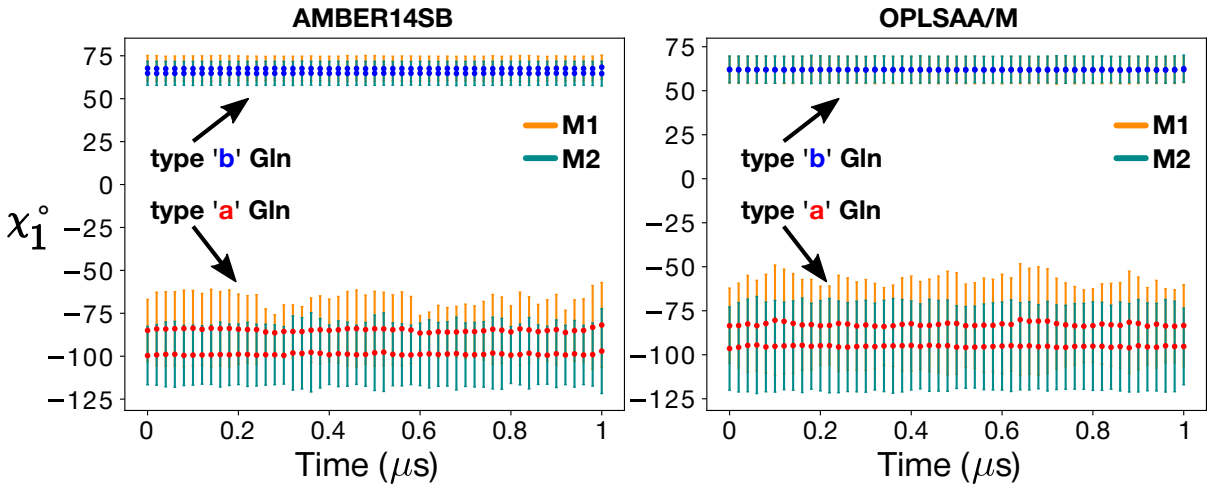

Figure S4. Both polyQ$_{15}$ models **M1** (orange) and **M2** (green) remained stable throughout 1-μs MD simulations in both the AMBER14SB (left) and OPLSAA/M (right) force fields. Shown are the mean χ_1 dihedral angles in type "a" (red) and "b" (blue) Gln side-chains, calculated over consecutive 20-ns windows; error bars show standard deviation.

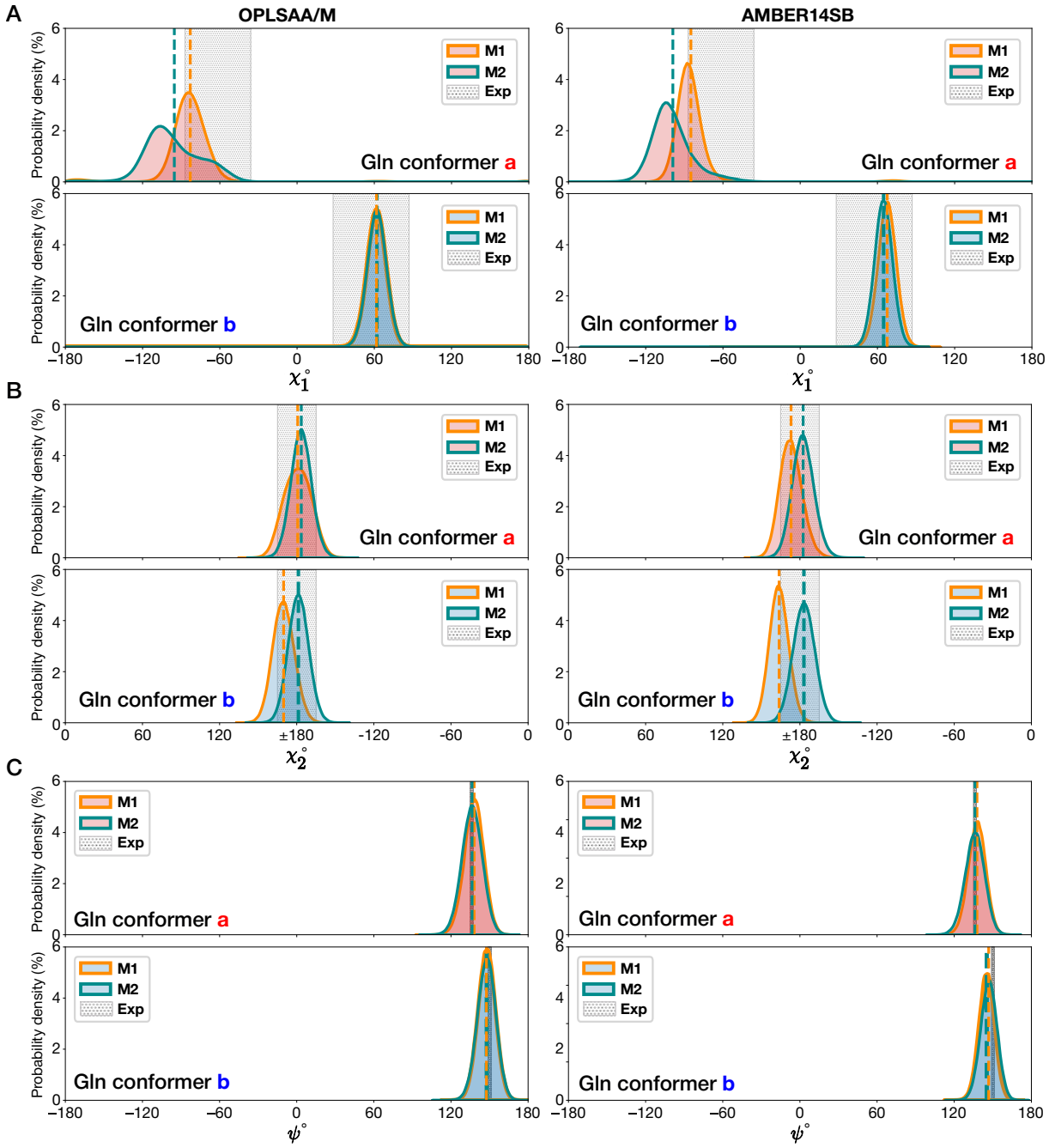

Figure S5. Distributions of the (A) χ_1, (B) χ_2, and (C) ψ dihedral angles in the **M1** (orange lines) and **M2** (green lines) models of the $D_2Q_{15}K_2$ fibril obtained from 1-μs MD simulations using the OPLSAA/M (left) and AMBER14SB (right) force fields. The upper panels show the type "a" conformers (red-shaded distributions) and the lower panels the type "b" (blue-shaded). The gray-shaded regions represent the ssNMR-informed constraints. The dashed vertical lines depict the mean values of the corresponding dihedral angles.

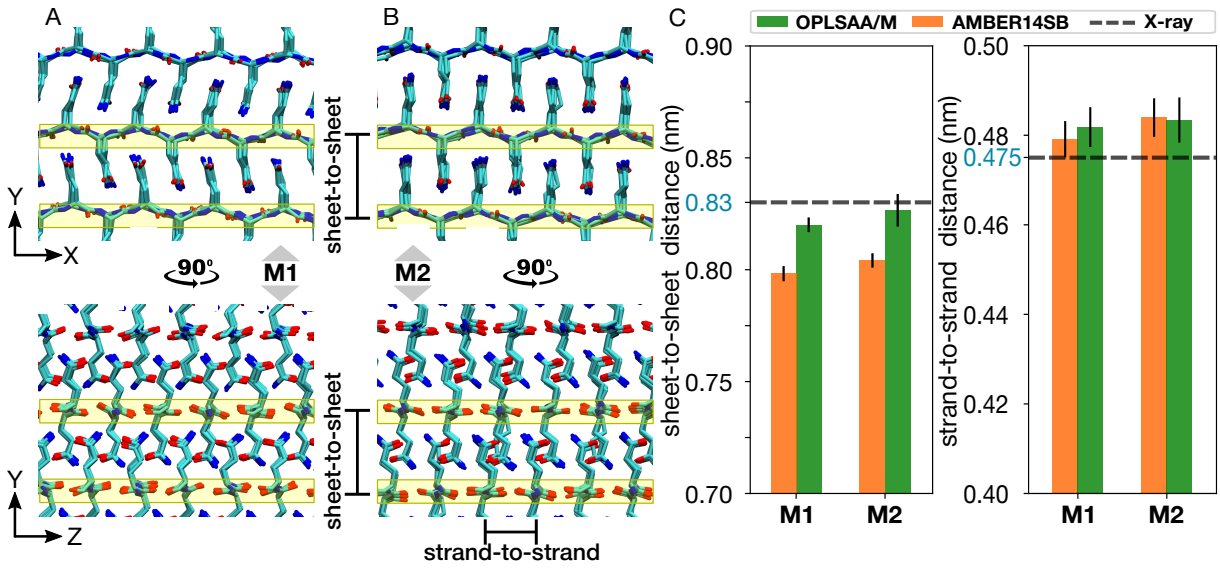

Figure S6. Illustration of the structural differences between the (A) **M1** and (B) **M2** models, visualized through views along the fibril axis z (top panels) and along the β-strand direction x (bottom panels). The yellow boxes highlight backbone atoms. (C) Sheet-to-sheet (left panel) and strand-to-strand (right panel) distances calculated for both the **M1** and **M2** models of polyQ$_{15}$ using the AMBER14SB (orange) and OPLSAA/M (green) force fields. The dashed black horizontal lines corresponds to the data obtained from X-ray experiments Perutz et al. (2002); Sikorski and Atkins (2005). Error bars represent standard deviation.

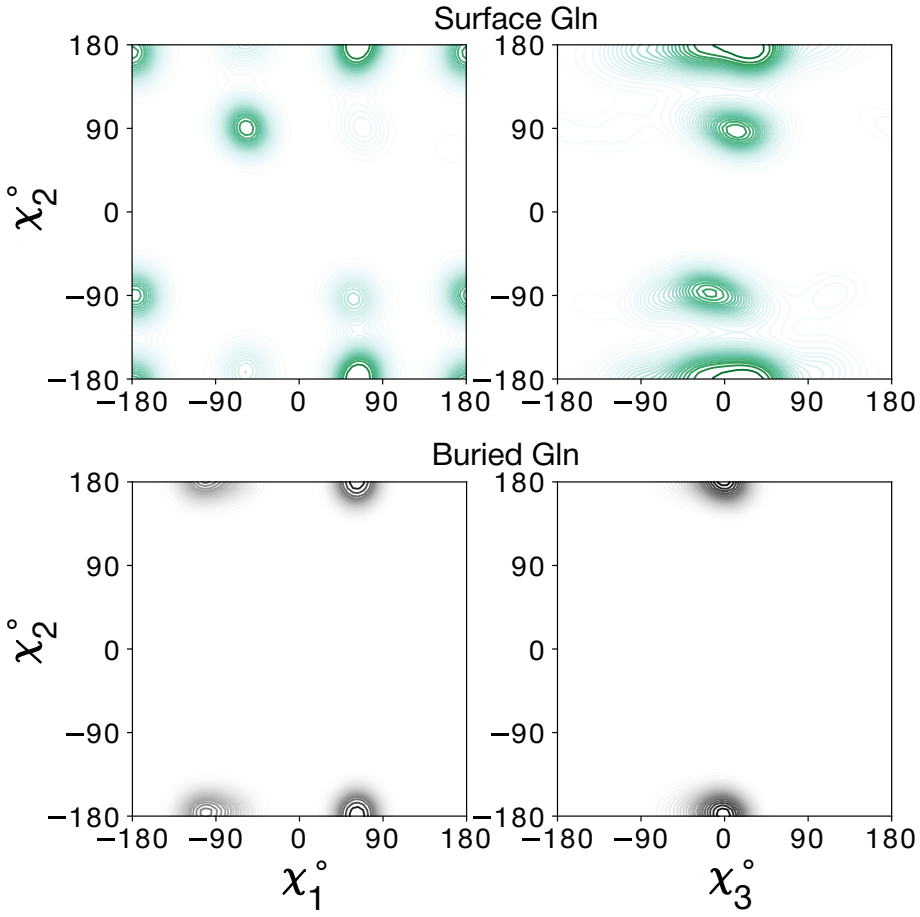

Figure S7. Distributions of the side-chain dihedral angles for Gln residues in the **M2** model of the polyQ$_{15}$ fibril. The water-facing side-chains (top) show more disorder than those internal to the polyQ amyloid core (bottom), but are nonetheless constrained to eight varyingly prominent specific rotamer states.

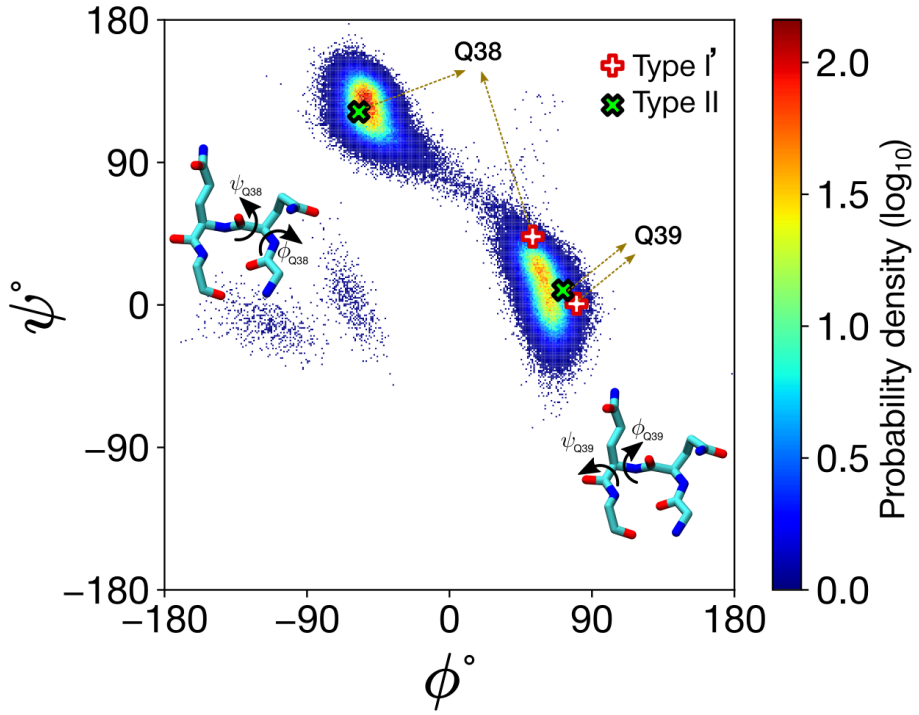

Figure S8. Ramachandran plot for the β-turn residues (Q38 and Q39) in the Q44-HttEx1 polyQ amyloid core. The distribution was obtained over the last 1 μs of the 5-μs MD simulation. The β-turn was initially prepared as a type I' conformer (whose canonical dihedrals are indicated with the red/white plus-signs); however, during the simulation it transitioned to the type II (black/green crosses).

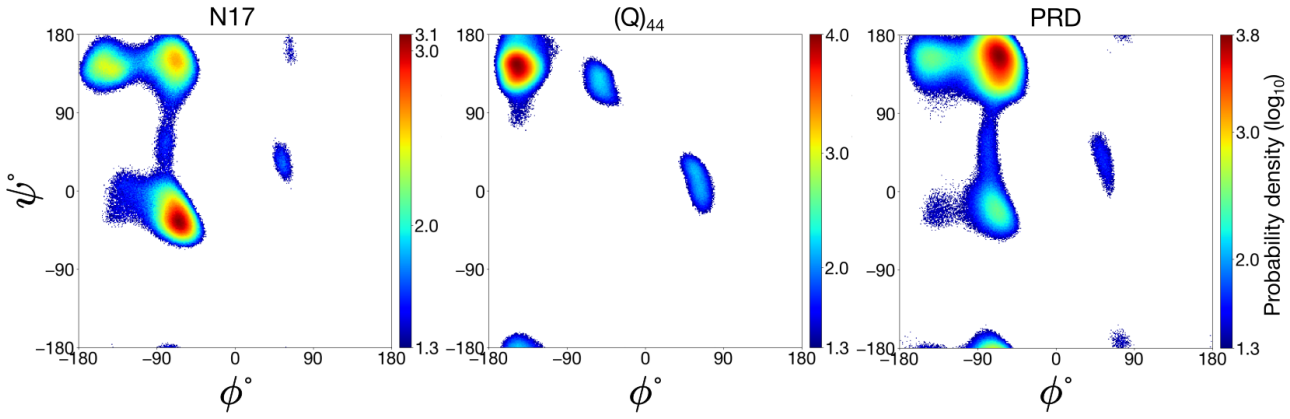

Figure S9. Ramachandran plots of the N17, Q_{44}, and PRD domains elucidate the characteristic secondary structures present in these three disparate domains of the HttEx1 protein: α-helical, β-sheet, and PPII-helical, respectively. The distributions were obtained over the last 1 μs of the 5-μs MD simulation. For clarity, $\log_{10}$(probability density) values below 1.3, characterized primarily by noise, have been excluded.

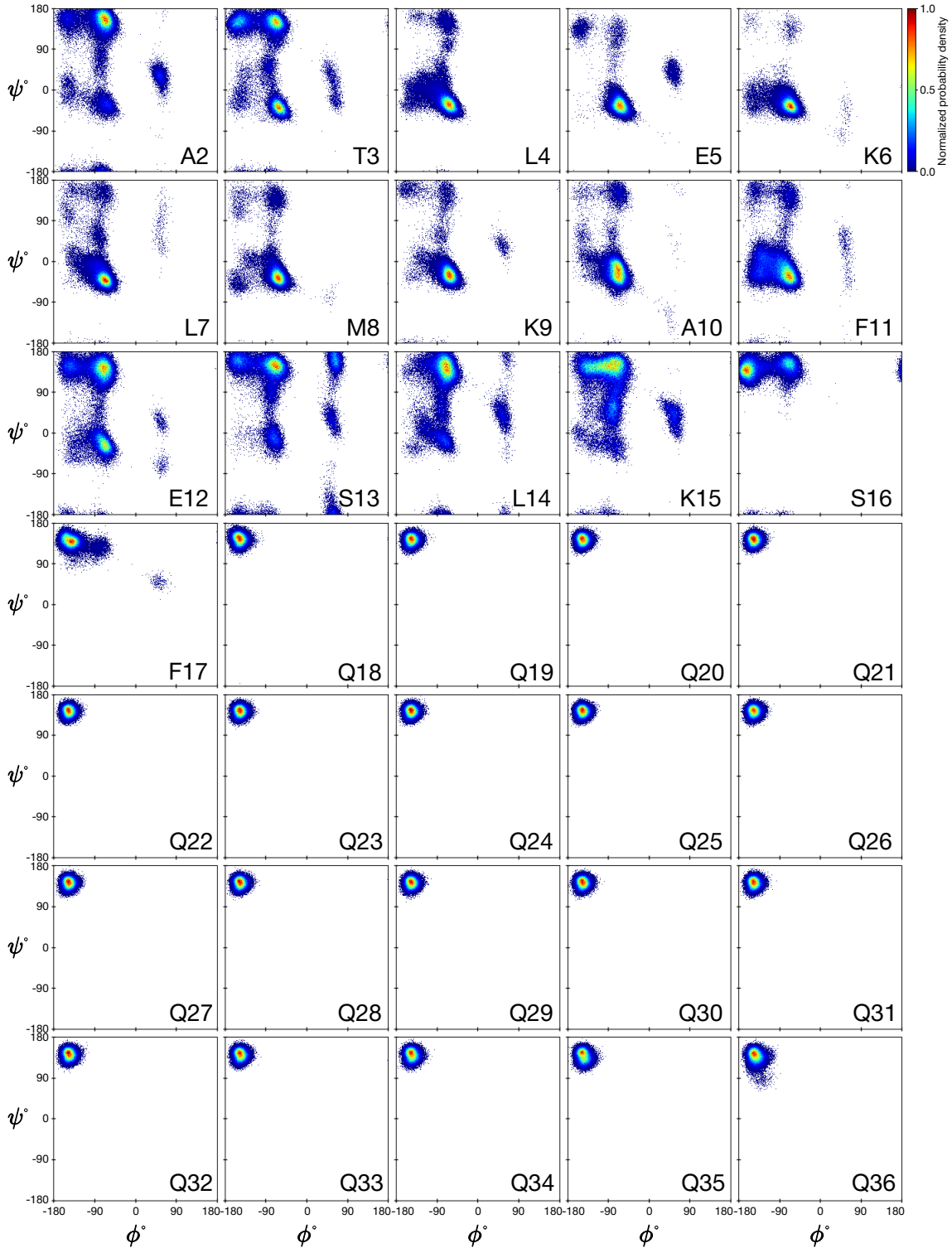

Figure S10. Ramachandran plots for residues 2–36 in the Q44-HttEx1 fibril. The plots showcase the conformational space explored by each individual residue, revealing variations in backbone torsion angles across the protein structure. The colorbar is calibrated such that a normalized value of 1 corresponds to the bin with the highest frequency among all protein residues. Each bin spans one degree in both ϕ and ψ. The distributions were obtained over the last 1 μs of the 5-μs MD simulation.

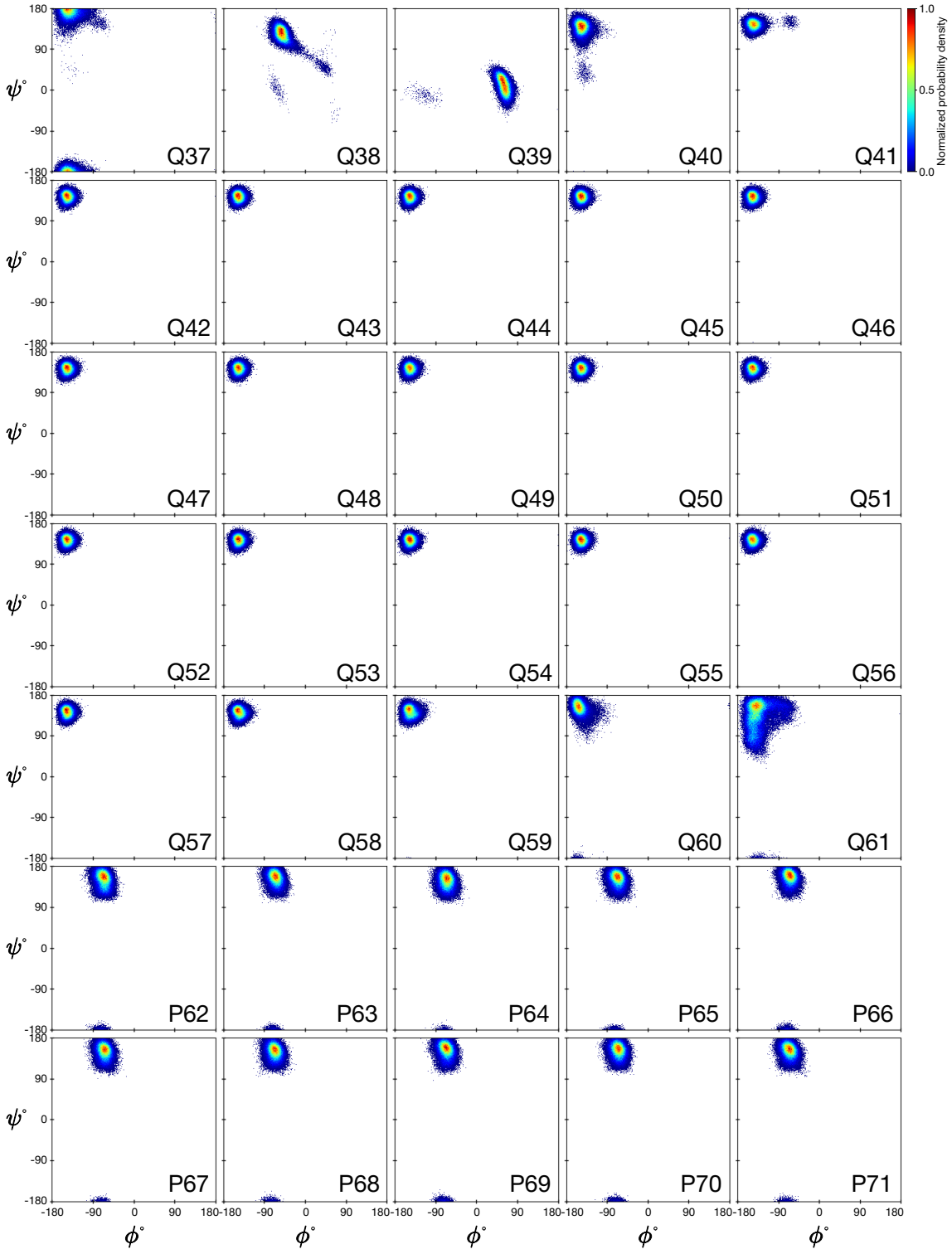

Figure S11. Ramachandran plots for residues 37–71 in the Q44-HttEx1 fibril. The plots showcase the conformational space explored by each individual residue, revealing variations in backbone torsion angles across the protein structure. The colorbar is calibrated such that a normalized value of 1 corresponds to the bin with the highest frequency among all protein residues. Each bin spans one degree in both ϕ and ψ. The distributions were obtained over the last 1 μs of the 5-μs MD simulation.

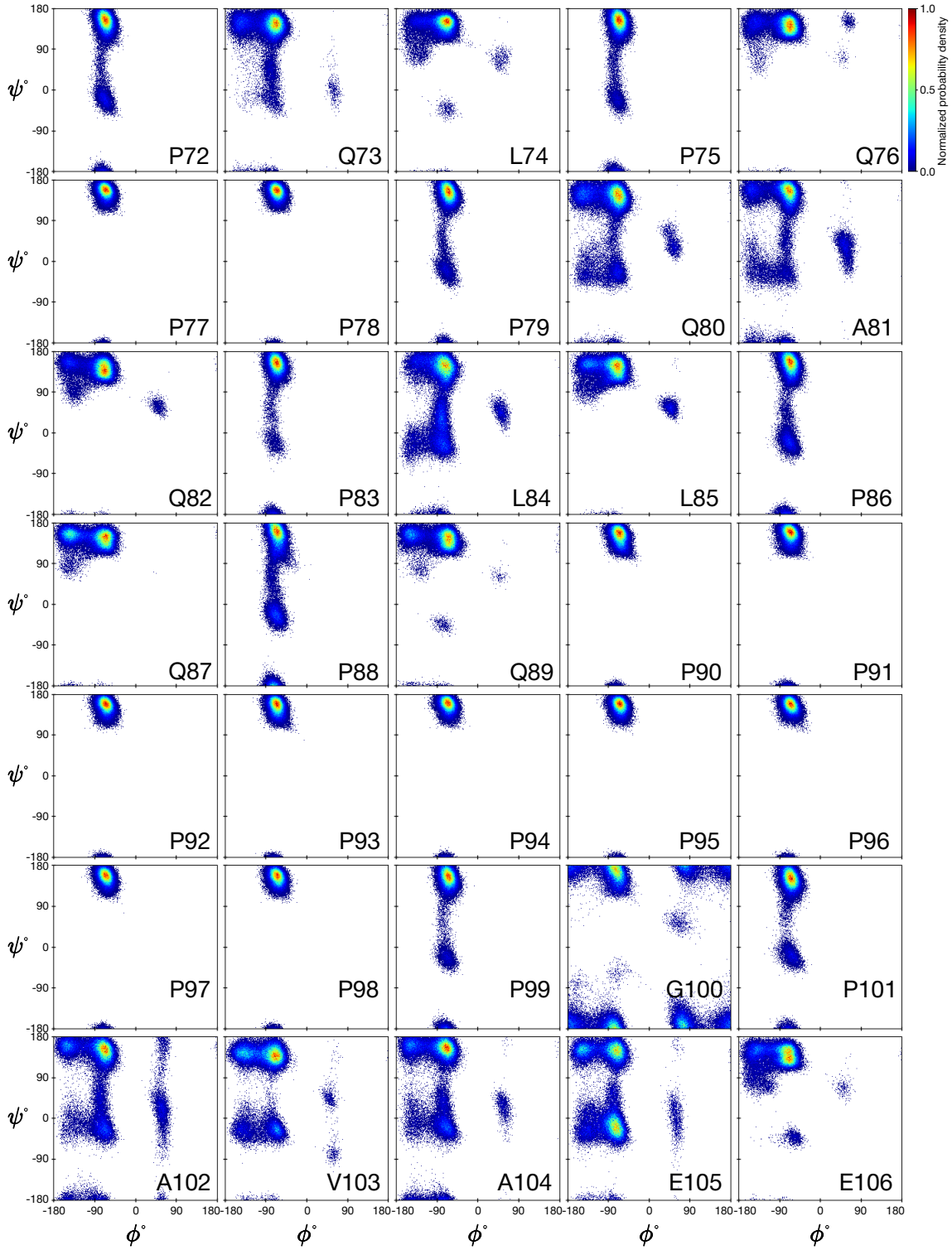

Figure S12. Ramachandran plots for residues 72–106 in the Q44-HttEx1 fibril. The plots showcase the conformational space explored by each individual residue, revealing variations in backbone torsion angles across the protein structure. The colorbar is calibrated such that a normalized value of 1 corresponds to the bin with the highest frequency among all protein residues. Each bin spans one degree in both ϕ and ψ. The distributions were obtained over the last 1 μs of the 5-μs MD simulation.

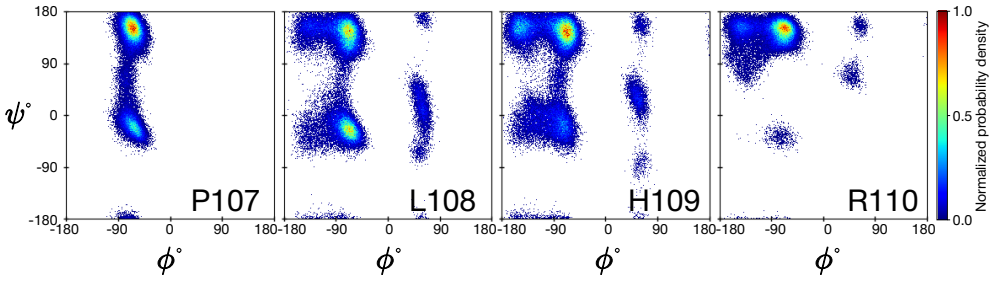

Figure S13. Ramachandran plots for residues 107–110 in the Q44-HttEx1 fibril. The plots showcase the conformational space explored by each individual residue, revealing variations in backbone torsion angles across the protein structure. The colorbar is calibrated such that a normalized value of 1 corresponds to the bin with the highest frequency among all protein residues. Each bin spans one degree in both ϕ and ψ. The distributions were obtained over the last 1 μs of the 5-μs MD simulation.

REFERENCES

Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, et al. Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14** (2017) 71–73.

Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of chemical theory and computation* **11** (2015) 3696–3713.

Robertson MJ, Tirado-Rives J, Jorgensen WL. Improved peptide and protein torsional energetics with the opls-aa force field. *Journal of chemical theory and computation* **11** (2015) 3499–3509.

Perutz MF, Finch JT, Berriman J, Lesk A. Amyloid fibers are water-filled nanotubes. *Proceedings of the National Academy of Sciences* **99** (2002) 5591–5595.

Sikorski P, Atkins E. New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. *Biomacromolecules* **6** (2005) 425–432.