

Supplementary materials and methods

Protein expression and purification

A truncation of the longest human tau isoform httau40 was used throughout this study. It contains all four microtubule binding repeat domains and the C-terminus. It consists of 187 residues (255-441) of httau40 and a 6xHis tag, giving it a total molecular weight of 22 kDa. Concentrations were measured using a molar extinction coefficient of $2.8 \text{ mM}^{-1} \text{ cm}^{-1}$ at 274 nm. This truncation is referred to as tau throughout the text.

The different tau mutants were created by site-directed mutagenesis. DNA constructs were amplified using XL1-Blue *Escherichia Coli* (*E. Coli*) cells, and purified with StrataPrep Plasmid MiniPrep Kit (Agilent). Sequences of each tau construct were verified prior to protein expression via Sanger sequencing (Genewiz).

Expression and purification was largely based on the protocol described by Eschmann et al.¹, and is briefly summarized here. DNA constructs were transfected into BL21 (DE3) *E. Coli* cells, which were then streaked on LB-agar plates containing kanamycin (50 $\mu\text{g}/\text{mL}$, final) and incubated for 16 to 18 hours at 37 °C. Starter cultures (10 mL LB-liquid, 35 $\mu\text{g}/\text{mL}$ kanamycin) were made from isolated colonies and incubated for 16 to 18 hours at 37 °C with shaking at 200 rpm. Liter cultures (1 L LB-liquid, 50 $\mu\text{g}/\text{mL}$ kanamycin) were inoculated with the starter cultures and incubated at 37 °C with shaking at 200 rpm. Expression was induced by adding isopropyl β -D-thiogalactopyranoside (1 mM, final) when an optical density measured at 600nm was between 0.6 and 0.8.

Cells in the expression cultures were pelleted via centrifugation for 20 minutes at 5,000 rpm and 4 °C. These were re-suspended in 30 mL Lysis Buffer (50 mM tris-HCl, 100 mM NaCl, 0.5 mM DTT, 0.1mM EDTA, 0.5 mM DTT (added immediately prior to use)) with a dissolved protease inhibitor tablet (Thermo Scientific Pierce Protease inhibitor tablets, prod #88266) and phenyl-methane sulfonyl fluoride (PMSF) (1 mM, final). Chicken egg-white lysozyme (1.3 mg/mL, final), DNase (10 $\mu\text{g}/\text{mL}$, final), and MgCl_2 (10 mM, final) were added for enzymatic lysis. Physical lysis was performed with 3 rounds of liquid nitrogen freeze fracture. Cellular debris was pelleted via centrifugation for 10 minutes at 13,000 g and 4 °C. The supernatant containing tau was isolated, more PMSF was added (>1 mM, final), and was incubated in a heat-bath at ~65 °C for 12-15 minutes to precipitate unwanted proteins. PMSF was added again (>1 mM, final), and the samples were chilled on ice for 20 minutes. Precipitated unwanted proteins and other debris were pelleted via centrifugation for 10 minutes at 13,000 g and 4 °C.

The supernatant was mixed with Ni-NTA agarose resin equilibrated with buffer A (20 mM sodium phosphate, pH 7.0, 500 mM NaCl, 10 mM imidazole, 0.1 mM EDTA, pH 7.4), and allowed to bind with gentle shaking at 4 °C overnight. Resin was washed with 10 resin volume (RV) of buffer A 3 times, then loaded into an elution column. The resin was washed successively with 10 RV of (1) Buffer A, (2) buffer B (20 mM sodium phosphate pH 7.0, 1 M NaCl, 20 mM imidazole, 0.1 mM EDTA, and 0.5 mM DTT (added immediately prior to use)), (3) buffer A, and (4) buffer B' (20 mM sodium phosphate pH 7.0, 1 M NaCl, 50 mM imidazole, 0.1 mM EDTA, and 0.5 mM DTT (added immediately prior to use)), before elution with 20 mL buffer C (20 mM sodium phosphate pH 7.0, 100 mM NaCl, 500 mM imidazole, 0.1 mM EDTA, and 0.5 mM DTT (added immediately prior to use)). The purified tau in buffer C was exchanged into final buffer (20 mM ammonium acetate, 100 mM NaCl, pH 7.0) using a PD10 desalting column (GE Healthcare). If desired, tau purified peptides were concentrated further using Amicon Centrifugal Concentrators with a 10 kDa cutoff.

Spin labelling

Proteins were spin-labelled using MTSL ((1-Acetoxy-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) Methanethiosulfonate) from Toronto Research Chemicals. MTSL attaches to free thiol groups on cysteine residues. Prior to labelling, samples were treated with 2-5 mM DTT to ensure disulfide bonds were reduced. DTT was removed using a PD10 desalting column, and 10 to 15 x molar excess MTSL to free

cysteine was added for labelling. Labelling was performed with gentle shaking at room temperature, overnight. Excess MTSL was removed using a PD10 desalting column.

Labelling efficiency was measured using the protein concentration calculated with a molar extinction coefficient of $2.8 \text{ mM}^{-1} \text{ cm}^{-1}$ at 274 nm, and the spin-label concentration via double integral of the EPR spectrum. All efficiencies were between 50 and 60%.

Formation and characterization of amyloid aggregates

Purified and labelled tau protein in D₂O or H₂O-based final buffer (20 mM ammonium acetate, 100 mM NaCl, pH 7.0) was mixed with heparin (15 kDa average molecular mass, Galen Laboratory Supplies) at a molar ratio protein:heparin 4:1. The solution was then left at room temperature for 24 h before measurements. Transmission electron microscopy was used to visualize aggregates, using a JEOL-1230 model microscope coupled to an ORCA camera. Sample was typically diluted to 50 μM before deposition on a TEM grid (300 mesh, Formvar/copper, Electron Microscopy Sciences) for 1 min. The grid was then rinsed with water before adding 2% uranyl acetate for 1 minute. The aggregation kinetic was followed using Thioflavin T (ThT) fluorescence. ThT was added to the protein solution at a final concentration of 20 μM and fluorescence (excitation 440 nm, emission 485 nm) was measured over time in a micro plate reader Biotek Synergy 2. The ratio tau mutant:cysteine-less was 1:10 (to match conditions used for DEER samples) and the absolute protein concentrations were 7 μM mutants and 70 μM cysteine-less. The standard deviation, shown only on the last data point for clarity, was calculated from 3 or 4 replicates.

DEER experiments

Proteins were all transferred into D₂O – based buffer (100mM NaCl, 20mM ammonium acetate) using Amicon Centrifugal Concentrators with a 10 kDa cutoff to reach a D₂O concentration of at least 99%. For each distance measurement between residues X and Y, we created a tau protein mutant that had cysteine residues solely at position X and Y. The labelled protein was mixed with cysteine-less protein at a ratio of 1:10 (see Fichou *et al.*² for more details). For the measurements carried out after aggregation, heparin (4:1 tau:heparin molar ratio) was added at this step and the sample was left at room temperature for 24h. The d₈-glycerol was added to reach a final solvent composition of 70:30 D₂O / d₈-glycerol mass ratio. Final concentration of labelled and cysteine-less proteins were 60 μM and 600 μM , respectively, and sample volume was around 50 μl . The sample was transferred into a 3mm quartz capillary and flash-frozen in liquid nitrogen. The frozen sample was then transferred to a precooled resonator for the EPR measurements.

The four-pulse DEER experiments were carried out at 85 K using the Q-band Bruker E580 Eleksys pulse EPR spectrometer operating at ~ 34 GHz equipped with a TWT amplifier (300 W). The measurements were obtained using the following DEER pulse sequence: $\pi/2_{\text{obs}} - \tau_1 - \pi_{\text{obs}} - t - \pi_{\text{pump}} - (\tau_2 - t) - \pi_{\text{obs}} - \tau_2 - \text{echo}$. Rectangular pulses were used in the observe channel with pulse lengths set to $\pi/2_{\text{obs}} = 10\text{-}12$ ns and $\pi_{\text{obs}} = 20\text{-}24$ ns. A π_{pump} pulse was set to a chirp pulse with a length of 20-24 ns and a frequency width of 133MHz. The observe frequency was set 150MHz higher than the center of the pump frequency range. τ_1 was 180 ns and τ_2 was set between 2.4 μs and 4 μs . The DEER signal was accumulated for 4h to 12h and 12h to 24h for the non-aggregated and aggregated samples, respectively. The background-subtracted data were then fitted assuming a Gaussian distribution of the inter-spin distances. The analysis procedure was carried out using the LongDistance software (<http://www.chemistry.ucla.edu/directory/hubbell-wayne-1>).

ESEEM experiments

The sample preparation was the same as described for the DEER experiments.

Two-pulse echo decay experiments were carried with a $\pi/2 - \tau - \pi - \tau - \text{echo}$. The pulse length of rectangular $\pi/2$ and π pulses were 10-12 ns and 20-24 ns, respectively, and the Frequency was set to 33.85 GHz, τ was incremented from 0 to 8 μs with 8 ns increments. The resulting two-pulse ESEEM time

domain data were phase corrected, background subtracted using a biexponential decay, apodized by a Hamming function, zero filled to 512 points and Fourier transformed using CTA FFT to obtain the ESEEM spectra. The maximum of the ^2H ESEEM peak at 7.9 MHz are plotted in figure 4. Error bars on figure 4 represent standard deviation of the ESEEM noise, calculated above 40MHz.

Supplementary figures and discussion

Characterization of tau aggregation ~~aggregation~~

Filament formation induced by heparin was monitored by thioflavin T fluorescence and transmission electron microscopy (Figure S1). The fibrillation kinetic of the four mutants showed in figure 2 are plotted in figure S1A. All labelled mutants, as well as the label-free protein, show a very similar kinetic, suggesting little influence of the mutations/labels on the aggregation pathway. Note that the fluorescence plateau was reached before 24h, time at which samples were measured by DEER and referred to as “after aggregation”. TE micrographs of cysteine-less tau are showed in figure S1B. Some filaments clearly exhibit a twist (intermittent dark and bright filament edges) while other don't. Within the resolution of our TEM pictures, we could not observe significant shape differences between the different mutants (data not shown).

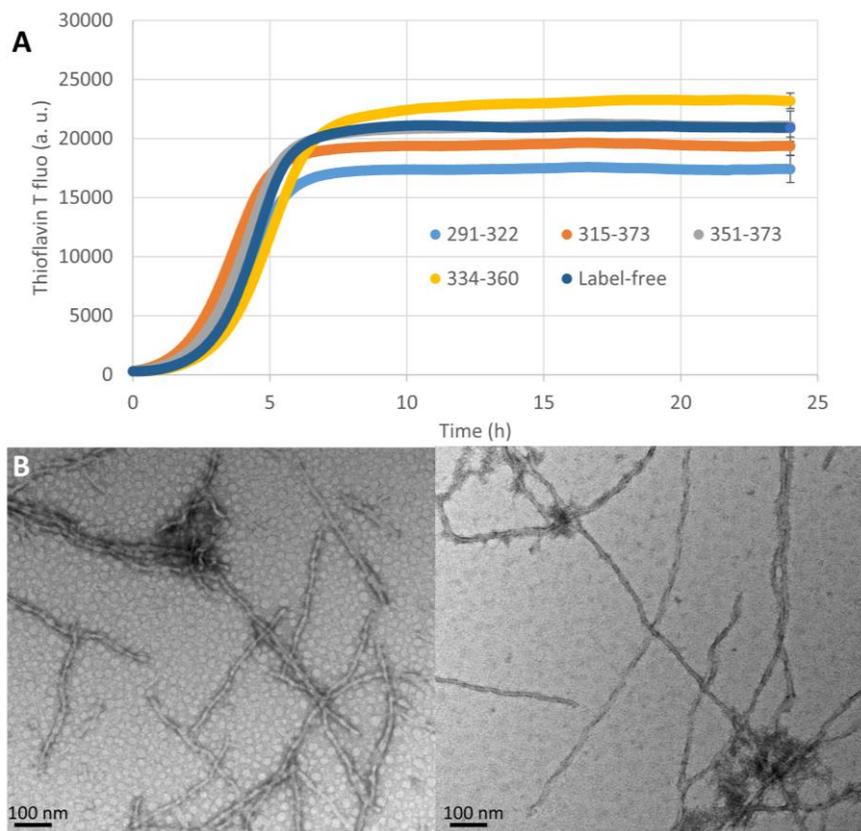


Figure S1: (A) Thioflavin T fluorescence of tau samples measured in figure 2. Heparin was added at time $t=0$ h. Each sample was prepared similarly than for DEER experiments, i.e. with a molar ratio of one labelled protein for 10 non-labeled proteins (cysteine-less). Curve referred to as label-free (blue) represent 100% non-labeled tau. For clarity, error bars representing standard deviation are plotted only on the last data point. (B) TEM pictures of non-labeled tau, each picture representing a different batch of protein.

Choice of spin label position for DEER spectroscopy

We used rational design to produce tau constructs with 4 mutations; C291S, C322S, and 2 additional cysteine mutations to create a DEER capable pair. DEER excels at measuring distances between 1.5 and 5nm, so by computing residue pair distances from a tau (htau40 isoform) solution ensemble³ (Figure S2), we could eliminate pairs that would be too close, or far apart for accurate measurement. Additionally, to give the best resolution possible, we estimated the distance changes expected from solution state to an AD PHF protofilament (PDB ID 5O3L; figure S3). After considering constraints on amino acids deemed to be detrimental to mutate, such as residues internal to the β -sheets, we designed mutations that would provide the largest change in distance.

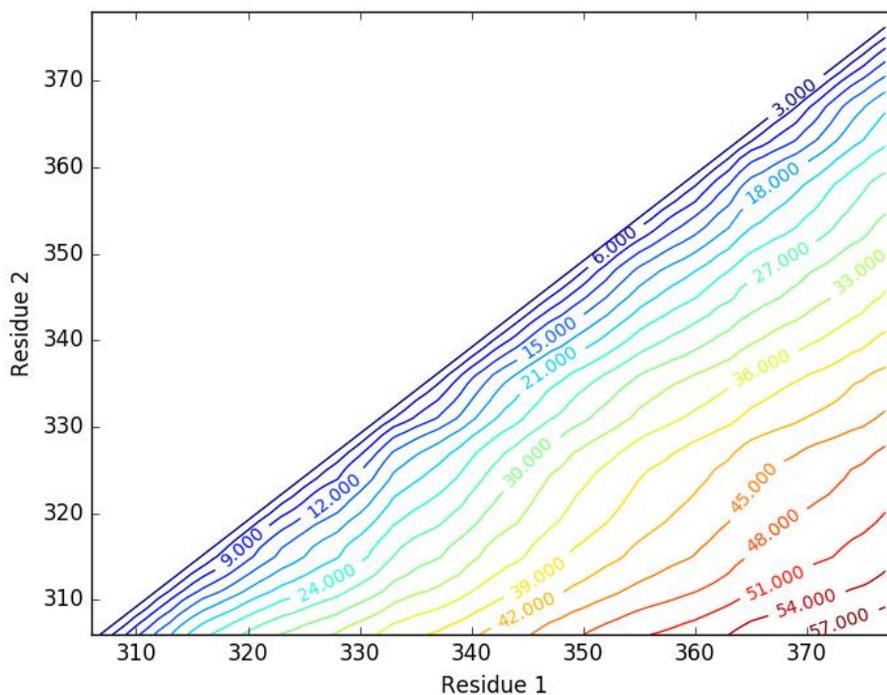


Figure S2: Average residue pairwise distances ($C\alpha$ - $C\alpha$) in the range 303-378 computed from a solution state ensemble of the tau isoform htau40³. Only pairs consisting of residue 1 > residue 2 are shown for simplicity. Contour map represents the pairwise distance in Angstroms with a step of 3Å between each line. Cold to warm colors depict small to large distances, respectively.

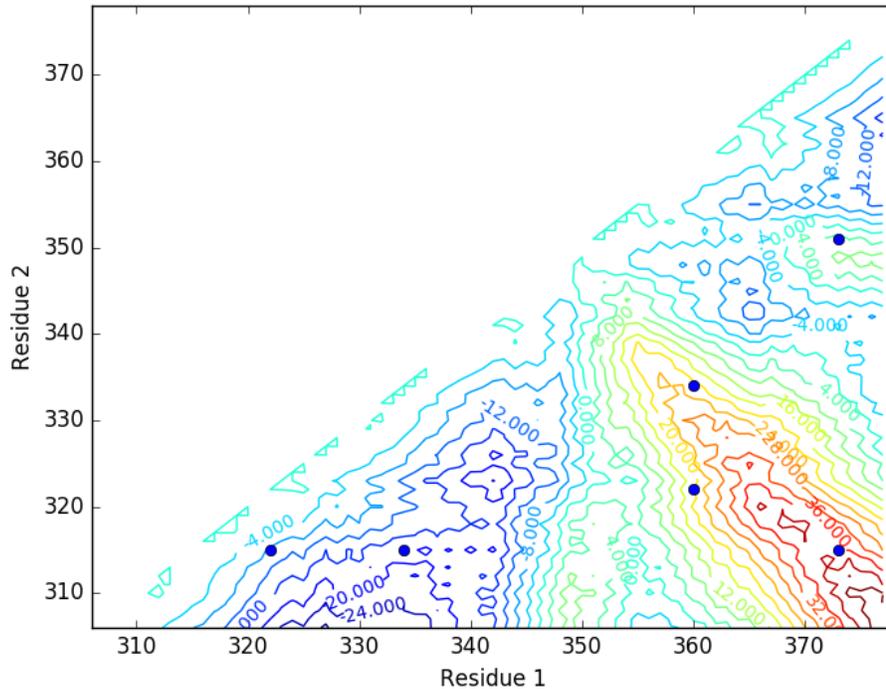


Figure S3: Pairwise distance change from solution state to AD filaments. The pairwise changes were calculated by subtracting the PHF distances ($C\alpha-C\alpha$) from the average solution distances shown in figure S2. Only pairs consisting of residue 1 > residue 2 are shown for simplicity. Contour map shows the distance shift in Angstrom with a step of 4 Å between each line. Positive numbers reflect a decrease of the distance from solution state to PHF. Blue dots show pairwise distances we have measured with DEER in this study.

Estimating the distance between C291 and C322

Fitzpatrick *et al.*⁴ reported the presence of additional electron density along the R3 region. They suggested that, in the case of 4R tau, it originates from R2 forming an anti-parallel β -sheet between residues 290-305 and the R3 region of the paired helical filaments. Based on the filament structure (PDB ID 5O3L) and the EM map (EMDB ID 3741), C322 is measured to be 3.4 nm away from C291.

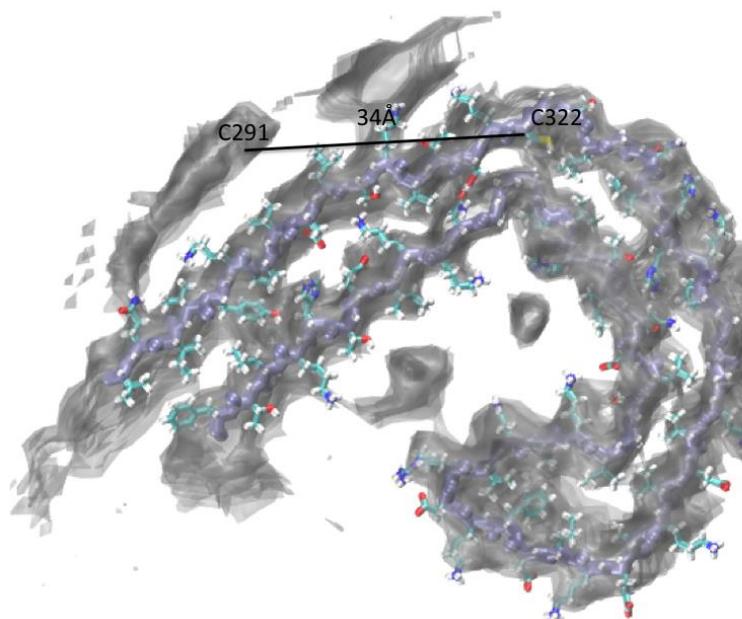


Figure S4: Estimation of the distance between C291 and C322 in 4R PHF tau is found to be 34 Å.

Comparison between DEER measurement and simulations using RotamerConvolveMD

The use of MTSL as a spin-label for DEER measurements introduces multiple additional rotational degrees of freedom inherent to the MTSL structure. The flexibility of MTSL artificially broadens the distance distribution measured by DEER. We used a python package, RotamerConvolveMD^{5,6}, to simulate the DEER signal of tau with MTSL spin labels in both a PHF conformation⁴, and in solution based on an ensemble of conformations generated for httau40³. RotamerConvolveMD appends a library of MTSL onto label positions, and calculates a distribution function based on all possible distance pairs that do not contain any steric hindrance between the spin label and the protein. This simulation allows us to compare structural data from cryo-EM and other measurement techniques to DEER data directly. For each spin label pair we measured with DEER, we also simulated the expected distance distribution for the solution state. We found good agreement between DEER measurements and the calculation from the solution ensemble (figure S5 and S6). Only distances between 315 and 373 show a significant discrepancy, which can be explained by long expected distances that cannot be measured by DEER. In addition to the measured and simulated solution state in figure S5 and S6, we also overlaid on figure S6 the expected and measured distances after aggregation (same data shown in figure 2). The good agreement between simulations and measurements in the solution state but not in the aggregated state reinforces the confidence in the DEER measured distances and thus in the conclusion that heparin-induced filaments are significantly different from the AD filament.

Broadening due to MTSL rotations

Additionally, the simulation of the MTSL rotations tethered to the PHF and SF enable the evaluation of the distance broadening due to MTSL flexibility. Indeed, with no rotational contributions from MTSL, any pairwise distance distribution on a PHF or SF would be a Dirac distance. However, the calculated distance distributions plotted in figure 2 show distribution broadening of 2.1 nm for a probability value greater than 1% of the most probable value. In contrast, DEER measurements showed

distinct broadening beyond what can be attributed to the spin label motion. Indeed, the Gaussian distance distributions showed in figure 2 possess a full width half maximum between 3.0 and 4.9 nm. This broadening must then originate from multiple tau conformations, reflecting multiple tau aggregate structures present in the sample.

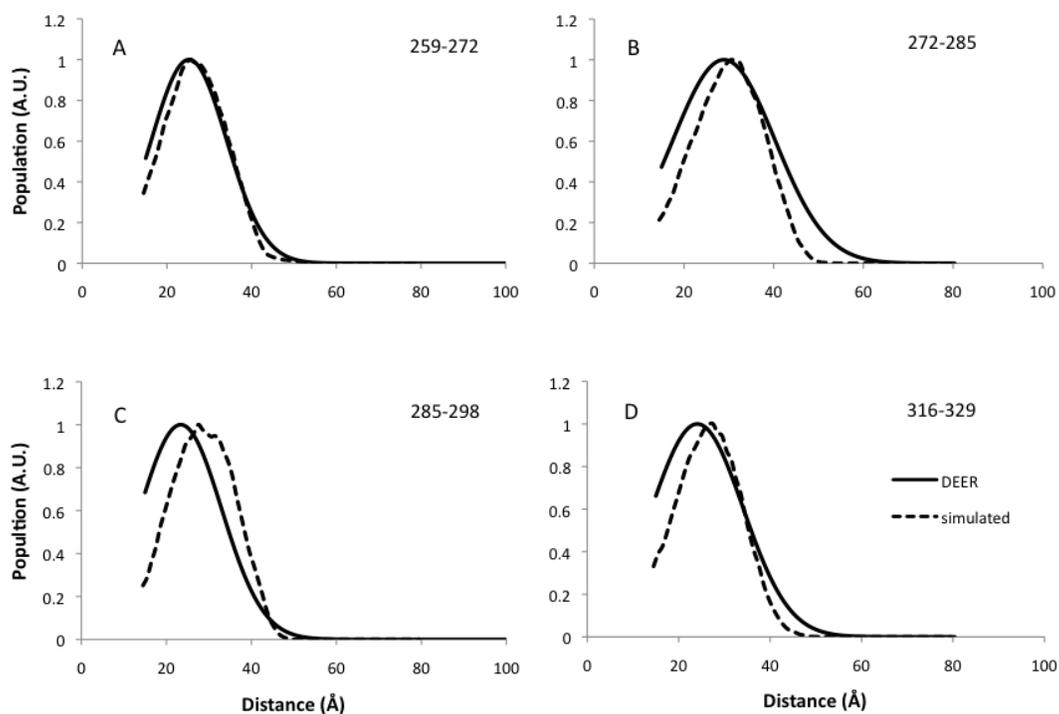


Figure S5: RotamerConvolveMD simulations of the expected distance distribution of MSTL spin labels in the ensemble of tau conformers (dashed line) shows good agreement with DEER measurements (solid line). Residue pair numbers are annotated in the upper right corner of each graph.

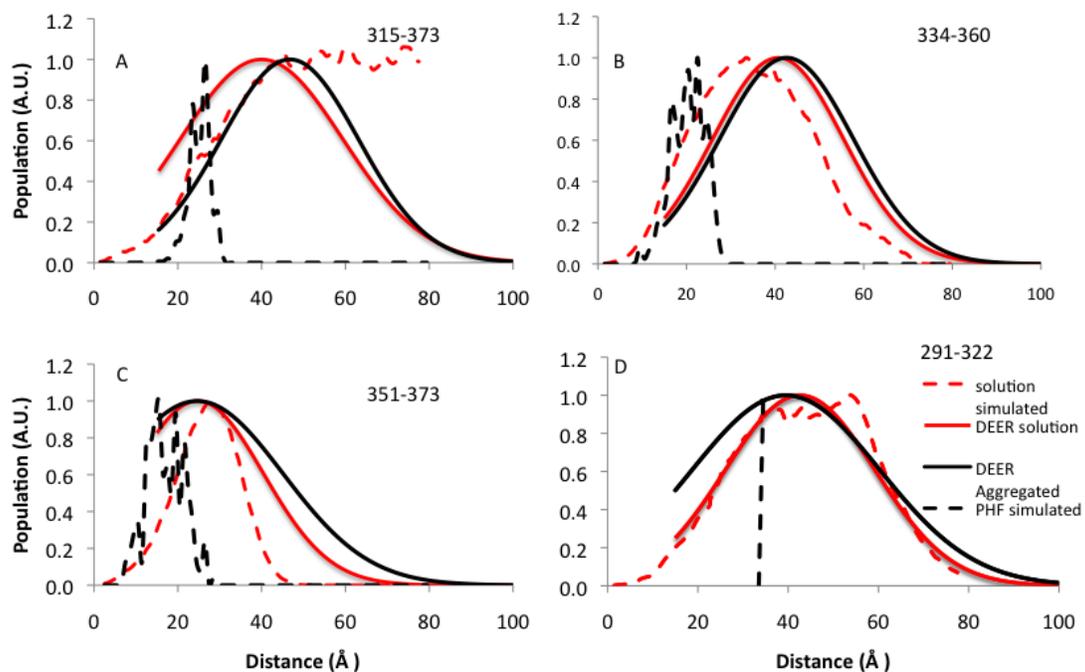


Figure S6: RotamerConvolveMD simulations of the expected distance distribution of MSTL spin labels in the ensemble of tau conformers (dashed red) and DEER measurements of unaggregated tau (solid red). A significant discrepancy is visible for sites 315-373 as the calculated large distances fall above the accessible DEER range. Simulated (dashed black) and measured (solid black) distance distributions for AD filaments (equimolar mixture of PHF and SF) and heparin filaments, respectively, are plotted as well. Residue pair numbers are annotated in the upper right corner of each graph.

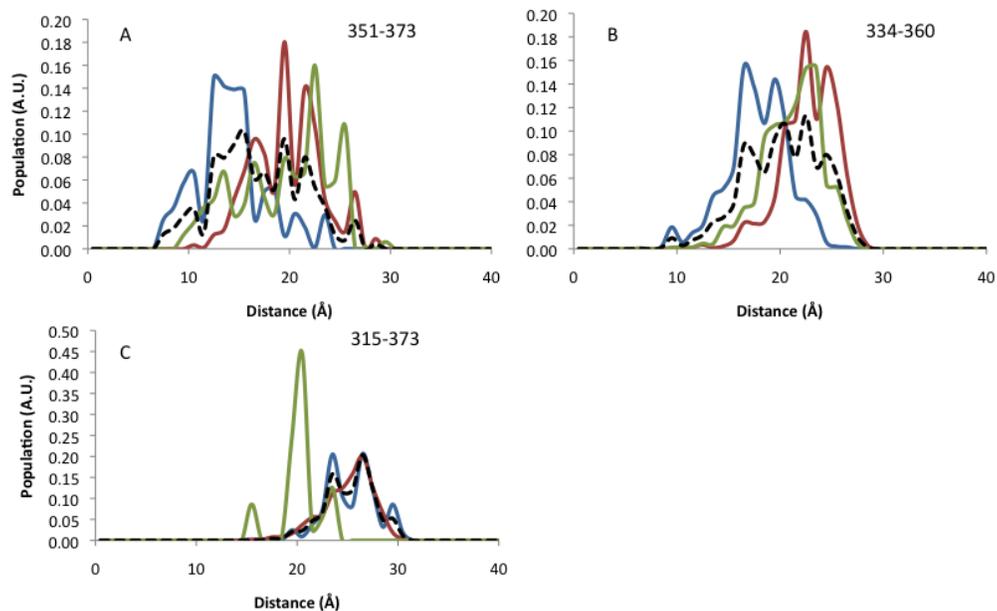


Figure S7: DEER distance distribution simulated for PHF (blue, PDB ID 5O3L), SF (red, PDB ID 5O3T), pronase treated PHF (green, PDB ID 5O3O), and a 1:1 mixture of SF and PHF (dashed black). A mixture of PHF and SF provides a broader distribution than any single conformation, and is used for comparison to DEER measurements throughout the text.

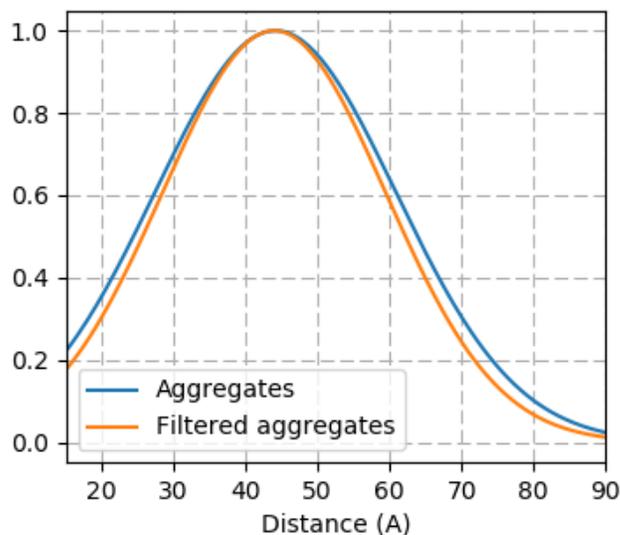


Figure S8: DEER distance measurement of aggregated G303C-A382C tau mutant following the same protocol as in figure 2 and 3 (blue curve), and after spinning down the sample for 10 min at 14,000 g in a 100kDa-cut off Amicon Centrifugal Concentrator (orange curve). The similarity of the two curves shows that the measured distances have no contribution from mono- to tetra-mer species.

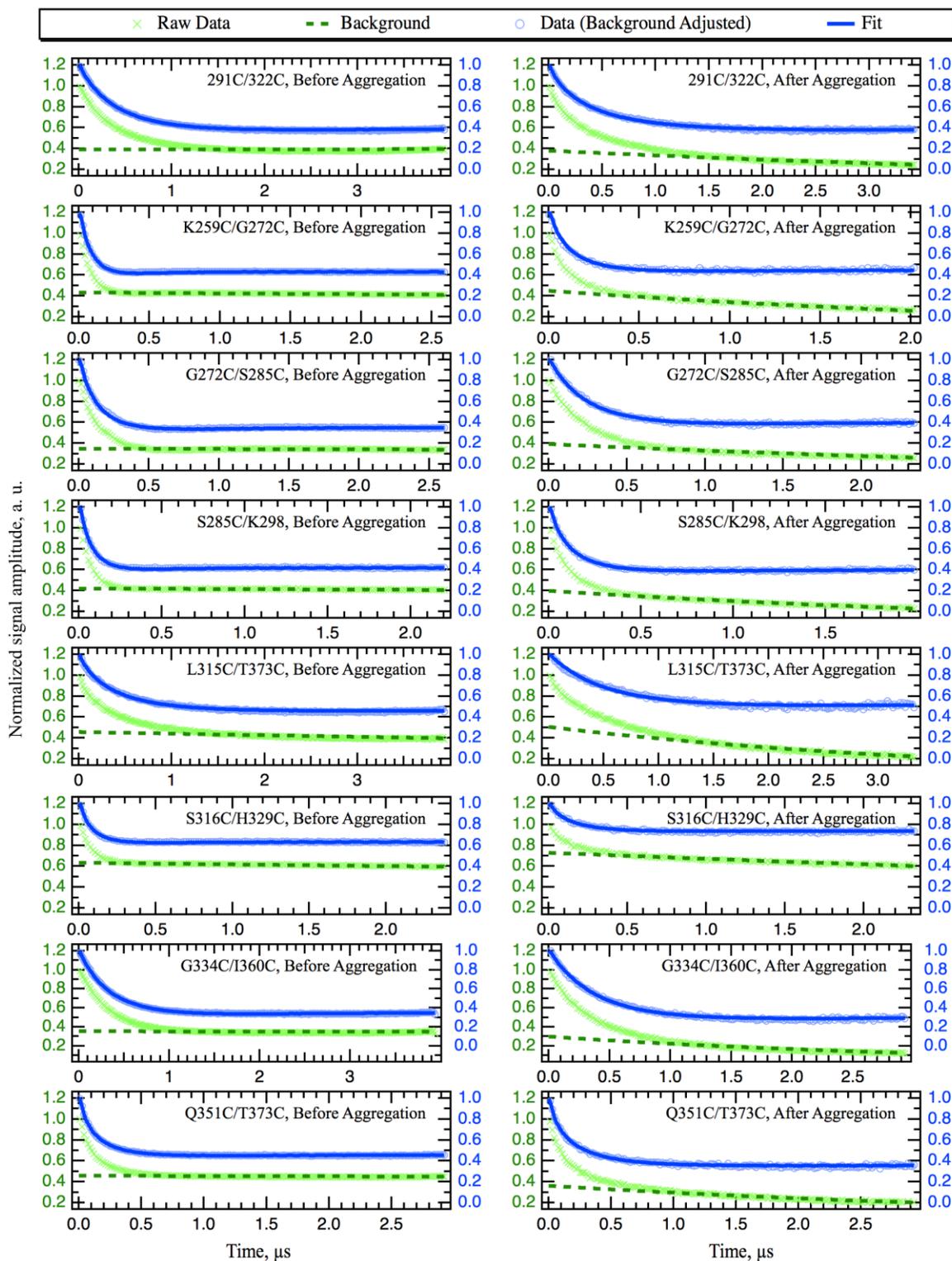


Figure S9: DEER time domain traces. The raw data (light green, left scale) are corrected with a 3D background (green, dashed). The background-subtracted data (light blue circles, right scale) are fitted (blue solid line) with a Gaussian-distribution model to obtain the distance distributions plotted in figure 2 and 3. Both raw and background subtracted data are normalized to 1.

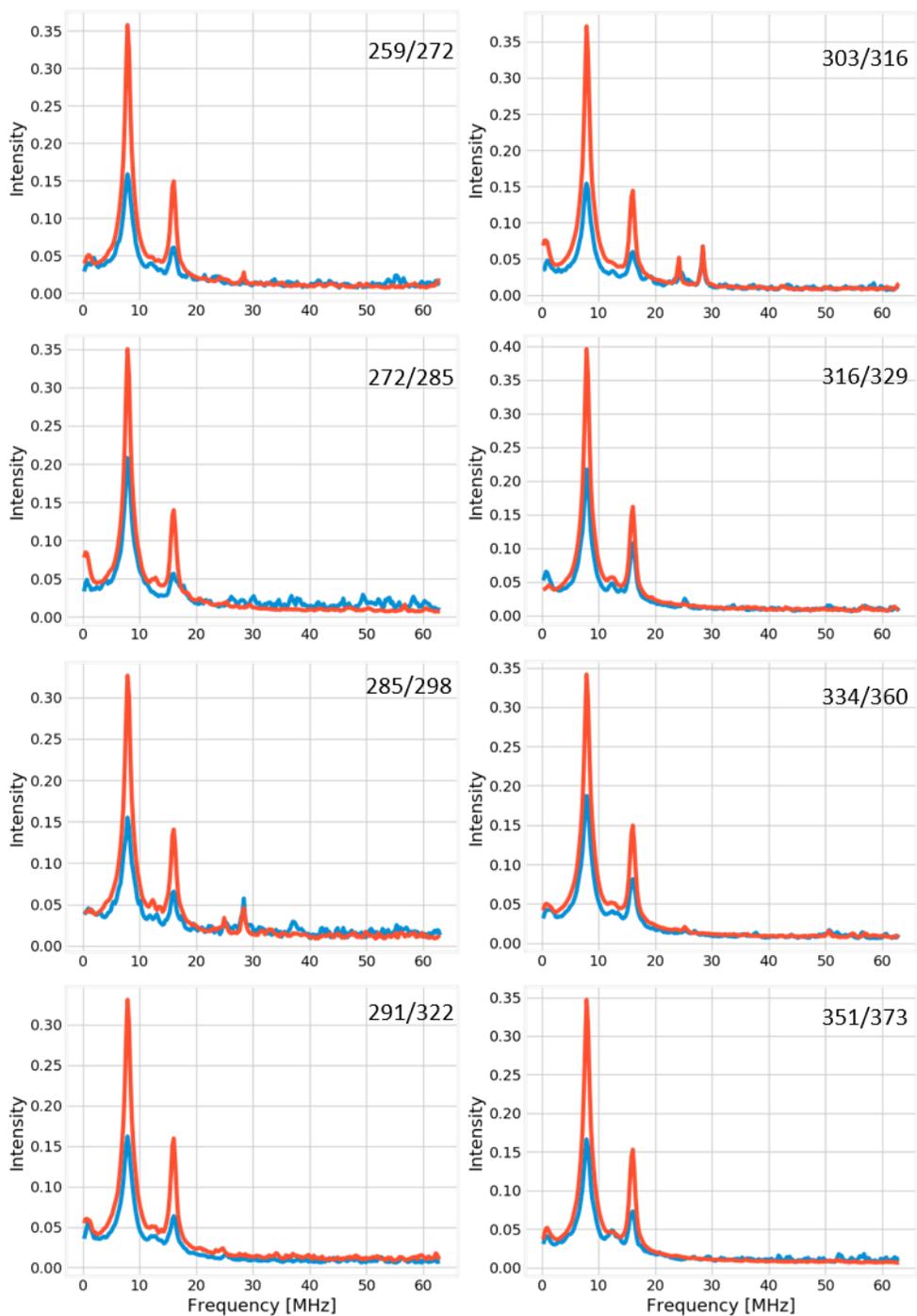


Figure S10: ESEEM spectra of tau before (red) and after (blue) aggregation. For each graph, the label positions are specified in the upper right corner. The two peaks around 7.9 and 15.8 MHz originate from ^2H ESEEM.

Reference :

- 1 N. Eschmann, E. Georgieva, P. Ganguly, P. Borbat, M. Rappaport, Y. Akdogan, J. Freed, J.-E. Shea and S. Han, *Sci. Rep.*, 2017, 44739.
- 2 Y. Fichou, N. A. Eschmann, T. J. Keller and S. Han, *Methods Cell Biol.*, 2017, **141**, 89–112.
- 3 M. Schwalbe, V. Ozenne, S. Bibow, M. Jaremko, L. Jaremko, M. Gajda, M. R. Jensen, J. Biernat, S. Becker, E. Mandelkow, M. Zweckstetter and M. Blackledge, *Structure*, 2014, **22**, 238–249.
- 4 A. W. P. Fitzpatrick, B. Falcon, S. He, A. G. Murzin, G. Murshudov, H. J. Garringer, R. A. Crowther, B. Ghetti, M. Goedert and S. H. W. Scheres, *Nature*, 2017, **547**, 185–190.
- 5 Y. Polyhach, E. Bordignon and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2356–2366.
- 6 L. S. Stelzl, P. W. Fowler, M. S. P. Sansom and O. Beckstein, *J. Mol. Biol.*, 2014, **426**, 735–751.