PNAS

Supporting Information for

Precision Proteoform Design for 4R Tau Isoform Selective Templated Aggregation

Andrew P. Longhini[1,2], Austin Dubose[3], Samuel Lobo[4], Vishnu Vijayan[3], Yeran Bai[1,2,5], Erica Keane Rivera[1,2], Julia Salar-Jarque[1,2], Arina Nikitina[1,2], Daniel C. Carrettiero[1,2,6], Matthew T. Unger[1,2], Olivia R. Sclafani[1,2], Valerie Fu[1,2], Emily R. Beckett[1,2], Michael Vigers[3], Luc Buee[7,8], Isabelle Landrieu[9,10] Scott Shell[4], Joan E. Shea[3,11], Songi Han[3,4,12], Kenneth S. Kosik[1,2,12,*]
Kenneth S. Kosik
Email: kosik@lifesci.ucsb.edu

This PDF file includes:

Supporting Text
Extended Methods
Figures S1 to S9
SI References

Other supporting materials for this manuscript include the following:

Supplementary Text

Cellular Effects of Seeded Tau Aggregates

To examine the possibility that tau aggregates were associated with lysosomes or autophagosomes, cells were stained with either a Lamp1 or a LC3 antibody. Small, punctate signals typical of lysosomes or autophagosome did not substantially co-localized with the aggregate signals (Figure 6C). The Manders' coefficient for tau to Lamp1 overlap was 0.12 ± 0.1, and for Lamp1 overlap to tau was 0.055 ± 0.05 (N=15 cells), indicating minimal overlap. For tau to LC3, the overlap was 0.41 ± 0.2, and LC3 to tau overlap was 0.44 ± 0.1 (N=8 cells), suggesting modest colocalization. Interestingly, a substantial overlap was observed between LC3 and the largest tau aggregates, with a Manders' coefficient of 0.90 ± 0.2 for tau overlap to LC3 under specific thresholds that capture only the largest tau aggregates (Figure S7A). Thus, there is some attraction of LC3 to the larger aggregates.

The large perinuclear aggregates of tau protein observed in our jR2R3 P301L-seeded cells hinted at the involvement of cellular degradation mechanisms beyond the capacity of traditional autophagy or proteasomal pathways to clear proteins (1), and thereby necessitate the deployment of alternative clearance strategies such as aggresomes (2, 3). To examine this possibility, we stained cells expressing aggregate induced by jR2R3 P301L with antibodies targeting SQSTM1 (also known as P62) and VCP (also known as P97). Both proteins have been implicated in previous studies as colocalizing with tau inclusions, suggesting a potential role in the degradation of tau aggregates (4–11). Immunofluorescence analysis (Figure S7B, S7C) revealed that both SQSTM1 and VCP exhibited pronounced colocalization with the large tau aggregates in the jR2R3 P301L-seeded cells. This observation aligns with existing literature that demonstrates the colocalization of SQSTM1 and VCP with tau inclusions (7–12), suggesting their involvement in the degradation of tau aggregates.

Dynamics and Heterogeneity of Tau Aggregates

Before the introduction of seeds, all cell lines—including WT and P301L mutants—displayed typical microtubule staining patterns (Figure 2A Left, S2B, S2C). In cells with mature aggregates, which can range in size from large (up to 20 μm) to minute, diffraction-limited specks (<250 nm), the population of tau bound to microtubules was greatly diminished (Figure 2C). Aggregates larger than 1 μm in size displayed a diverse range of morphologies, with some appearing elongated and others lumped together. This irregularity is reflected in a variable mean circularity value centered at 0.55 ± 0.25. The population of cells with the most aggregates were entirely devoid of tau microtubule staining (Figure 2E) and by four days in culture such cells dominated. In this experiment, cells actively expressing new monomeric tau preferentially partitions into aggregates rather than binding to existing microtubules suggesting that the aggregate has a greater affinity for tau than the microtubules. However, after staining with an anti-α-tubulin antibody, an intact microtubule network remained presumably associated with other microtubule-associated proteins (Figure S8A).

To systematically investigate the temporal evolution of tau aggregation, live cells were monitored for a period of 10 hours, with images of tau aggregates captured at 30-minute intervals. Aggregates did not appear to fuse but sometimes clumped together. Fragmentation of the aggregates was not observed. A few aggregates were spherical, but most were irregular in appearance. A few had a linear appearance that resembled the seeded aggregates previously observed in proximity to microtubules (13). Individual cells were tracked, allowing for the quantification of the total area occupied by aggregates, the average size of aggregates, and the number of aggregates per cell. To enhance data reliability, we excluded cells undergoing division or exhibiting signs of cell death.

Our analysis revealed that both the total area occupied by aggregates per cell and the number of aggregates per cell followed a bi-exponential distribution (Figure S8B). We then classified changes in total aggregate area within individual cells and found that the total aggregate area decreased in 48 cells, remained stable in 267 cells, and increased in 86 cells over 10 hours (Figure S8C). We applied K-means clustering analysis to the trajectories of

individual cell total area, effectively segregating the data into three distinct subpopulations based on aggregate dynamics (Figure S8D). We plotted these subpopulations and observed clear patterns of decline, stability, and growth in aggregate areas for the respective clusters (Figure 8E).

Extended Methods

Cloning of Tau Constructs

We assembled mClover3-2N4R-WT constructs using the HiFi Reaction Kit from NEB. Additionally, we generated various isoforms such as mClover3-0N3R-WT, mClover3-0N4R-WT along with truncations including mClover3-Tau187(4R)-WT, mClover3-Tau187(3R)-WT through deletion PCRs and KLD reactions using NEB kits.

Using site-directed mutagenesis, we created mutant constructs, specifically mClover3-Tau187P301L, mClover3-0N4R-P301L, and mClover3-2N4R-P301L. Additionally, using similar techniques, we created mClover3-Tau187(3R)-P270L, mClover3-Tau187(3R)-P270L-E264D, mClover3-Tau187(3R)-P270L-E264D-L266I, mClover3-Tau187(3R)-P270L-E264D-L266I-V269Q, and mClover3-Tau187(3R)-P270L-E264D-L266I-V269Q-S274K

Constructs with dual tau expression, such as mRuby3-Tau187-P301L-P2A-mClover3-Tau187P301L, mRuby3-Tau187(3R)-P2A-mClover3-Tau187-P301L, and mRuby3-Tau187(3R)-P2AmClover3-Tau187(3R), were synthesized using HiFi reactions with the previously mentioned constructs as templates.

We procured pET28a-Tau187-WT-His and pEt28a-Tau187-P301L-His constructs from Genescript and constructed pET28a-Tau187(3R) using deletion PCR.

For DEER experiments, we generated Tau187 Cysless constructs with one or two cysteine mutations using site-directed mutagenesis.

Tau187 Purification

We transformed E. coli BL21 (DE3) cells with various plasmids and stored the cells as frozen glycerol stocks at −80°C. An overnight growth of cells in 10 mL Luria Broth preceded the inoculation of 1 L of fresh LB. The incubated the cells at 37°C at 200 rpm, supplementing with 10 µg/mL kanamycin, until an optical density at λ = 600nm of 0.6–0.8 was achieved. We induced expression by introducing 1 mM isopropyl-ß-D-thiogalactoside for 2–3 h. Post expression, cells were harvested via centrifugation at 4500 g for 20 min.

Harvested cells were resuspended in lysis buffer (Tris-HCl, pH = 7.4, 100 mM NaCl, 0.5 mM DTT, 0.1 mM EDTA) supplemented with 1 Pierce protease inhibitor tablet per 5 mL lysis buffer, 1 mM PMSF, 2 mg/mL lysozyme, 20 µg/mL DNase and 10 mM MgCl2, and incubated on ice for 30 min. The samples underwent three freeze-thaw cycles using liquid nitrogen, followed by 10-minute centrifugation at 10,000 rpm to discard cell debris. Another, 1 mM PMSF was added, and samples were then heated at 65°C for 12 minutes and cooled on ice for 20 minutes before undergoing another round of centrifugation to remove the precipitant.

The clear supernatant was incubated with pre-equilibrated Ni-NTA resins overnight in buffer A (20 mM sodium phosphate, pH = 7.0, 500 mM NaCl, 10 mM imidazole, 100 µM EDTA). We subsequently transferred the resins to a column and washed with 20 mL of buffer A, 25 mL buffer B (20 mM sodium phosphate, pH = 7.0, 1 M NaCl, 20 mM imidazole, 0.5 mM DTT, 100 µM EDTA). We collected the purified protein through elution with 15 mL of buffer C (20 mM sodium phosphate, pH = 7.0, 0.5 mM DTT, 100 mM NaCl, 300 mM imidazole). SDS-PAGE confirmed the purity of the fractions. After buffer exchange into a DTT-free working buffer (20 mM ammonium acetate, pH 7.0), we immediately froze the proteins and stored them at −80°C.

Thioflavin T Experiments

Thioflavin (ThT) experiments were conducted using a BioTek Synergy 2 fluorescent plate reader. Each well of a 384-well plate (Corning, low volume non-binding surface, black with clear flat bottom) contained 50 µM of SLS or Tau monomer, 20 µM ThT, 12.5 µM heparin or 12% seed, 20 mM HEPES buffer (pH 7.4), making a total volume of 20 µL per well.

The plate reader was set to a temperature of 37°C and allowed time to equilibrate. Subsequently,

ThT fluorescence intensity was recorded at excitation and emission wavelengths of 440 nm and 485 nm, respectively. Measurements were taken every 2 minutes until a plateau in fluorescence intensity was observed. Each experiment was performed with five replicates and repeated three times using independent samples, typically spanning a total time of 48 hours.

TEM Analysis

For transmission electron microscopy (TEM) analysis, 5 µL of recombinant tau fibril samples were placed onto a glow-discharged copper grid (Electron Microscopy Science, FCF-200-Cu) for 20 seconds before blotting dry with filter paper. The samples were then stained by applying 5 µL of 1.5% (w/v) uranyl acetate solution, immediately blotting dry. An additional 5 µL of uranyl acetate solution was added for 60 seconds, followed by blotting dry. The samples were examined using a Thermo Scientific Talos G2 200X TEM/STEM microscope operating at 200 kV and room temperature. The grids were imaged using a Ceta II CMOS 4k x 4k camera at the specified magnifications.

Fluorescence-guided OPTIR imaging setup, spectral measurement, and data processing.

The mid-IR laser used as the photothermal pump source was pulsing at 100 kHz repetition rate with 1% duty cycle. The visible probe source was a continuous wave laser with a center wavelength of 532 nm. The counter-propagation geometry of the mid-IR and the visible beam was used. The mid-IR beam was focused on the sample plane below the sample substrate (CaF_2) with a reflective objective (40x, 0.78NA, Pike Technologies), and the visible beam was focused on the sample plane from the top with a refractive objective (50x, 0.8NA, Olympus). Epi-detected light was collected and focused on a photodiode. The OPTIR signal was demodulated with a lock-in amplifier at 100 KHz. The system was enclosed and purged under gentle nitrogen flow to minimize water vapor interference of the spectra interpretation. For co-registered widefield fluorescence imaging, a filter cube set suitable for mClover3 excitation and emission was used. Typical power at the sample for mid-IR and visible were a few mW depending on the sample locations.

All raw spectra were normalized with mid-IR power spectra to reveal the sample-related IR peaks. For the OPTIR spectra presented in the manuscript, OPTIR spectra were normalized to OPTIR intensity at 1654 cm^{-1}. For secondary structure quantification, β-sheet to α-helix ratio was calculated using OPTIR intensity at 1634 cm^{-1} and 1654 cm^{-1}. Lipid contents were quantified by normalizing OPTIR intensity at 1740 cm^{-1} to 1654 cm^{-1}. For each cell, we acquired at least 8 spectra at different puncta and puncta-free locations, where the mean intensity of the ratioed value was used to represent a single cell. Ten cells were used to perform statistical analysis.

Object Segmentation and Tracking

All image processing was carried out using Python 3.10.12. The time series of cell images featuring puncta were loaded utilizing the OpenCV-python 4.7.0.72 package. The analysis of these images was conducted using scikit-image 0.19.3 package.

Initially, nuclei were identified by applying Li thresholding (14) to the Hoechst intensity image. Objects with size smaller than 100 μm^2 were excluded from further analysis as noise, and the remaining connected regions were labeled as separate nuclei.

Subsequently, the mClover3-Tau187 fluorescence intensity images, after being subjected to blurring, were used for the detection of cell bodies through mean thresholding (15). The masks for nuclei and cell bodies were combined to form complete cell masks. The watershed algorithm was employed, using labeled nuclei as seeds, to delineate individual cells.

Cells obtained through this process were tracked across successive frames using centroid tracking. In this approach, a centroid in a new frame was associated with the nearest centroid in the preceding frame, marking them as the same object.

Puncta were detected using the Laplacian of Gaussian blob detection method. In each frame, puncta located within the mask of a cell object were assigned to that cell.

Computational Modeling

REMD Simulations

Replica-exchange molecular dynamics (REMD) simulations were carried out using the Gromacs package (versions 2019.6 and 2020.1) (16,17). Simulations were conducted in rhombic dodecahedral boxes of approximately 7.5 nm length, comprising 19-mer peptide(s), around 7860 TIP4P-D water molecules, and neutralizing chlorine ions. Peptides were simulated with zwitterionic termini. In the case of monomer simulations, the peptides began in extended conformations. For dimer simulations, the peptides were initialized in the two most common monomer conformations based on a Daura clustering analysis (18), and they were separated by a couple of nanometers.

The force field AMBER99SB-disp (19) was used, and periodic boundary conditions were applied along all three box vectors. The simulations included an initial 5 ns run in the NPT ensemble at 1 bar pressure and 300 K. The v-rescale thermostat (0.5 ps time constant for temperature coupling) and the Parrinello-Rahman barostat (3 ps time constant for pressure coupling) were used for this step. Next, 5 ns long NVT simulations with the Nose-Hoover thermostat (3 ps time constant for temperature coupling) were performed at 300 K temperature using the average box-size from the final three ns of the previous NPT step (20,21). Throughout the simulations, the Particle Mesh Ewald (PME) method was used for calculating the electrostatic interactions with a grid-spacing of 0.12 nm (22). A time step of 2.5 fs was employed using a leap-frog integrator (23). A cutoff distance of 1.2 nm was set for all nonbonded interactions. For simulations excluding MTSL, the LINCS algorithm was used to constrain bonds with hydrogens (24). Water molecules were maintained fully rigid using the SETTLE algorithm (25).

Subsequently, 60 replica systems were created at temperatures ranging from 300 K and 456 K: The temperatures of the replicas were: 300.0, 301.9, 303.8, 305.7, 307.7, 309.7, 311.7, 313.7, 315.8, 317.9, 320.0, 322.1, 324.3, 326.5, 328.7, 330.9, 333.2, 335.5, 337.8, 340.1, 342.5, 344.9, 347.3, 349.7, 352.2, 354.7, 357.2, 359.7, 362.3, 364.9, 367.5, 370.1, 372.8, 375.5, 378.2, 380.9, 383.7, 386.5, 389.3, 392.1, 395.0, 397.9, 400.8, 403.7, 406.7, 409.7, 412.7, 415.8, 418.9, 422.0, 425.2, 428.4, 431.6, 434.9, 438.2, 441.6, 445.0, 448.5, 452.0, 455.6 K. Each replica underwent a 5 ns NVT equilibration run per without exchanging replicas.

Following this, the replicas were subjected to 5 ns equilibration runs with replica exchanged occurring every 3 ps, achieving an average exchange probability of about 30%. This was followed by production REMD simulations ranging from 500-750 ns per replica. The initial 40 ns were discarded for further equilibration, and certain simulations employed longer equilibration times as discussed in the next paragraph.

Equilibration times were determined using a heuristic that maximizes the number of effectively uncorrelated measurements. The autocorrelation function was integrated for equilibration times in 10 ns increments starting from 50 ns, and the equilibration time with the most effectively uncorrelated measurements was chosen. Dimer simulation equilibration times ranging from 80 to 140 ns were selected by maximizing the number of effectively uncorrelated intermolecular hydrogen bond counts. Autocorrelation functions were measured using pymbar's timeseries module. Confidence intervals were determined with bootstrapping. Residues were assigned secondary structures using the DSSP algorithm in mdtraj. ChimeraX and Pymol were employed for visualizing conformations. Hydrogen bonding analysis was conducted in MDAnalysis, utilizing distance cutoffs of 3.5 Å and angle cutoffs of 140° for determining hydrogen bonds. Energy landscapes were generated by measuring the 2D probability density, taking a logarithm, scaling by Boltzmann's constant and temperature, and applying Gaussian interpolation to smooth the landscape.

After running the REMD simulations we clustered the jR1R3 and jR2R3 monomer conformations using the Daura algorithm, and the top two clusters had two pairs of intramolecular backbone hydrogen bonds either between K298-Q307 & aa305-aa300 for Cluster 1 or K298-V309 & Q307-aa300 for Cluster 2. We performed the Daura algorithm with GROMACS (`gmx cluster … -method gromos`); and selected backbone coordinates and used a 2Å RMSD cutoff for clustering (18).

Water Triplet Angle Analysis

The water triplet angle distribution refers to the distribution of angles formed by the oxygen of a central water and the oxygens of each pair of its neighboring water molecules. This distribution has been demonstrated to reveal patterns in the structing of water in response to different solutes and interfaces (26,27). In this study, the distribution of hydration water was computed to be 4.25 Å around the heavy atoms of each residue's backbone or side chain. The 4.25 Å cutoff corresponds to the second minimum in the radial distribution function (RDF) between the heavy atoms of the backbone or side chain and the oxygen atoms of water. Neighboring water molecules were defined as those having oxygen atoms within 3.4 Å of the central water molecule. The hydrophilic signature was defined as the fraction of hydration waters around each peptide that form angles between 4550° with adjacent waters, normalized by the fraction of bulk waters forming angles between 45-50° with adjacent waters. Conformations were analyzed based on data extracted every 10 ps from REMD simulations. Each frame contained statistically independent water triplet angle measurements, which were based on measured autocorrelation times. Confidence intervals of 67% were determined though bootstrapping with over 1000 independent measurements taken with replacement.

Spin Labeling of Tau

Tau187, with its native cysteines mutated to serines and the specified residues mutated to cysteines, was expressed and spin-labeled to investigate distances between two sets of target residues (residues 351 and 373, and residues 334-360). The native cysteines at 291 and 322 were mutated to serines. Additionally, a similar tau construct without any cysteines (referred to as tau-cysless) was expressed and purified. The protein was spin-labelled using MTSL ((1-Acetoxy-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl) Methanethiosulfonate), which was purchased from Toronto Research Chemicals.

Before the labeling process, the samples were treated with a10x molar excess of DTT to reduce any cysteine residues. The DTT was subsequently removed using a PD-10 desalting column. Subsequently, a 10x molar excess of MTSL was added to the free cysteine residues and the mixture was incubated with the protein at 4˚C overnight. After the incubations, excess MTSL was removed using a PD-10 desalting column.

The labelling efficiency, which is defined as the molar ratio of attached spin labels to cysteine residues, was found to be between 50-60% for double-cysteine mutants and approximately 90% for the single-cysteine mutant.

Sarkosyl Extraction from Tau187 Cell Lines and AD Brains

jR2R3 P301L fibrils were transfected in H4 cells that overexpressed the mClover3-Tau187-P301L construct, as previously described. After incubating the transfected cells overnight, the cells were washed and collected. Tau187 puncta were extracted from a solution containing 2% sarkosyl in A68 buffer (10 mM Tris-HCl pH 7.5, with 10% sucrose, 0.8 M NaCl, 1mM EGTA). The A68 buffer was supplemented with PMSF and a protease inhibitor. A manual cell scraper was employed to collect the cell extracts. The samples were then sonicated for 15 seconds at 30% amplitude and incubated at 37˚ C for 30 minutes. After incubation, cell extracts were centrifuged at 113,000 g for 20 minutes at 25˚C. The sarkosyl-soluble fraction was collected, and the remaining pellets were washed with 30mM Tris-HCL (pH 7.5). The samples were then centrifuged again at 113,000 g for 20 minutes at 25˚C. The sarkosyl-insoluble fractions were resuspended in 30 mM Tris-HCl (pH 7.5) and sonicated for 15 seconds at 30% amplitude.

Brain samples used in this study were sourced from the New York Brain Bank at Columbia University. The specific region of the brain from which the samples were obtained was the BA38 temporal pole, and the specimens showed neuropathologic changes associated with Alzheimer's disease. Before processing, the samples were equilibrated to -20˚C for 30 minutes, and then sliced into 0.6 micron-wide sections to accumulate a total of 0.5 grams of material. The brain tissue was homogenized in 2% sarkosyl in A68 buffer (10 mM Tris-HCl, pH 7.5, containing 10% sucrose, 0.8 M NaCl, 1 mM EGTA). At 6% (w/v) using a glass Dounce homogenizer. PMSF and

protease inhibitors were added to the A68 buffer. After homogenization, the samples were centrifuged at 113,000 g for 20 minutes at 25˚C. The sarkosyl-soluble fraction was collected, and the pellets were washed with 30 mM Tris-HCl (pH 7.5). Samples were centrifuged again at 113,000 g for 20 minutes at 25˚C. The sarkosyl-insoluble fractions were resuspended in 30 mM Tris-HCl (pH 7.5) and sonicated for 15 seconds at 30% amplitude.

Western Blotting

The protein concentrations of the samples were determined using a Pierce BCA Protein Assay Kit from Thermo Fisher Scientific. Standard SDS-PAGE was performed using both the sarkosyl soluble and sarkosyl-insoluble fractions derived from cell and AD brain extractions. The samples were loaded on a 4–20% Mini-PROTEAN® TGX™ Precast Protein Gel from Bio-Rad. Precision Plus Protein Dual Color Standards ladder from Bio-Rad was used as a reference for molecular weights. For the western blot staining, the MC1 antibody hybridoma conditioned media was employed, which was kindly provided by the lab of Peter Davies. The immunoblots were imaged using a Licor Gel Imaging Station and subsequently analyzed with Fiji imaging software.

DEER

To prepare the heparin fibrils for the DEER analysis, a sample mixture was created by mixing tau187(spin-labeled) and tau-cys-less at a 1:20 molar ratio (5 μM:100 μM) molar ratio and incubating it with heparin at a 1:4 molar ratio of tau to heparin. The mixture was incubated at 37°C for 24 hours. For preparing jR2R3 P301L seeded tau fibrils, instead of heparin, 50% seeds (by molar percentage) were added to the same mixture and incubated at 37°C for 24 hours. The formation of fibrils was verified using ThT binding assay and TEM analysis. After fibrillization, the fibrils were centrifuged at 15000 rpm to pellet them, and the supernatant was discarded. The pellets were washed three to four times with D_2O based 20 mM HEPES buffer to eliminate any remaining monomers. The final pellets were resuspended in 28 μL of D_2O based 20 mM HEPES buffer, mixed with 12 μL D8-glycerol (making up 30% of the total volume), and transferred to a 2 mm diameter quartz tube before being flash-frozen using liquid nitrogen.

The four-pulse DEER experiments were conducted at 65 K on a Q-band Bruker E580 Elexsys pulse EPR spectrometer operating at approximately 34 GHz and equipped with a 300 W Traveling Wave Tube (TWT) amplifier. The DEER pulse sequence used was: $p_{obs}/2 - t_1 - p_{obs} - (tp_{pump}) - (t_2\text{-}t) - p_{obs} - t_2 - echo$. Rectangular observe pulses were used with lengths set to $p_{obs}/2$ = 10-12 ns and p_{obs} = 20-24 ns. A chirp πpump pulse was applied with a length of 20-24 ns and a frequency width of 133 MHz. The observed frequency was set 150 MHz higher than the center of the pump frequency range. t_1 was 180 ns and t_2 was varied between 1.8 ms and 2.4 ms. The DEER experiment was accumulated over ~approximately 12 hours.

The DEER data analysis was carried out using LongDistances software with Tikhonov regularization. The background signal, which arises from non-specific spin-spin interactions other than the intramolecular spin pairs and can interfere with the data, was minimized by using spin dilution with non-spin labeled tau. Additionally, singly spin-labeled tau (at site 351) was used to prepare fibrils, and the resulting signal was fit to determine the background dimensionality, which was found to be approximately 2 for the fibrils. These values were used for background subtraction for the respective fibrils. Other parameters were kept constant for all data to enable accurate comparison between samples. The background-corrected signal was transformed to a distance distribution, $PP(RRRRRR)$, by Tikhonov regularization with non-negativity constraint employing LongDistances software by Altenbach. The error in $PP(RRRRRR)$ was estimated using the bootstrapping method available in the LongDistances software.

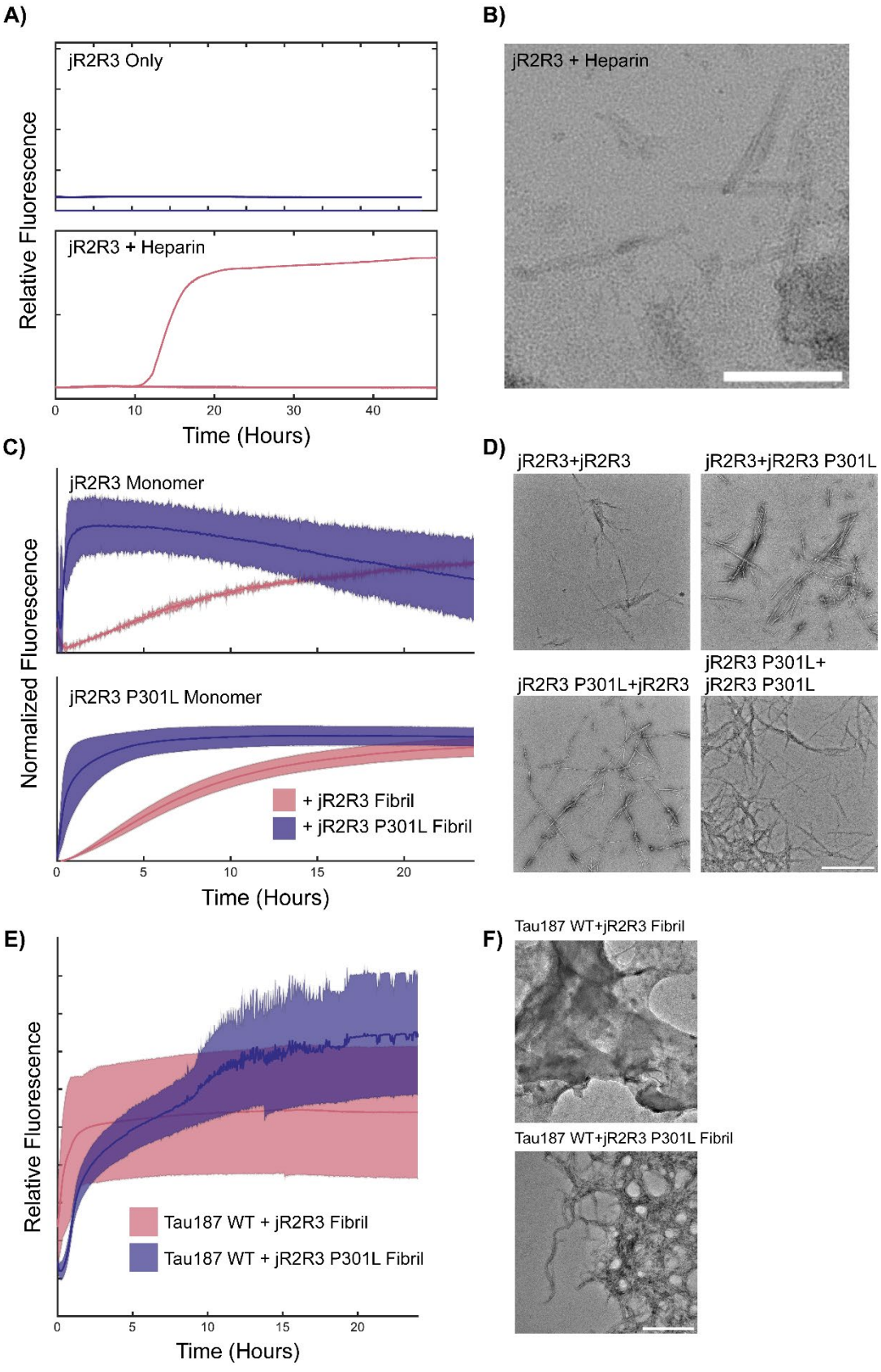
A)
jR2R3 Only
jR2R3 + Heparin
Relative Fluorescence
0
10
20
30
40
Time (Hours)
B)
jR2R3 + Heparin
C)
jR2R3 Monomer
jR2R3 P301L Monomer
+ jR2R3 Fibril
+ jR2R3 P301L Fibril
Normalized Fluorescence
0
5
10
15
20
Time (Hours)
D)
jR2R3+jR2R3
jR2R3+jR2R3 P301L
jR2R3 P301L+jR2R3
jR2R3 P301L+
jR2R3 P301L
E)
Tau187 WT + jR2R3 Fibril
Tau187 WT + jR2R3 P301L Fibril
Relative Fluorescence
0
5
10
15
20
Time (Hours)
F)
Tau187 WT+jR2R3 Fibril
Tau187 WT+jR2R3 P301L Fibril

Fig. S1. Monitoring Aggregation of Tau Constructs by ThT and TEM A) Among five jR2R3 reactions, only one demonstrated aggregation over a 48-hour period post-heparin addition (rose), as indicated by ThT. Isolated jR2R3 displayed no observable aggregation (purple). B) TEM imaging of heparin-seeded jR2R3fibrils revealed short, fragmented fibrillar structures. Scale bar: 200 nm. C) Both jR2R3 (rose) and jR2R3 P301L (purple) effectively seeded fibril formation in naïve monomers of themselves and of each other, with jR2R3 P301L promoting faster aggregation in both cases. D) TEM images of the four seeding reactions demonstrate clear fibrillar assemblies in each case. Scale bar: 500 nm. E) ThT fluorescence assays for Tau187 WT (rose) and Tau187 P301L (purple), demonstrating fibril formation when seeded by jR2R3. F) Corresponding TEM images of the seeding reactions, revealing the presence of fibrillar structures. Scale bar: 300 nm.

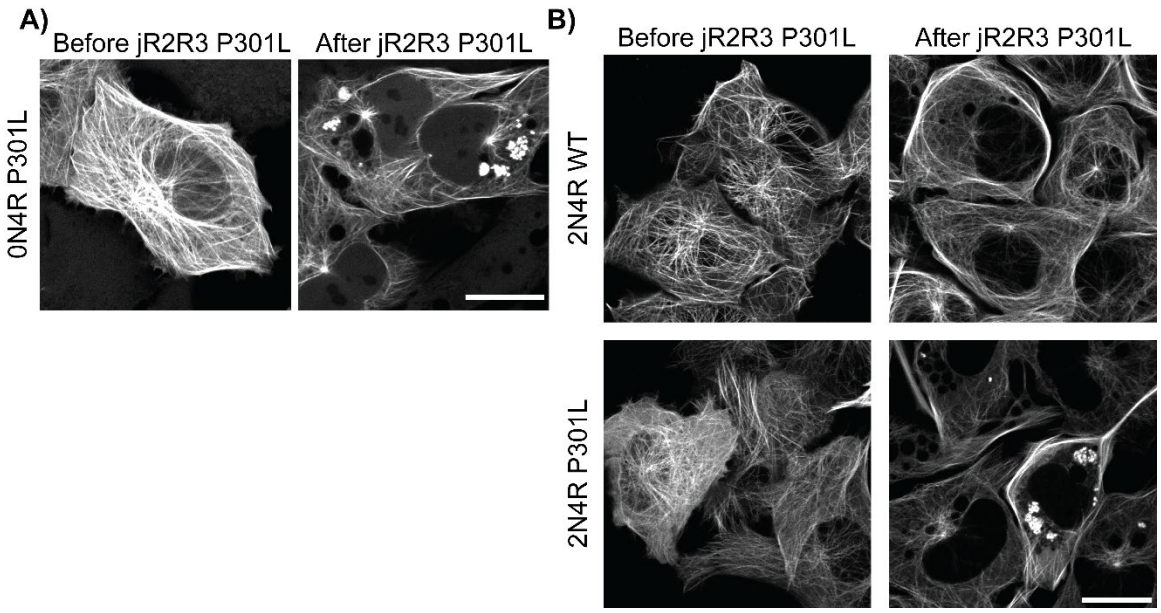

Fig. S2. Cellular Seeding of jR2R3 and jR2R3 P301L A) 0N4R P301L was seeded with jR2R3 P301L and puncta formation was observed after 12 hours. Puncta were observed in 13 ± 6% of cells. Scale bar: 20 μm. B) 2N4R WT and 2N4R P301L were seeded with jR2R3 P301L and puncta formation was observed after 12 hours. 2N4R had the least observed cells with at least one punctum, with 2N4R WT having 2 ± 1.7% and 2N4R P301L having 12 ± 2.4%. Scale bar: 20 μm.

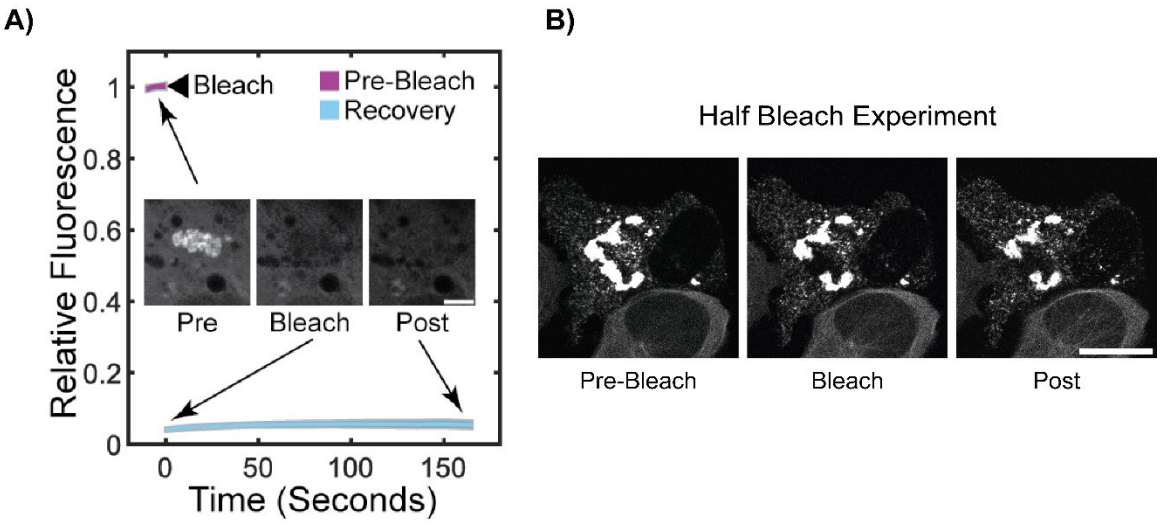

Fig. S3. FRAP Analysis of Tau Assemblies A) FRAP confirmed the presence of a highly immobile tau population within the puncta structures. Scale bar: 15 μm. B) Larger tau assemblies were partially subjected to FRAP, revealing no observable recovery. Scale bar: 15 μm.

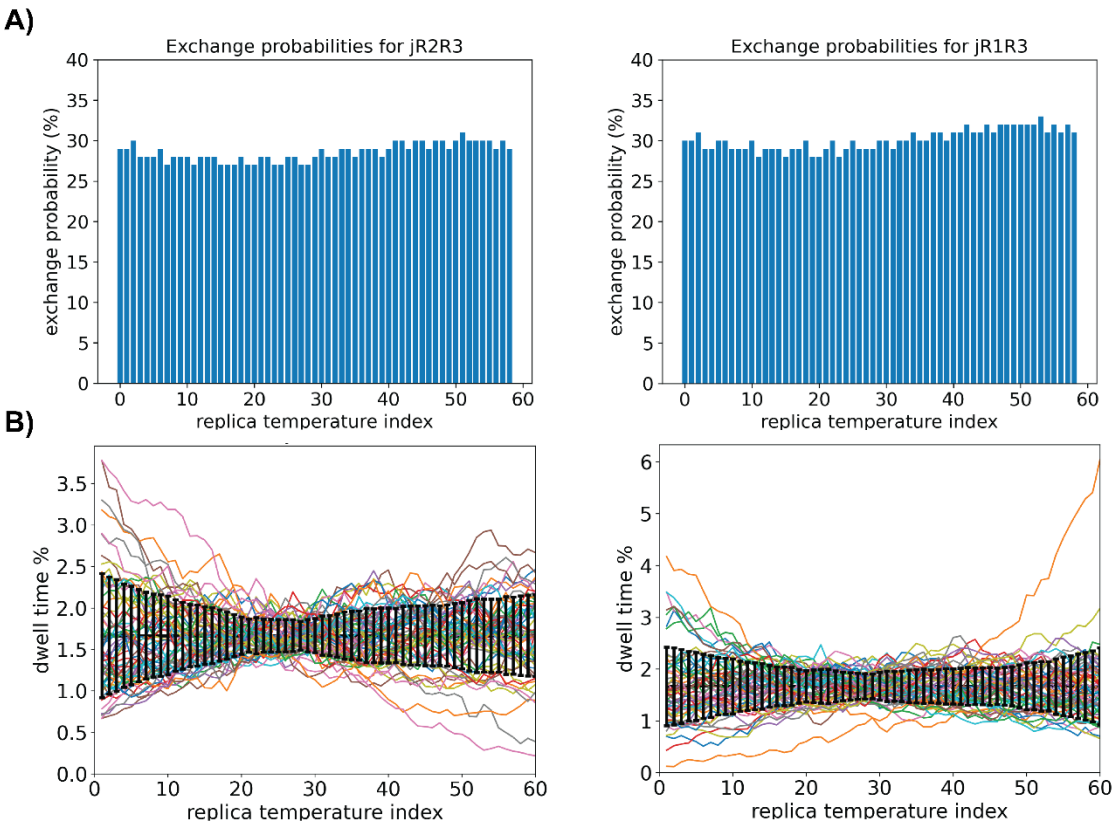

Fig. S4. Analysis of Replica Exchange Convergence A) Exchange probabilities for jR2R3 and jR2R3 P301L were plotted against the replica exchange temperature index. Uniform and high exchange probabilities were observed across all temperatures. B) Dwell times for jR2R3 and jR2R3 P301L plotted as a function of replica exchange temperature index showed a uniform sampling across all temperatures.

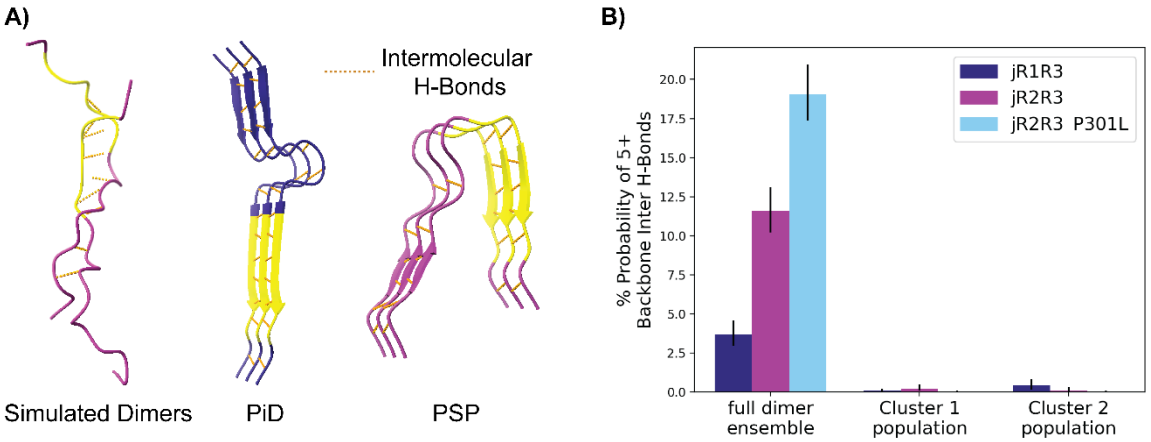

Fig. S5. Dimer Simulations Show Increased Intermolecular Hydrogen Bonding A) Representative conformations of dimer, with intermolecular hydrogen bonds depicted as yellow dots. A dimer conformation with both monomers in open states and featuring nine intermolecular backbone hydrogen bonds is shown on the left, compared to the structures from PiD (middle) and PSP (Right). B) Analysis of dimer simulations for jR2R3, jR2R3 P301L, and jR1R3, showing differences in intermolecular hydrogen bonding. jR1R3 exhibits reduced intermolecular hydrogen bonding compared to jR2R3 and jR2R3 P301L. Bars represent 90% confidence intervals. Oligomers with high hydrogen bonding (i.e., 5 or more intermolecular backbone hydrogen bonds) are rarely formed when one monomer is in a closed conformation.

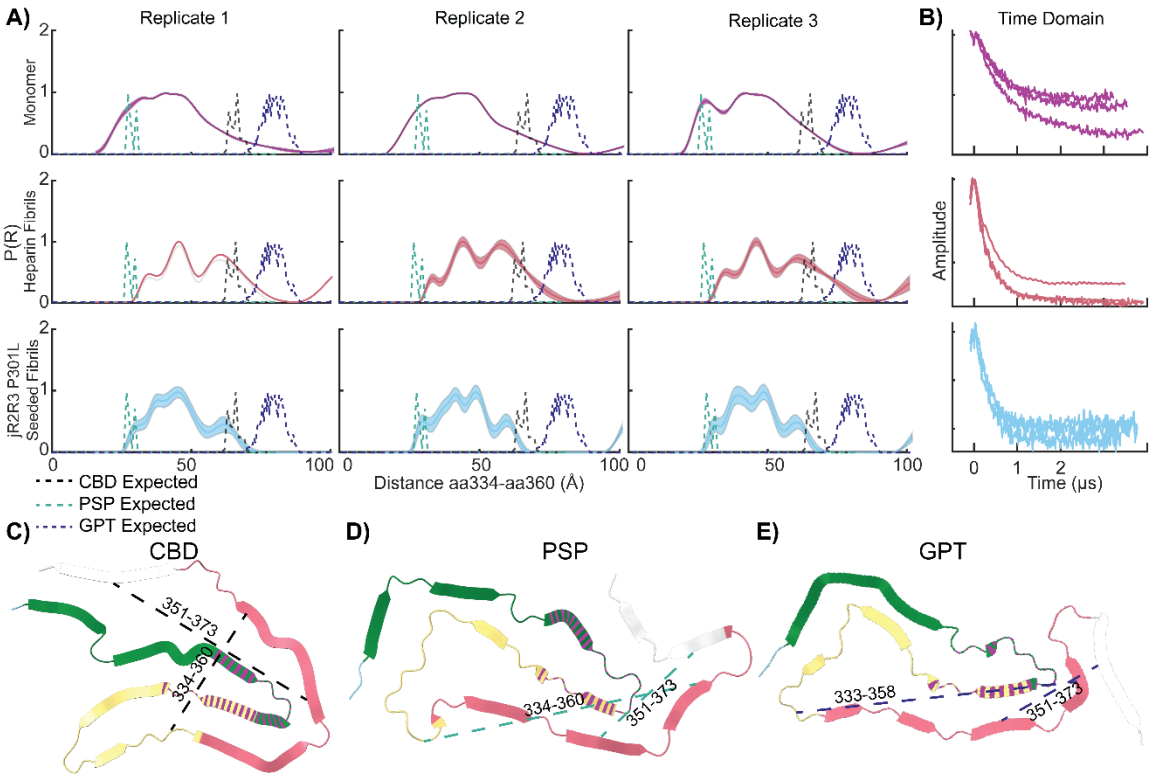

Fig. S6. Comparative Analysis of DEER Measurements for Monomeric and Fibrillar Tau Configurations aa334-aa360 Configurations A) DEER distance distribution plots (P(R)) for the 334-360 spin-pair distances in various tau configurations. The monomer is represented in purple, jR2R3 P301L seeded fibril in cyan, heparin fibril in rose, while the predicted distances for CBD are in black and PSP are in teal. Triplicates of each condition are plotted across the columns. B) The time domain data for each of the replicates plotted in (A). C) Hypothetical representation of distances expected if a pure CBD fold was sampled. D) Hypothetical representation of distances expected if a pure PSP fold was sampled. The figure provides insight into the conformational distinctions between tau monomers, jR2R3 P301L-seeded fibrils, and pathological tau configurations.

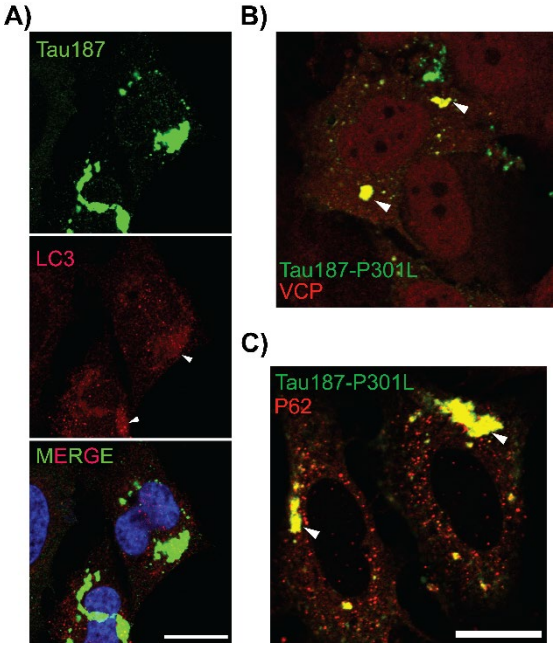

Fig S7: Aggregate Colocalization of Autophagosomal and Aggresome Markers A) Tau (green) co-localization with LC3 (red) indicates significant overlap with the largest tau aggregates (white arrows), suggesting a possible early stage of autophagy recruitment. Scale bar: 20 μm. B, C) Colocalization of protein degradation markers VCP (red, B) and P62 (red, C) with large tau aggregates (green). Scale bar: 20 μm.

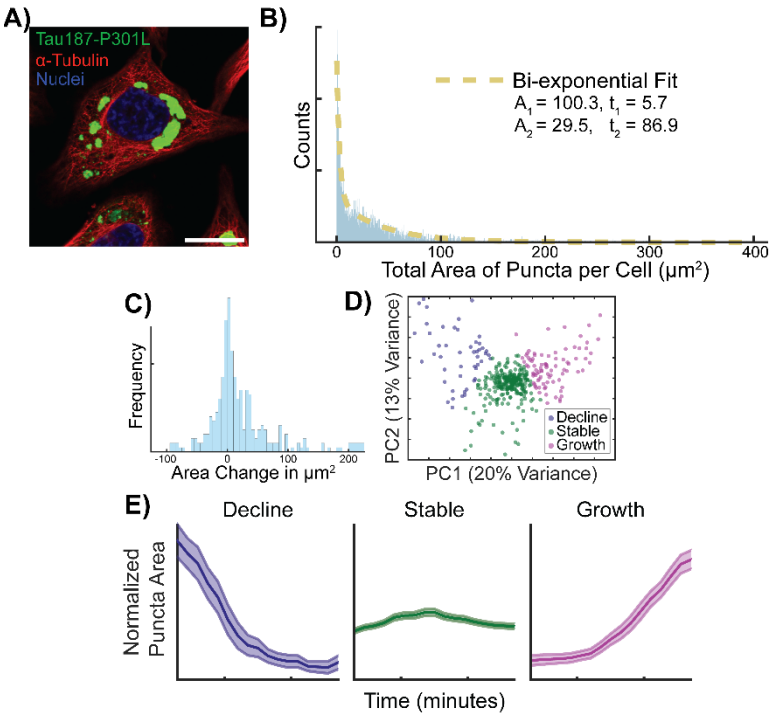

Fig. S8: Growth, Stability, and Shrinkage in Puncta A) Microscopy images showing intact microtubule network (α-tubulin staining in red) despite no visible tau binding to microtubules after large aggregates form (green). Scale bar: 20 μm. B) Histogram of the total puncta area per cell fit to a bi-exponential decay curve (sand dotted line). C) Monitoring cells with jR2R3 P301L-induced aggregates for 10 hours, imaging every 30 minutes. The total change in aggregate area per cell was plotted as a histogram. D) PCA analysis on the trajectories of the changing aggregate area; K-means clustering identified three clusters that represent cells with declining (indigo), stable (green), and growing (purple) aggregate areas. E) The average of the individual trajectories based on K-means clustering are plotted, showing decline, stability, or growth. Shaded area represents the SEM of the normalized, combined trajectories. E) Colocalization of protein degradation markers VCP and P62 (red) with large tau aggregates (green). Scale bar: 20 μm.

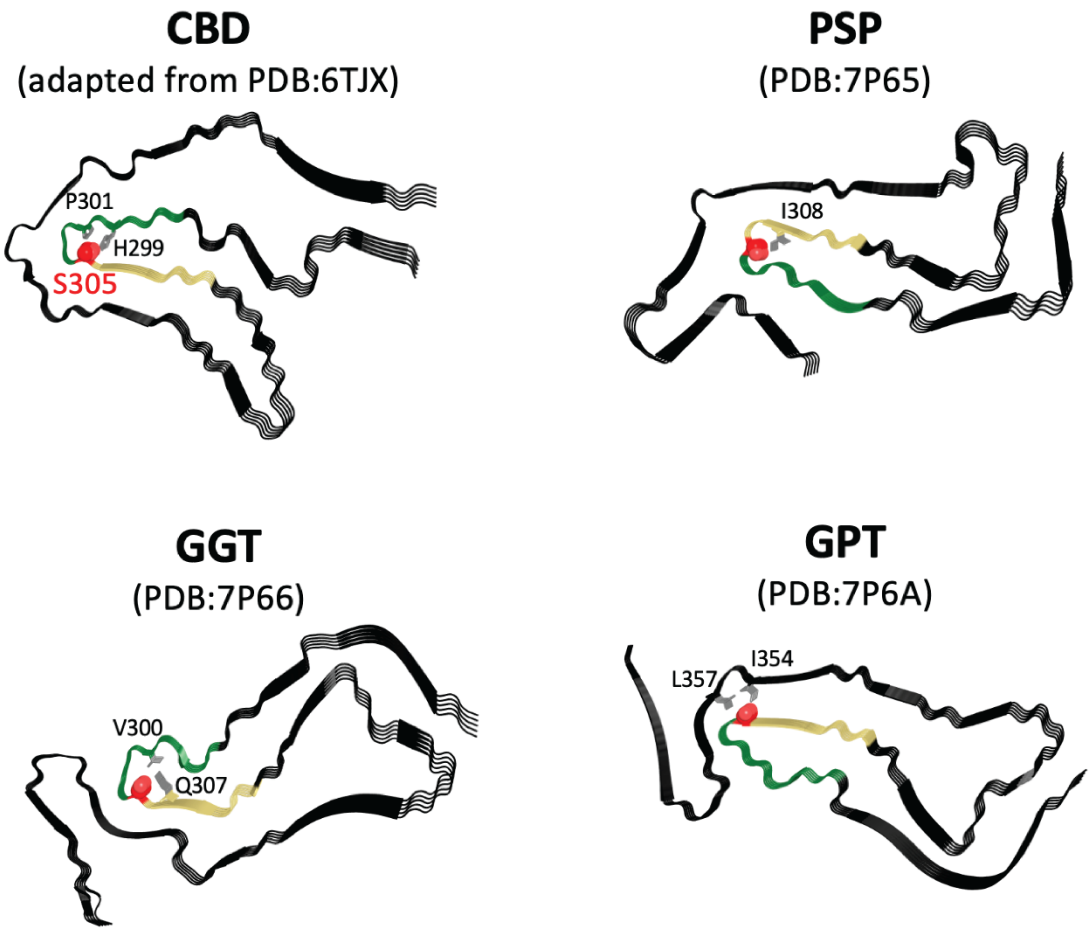

Fig. S9. Orientation of S305 in 4R Tauopathies The relative orientation of S305 (represented as red spheres) in CBD, PSP, GGT, GPT is illustrated. The region of the R2 repeat encompassed by jR2R3 is colored in green and that of R3 is colored in sand. In CBD, PSP, and GGT the S305 is oriented inwards towards the rest of the strand loop strand sequence while in GPT it points outwards and is stabilized by the second layer of the fibril. Residues closely interacting are shown as grey sticks.

SI References

1. N. V. Gorantla, S. Chinnathambi, Autophagic Pathways to Clear the Tau Aggregates in Alzheimer's Disease. *Cell Mol. Neurobiol.* **41**, 1175–1181 (2021).
2. M. A. Rahman, *et al.*, Autophagy Modulation in Aggresome Formation: Emerging Implications and Treatments of Alzheimer's Disease. *Biomedicines* **10**, 1027 (2022).
3. J. M. T. Hyttinen, *et al.*, Clearance of misfolded and aggregated proteins by aggrephagy and implications for aggregation diseases. *Ageing Res. Rev.* **18**, 16–28 (2014).
4. K. Zatloukal, *et al.*, p62 Is a Common Component of Cytoplasmic Inclusions in Protein Aggregation Diseases. *Am. J. Pathol.* **160**, 255–263 (2002).
5. D. C. Carrettiero, *et al.*, Stress routes clients to the proteasome via a BAG2 ubiquitin-independent degradation condensate. *Nat. Commun.* **13**, 3074 (2022).
6. P. Sánchez-Martín, M. Komatsu, p62/SQSTM1 – steering the cell through health and disease. *J. Cell Sci.* **131**, jcs222836 (2018).
7. J. R. Babu, T. Geetha, M. W. Wooten, Sequestosome 1/p62 shuttles polyubiquitinated tau for proteasomal degradation. *J. Neurochem.* **94**, 192–203 (2005).
8. I. Saha, *et al.*, The AAA+ chaperone VCP disaggregates Tau fibrils and generates aggregate seeds in a cellular system. *Nat. Commun.* **14**, 560 (2023).
9. J. Ramesh Babu, *et al.*, Genetic inactivation of p62 leads to accumulation of hyperphosphorylated tau and neurodegeneration. *J. Neurochem.* **106**, 107–120 (2008).
10. F.X. Blaudin de Thé, *et al.*, P62 accumulates through neuroanatomical circuits in response to tauopathy propagation. *Acta Neuropathol. Commun.* **9**, 177 (2021).
11. J. F. Abisambra, *et al.*, Tau Accumulation Activates the Unfolded Protein Response by Impairing Endoplasmic Reticulum-Associated Degradation. *J. Neurosci.* **33**, 9498–9507 (2013).
12. Y. Xu, S. Zhang, H. Zheng, The cargo receptor SQSTM1 ameliorates neurofibrillary tangle pathology and spreading through selective targeting of pathological MAPT (microtubule associated protein tau). *Autophagy* **15**, 583–598 (2019).
13. E. Dimou, *et al.*, Super-resolution imaging unveils the self-replication of tau aggregates upon seeding. *Cell Rep.* **42**, 112725 (2023).
14. C. H. Li, C. K. Lee, Minimum cross entropy thresholding. *Pattern Recognition* **26**, 617–625 (1993).
15. C. A. Glasbey, An Analysis of Histogram-Based Thresholding Algorithms. *CVGIP: Graphical Models and Image Processing* **55**, 532–537 (1993).
16. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **4**, 435–447 (2008).
17. S. Pronk, *et al.*, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
18. X. Daura, W. F. van Gunsteren, A. E., Mark, Folding—unfolding thermodynamics of a β-heptapeptide from equilibrium simulations. *Proteins* **34**, 269-280. (1999).
19. P. Robustelli, S. Piana, D. E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states. *Proceedings of the National Academy of Sciences* **115** (2018).
20. S. Nosé, A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics* **52**, 255–268 (1984).
21. W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **31**, 1695–1697 (1985).
22. U. Essmann, *et al.*, A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577–8593 (1995).
23. L. Verlet, Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **159**, 98–103 (1967).
24. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463–1472 (1997).
25. S. Miyamoto, P. A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **13**, 952–962 (1992).

26. B. C. Dallin, A. S. Kelkar, R. C. V. Lehn, Structural features of interfacial water predict the hydrophobicity of chemically heterogeneous surfaces. *Chemical Science* **14**, 1308–1319 (2023).
27. J. I. Monroe, M. S. Shell, Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *J. Chem. Phys.* **151**, 094501 (2019).