Amyloid-motif-dependent tau self-assembly is modulated by isoform sequence context

Sofia Bali[1,2], Ruhar Singh[2], Pawel M. Wydorski[1,2], Nico E. Van Nuland[3], Aleksandra Wosztyl[1,4], Valerie A. Perez[2], Dailu Chen[2], Jie Chen[2], Josep Rizo[4,5,6], Lukasz A. Joachimiak[2,5]

[1]Molecular Biophysics Graduate Program, University of Texas Southwestern Medical Center, Dallas, TX 75390, United States

[2]Center for Alzheimer's and Neurodegenerative Diseases, Peter O'Donnell Jr. Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, United States

[3]Instituto de Química Física Blas Cabrera (IQF), CSIC, E-28006 Madrid, Spain.

[4]Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, United States

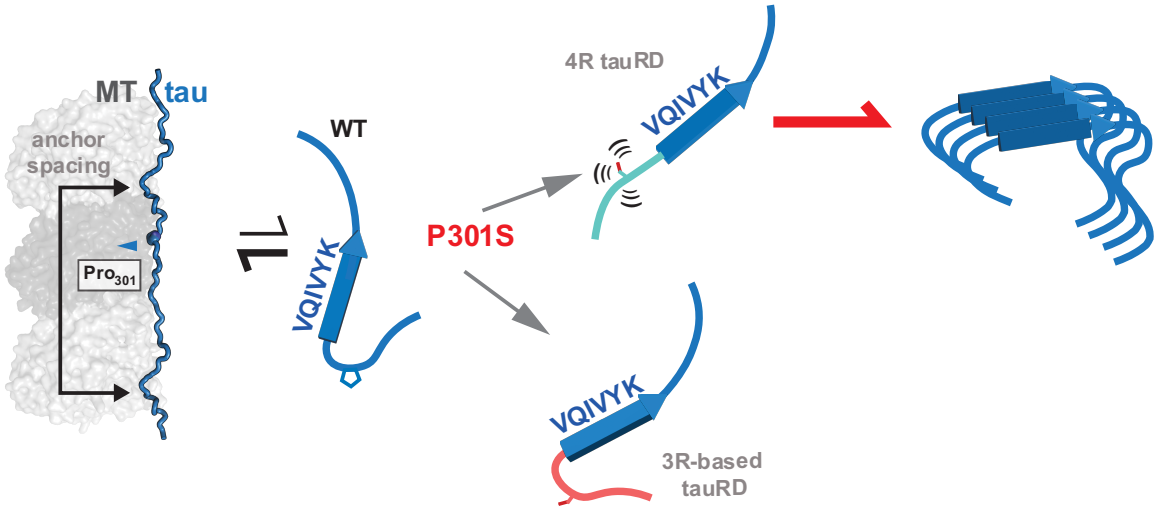
Southwestern Medical Center, Dallas, TX 75390, United States

[5]Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390, United States

[6]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75390, United States

Abstract

The microtubule-associated protein tau is implicated in neurodegenerative diseases characterized by amyloid formation. Mutations associated with frontotemporal dementia increase tau aggregation propensity and disrupt its endogenous microtubule-binding activity. However, the structural relationship between aggregation propensity and biological activity remains unclear. We employed a multi-disciplinary approach, including computational modeling, NMR, cross-linking mass spectrometry, and cell models to engineer tau sequences that modulate its structural ensemble. Our findings show that substitutions near the conserved 'PGGG' β-turn motif informed by tau isoform context reduce tau aggregation in vitro and cells and can even counteract aggregation induced by turn destabilizing disease-associated proline-to-serine mutations. Engineered tau sequences maintain microtubule binding and explain why 3R isoforms exhibit reduced pathogenesis compared to 4R. We propose a simple mechanism to reduce the formation of pathogenic species while preserving biological function, thus offering insights for therapeutic strategies aimed at reducing tau protein misfolding in neurodegenerative diseases.



INTRODUCTION

The deposition of the microtubule-associated protein tau is linked to a group of neurodegenerative diseases collectively known as tauopathies[1-3]. Recent cryo-EM tau fibril structures derived from patient material have uncovered how tau adopts distinct conformations (i.e., structural polymorphs) across different tauopathies[4-6], each able to propagate in a prion-like manner[7,8]. Despite these insights into the fate of tau protein, the mechanism by which it misfolds into different conformations in each disease is poorly understood. Prior studies have shown that the normally aggregation-resistant tau monomer can convert into an aggregation-prone conformation, termed a seed, which can then recruit inert tau monomers to form amyloid fibrils[9]. These seeds have been detected in young tauopathy mouse models before the detection of insoluble species[10] and can be isolated from Alzheimer's disease patient material[11]. Structural and biochemical studies comparing inert tau monomers and monomeric seeds have identified rearrangements of long-range charge complementary and short-range hydrophobic interactions mediated by amyloidogenic motifs[12].

Mutations in the *MAPT* gene that encodes tau are linked to frontotemporal dementia (FTD)[13,14], and these mutations primarily localize to the repeat domain of tau (tauRD)[15,16]. This domain comprises the region required for binding to microtubules (MT)[17]. Alternative splicing of the *MAPT* gene results in two isoforms of the tau repeat domain: 4R tauRD (243-380) and 3R tauRD (243-380, ΔR2:275-305)[18]. Notably, 4R tau isoforms aggregate with a higher propensity than 3R tau[19]. Tau amyloid deposits in tauopathies are mostly composed of either a mixture of 4R and 3R isoforms or 4R-only isoforms, with Pick's disease being the only tauopathy exclusively linked to the 3R tau isoform[1,20]. Experimental evidence from mouse models provides compelling insights into this isoform dynamics: reducing 4R tau isoforms reduces tau pathology and toxicity, whereas reducing 3R isoforms increases tau pathology[19]. This differential behavior of 3R and 4R tau isoforms presents avenues for therapeutic intervention, with current strategies exploring the use of antisense oligonucleotides[20,21]. A mechanistic understanding of how these isoforms alter tau's intrinsic aggregation properties will be crucial in developing novel regulatory strategies.

A serine or leucine mutation at proline position 301, which localizes to the ^{301}PGGG^{304} motif at the 3R/4R splice junction site, causes FTD[16] and is commonly used in animal, cellular, and in vitro models to study tau aggregation[22-26]. This mutation is also upstream of the ^{306}VQIVYK^{311} amyloid motif, which forms essential interactions in all tau fibril structures[27,28], while mutations in this motif abrogate tau aggregation[11]. Moreover, time-resolved cryo-EM analysis of disease-specific tau folds has uncovered early fibrillar intermediates stabilized by VQIVYK-VQVIYK interactions[29]. Our previous studies using minimal fragments have established that sequences from both 4R and 3R tau isoforms upstream of the ^{306}VQIVYK^{311} motif are sufficient to regulate amyloid motif aggregation. We observed that these sequences, which are upstream of the 'PGGG,' can transiently engage with ^{306}VQIVYK^{311,} and their interaction depends on the 'PGGG' sequence adopting a β-turn[24,30,31]. However, the P301S mutation from 4R tau can disrupt this aggregation inhibition, while the equivalent P270S mutation from 3R tau does not. The flexible nature of the 'PGGG' sequence enables the upstream sequence to adopt a range of conformations in tauopathy fibrils from disordered to folding onto the amyloid motif into different beta-arch geometries. Post-translational modifications in proximity to the amyloid motif have been shown to stabilize amyloid motif-mediated self-association into fibrils[32]. These findings underscore the complexity of the

structural ensembles sampled by 4R vs. 3R tau isoforms that may balance protective anti-aggregation and pro-aggregation interactions involving amyloid motifs[32-35].

Thus, a deeper investigation into how to regulate ^{306}VQIVYK^{311}-based tau aggregation will yield new insights into tuning tau ensembles to not only inhibit assembly but also to control tau folding into defined amyloid conformations. This study uses a multidisciplinary approach to understand how local sequence changes modulate the 'PGGG' structural motif to limit amyloid motif-dependent assembly. By leveraging the natural sequence composition of the 3R isoform, we evaluate the structural and chemical impacts that reduce aggregation while maintaining microtubule-stabilizing function. Our research presents a model for inhibiting 4R tau aggregation, potentially opening new avenues to develop targeted therapeutic strategies for tauopathies.

RESULTS

Differential aggregation propensity induced by sequence preceding the amyloid motif ^{306}VQIVYK^{311} in 3R vs. 4R tau.

The 4R and 3R tau isoforms differ in sequence upstream of the ^{306}VQIVYK^{311} motif at only five out of 11 amino acids (Fig. 1a). We compared ThT-based aggregation propensities of wild-type and FTD-linked proline to serine mutants in 4R and 3R peptides that span the inter-repeat region to repeat 3 (Fig. 1b-c and Supplementary Fig. 1a-d). We determined that the P301S R2R3$_{295\text{-}311}$ peptide assembled into amyloid fibrils rapidly starting at 100 μM; peptide WT R2R3$_{295\text{-}311}$ only aggregated at the highest concentration, 800 μM, while peptides WT R1R3$_{295\text{-}311\Delta R2}$ and P270S R1R3$_{263\text{-}311\Delta R2}$ did not aggregate at any concentration within the time frame of the experiment (Fig. 1c and Supplementary Fig. 1a-d). We evaluated the aggregation behavior of WT and P301S 4R tauRD (residues 244-380), as well as WT and P270S 3R tauRD (residues 244-380Δ275-305) (Supplementary Fig. 1e). These longer tau fragments replicated the differential proline to serine aggregation phenotypes observed in the P270S R1R3$_{263\text{-}311\Delta R2}$ and P301S R2R3$_{295\text{-}311}$ peptides.

We acquired TOCSY and NOESY spectra for the four peptides at 25ºC and 10ºC to investigate the local structural differences sampled by these four sequences (Fig 1b.). We performed proton chemical shift assignments for each sequence at 10Cº and transferred assignments to 25ºC (Supplementary Fig. 2a-d). For the aggregation-prone peptide, P301S R2R3$_{295\text{-}311}$, we confirmed that signal intensities did not significantly change throughout the experiment at 25ºC (Supplementary Fig. 1f). We observed small deviations when comparing the chemical shifts of WT R1R3$_{263\text{-}311\Delta R2}$ to P270S R1R3$_{263\text{-}311\Delta R2}$ and WT R2R3$_{295\text{-}311}$ to P301S R2R3$_{295\text{-}311}$ except for the expected shifts at the position preceding the proline to serine mutation (Q269 and V300 respectively, Supplementary Fig. 2a,b and Supplementary Table 1). We estimated the extent of b-sheet or α-helical content in our fragments based on Cα proton chemical shifts (Hα) using the neighbor-corrected secondary structure propensity calculator (ncSPC) with a neighbor correction window of 3[36,37].

All four peptides exhibited a similar secondary structure profile at 25ºC and 10ºC with more negative ncSP values at 25ºC (Fig. 1d-g; Supplementary Fig. 1g-j). The profiles showed negative scores consistent with an enhanced population of β-sheet structure compared to random coil for the left and right side of the 'PGGG'/'SGGG' motif (Fig. 1d-g). The central 'PGGG' motif (270-273) in WT R1R3$_{263\text{-}311\Delta R2}$ had a higher turn propensity than the motif (301-304) in WT R2R3$_{295\text{-}}$

$_{311}$. In the peptides containing the mutant 'SGGG' motif, P270S R1R3$_{263\text{-}311\Delta R2}$ had a similar turn propensity as its proline-containing counterpart. In contrast, the P301S R2R3$_{295\text{-}311}$ peptide showed no significant deviations from the random coil signal (Fig. 1d-f). Therefore, we observed subtle differences localized to the 'PGGG'/'SGGG' motif specific to the P301S R2R3$_{295\text{-}311}$ peptide.

A simple hypothesis is that the R1R3 sequence is less compatible with amyloid fibrils than the R2R3 sequence. To test this, we assessed the role of these sequences in the end-stage tau fibrils derived from tauopathy patients (Supplementary Fig. 1k). We observe that these sequences adopt a large variety of conformations within the fibril form: dynamic and unresolved in the fibril structure (AD/CTE), forming a beta arch that interacts on either the polar ("-Q-V-K" GGT/CBD/AGD,) or nonpolar ("V-I-Y" GPT/PSP,P301Sm) face of the amyloid motif, and only in the 3R derived fibril is this segment extended (PiD; Supplementary Fig. 1k). Comparing this segment to tau P301S fibrils extracted from a tauopathy mouse model, we observe that the mutant segment adopts yet a different beta-arch conformation that folds onto the "V-I-Y" face (Supplementary Fig. 1k). Overlaying these regions, we observe that the 'PGGG/SGGG' forms a pivot point where the flexibility of this motif enables the access of interactions between both faces of the 'VQIVYK' motif with different residues in the 294-305 element, with both the Proline and Serine at position 301. To evaluate the individual contributions of each residue, we compared the solvent-accessible surface area (SASA) to the estimated contribution of each residue to the fibril structure through an alanine[27] (Supplementary Fig. 1m-n). As previously described[27], we observe that the VQIVYK motif is buried and is important to the stability of the assembly (Supplementary Fig. 1m-n). In the AD and CTE tau folds, the sequence starts at the amyloid motif, and the preceding sequences are not ordered (Supplementary Fig. 1k). In the other structures, the upstream residues from R2 (D295, I297, V300, and S305) and R1 in the PiD structure (E264, L266, Q269, and K274) are less buried than the VQIVYK amyloid motif (Supplementary Fig. 1m). Still, their contribution to stability is variable (Supplementary Fig. 1n). Of all the residues upstream of VQIVYK, I297 in R2 or L266 in R1 appear to have the largest and most consistent contribution to the local stability of tau fibrils with the leucine and isoleucine residues largely isosteric (Supplementary Fig. 1k, m-n) except for the P301S tau structure where the Cγ of I297 forms a clash. Similarly, D295 in R2 and E265 in R1 are either buried and form salt bridges or face the solvent (Supplementary Fig. 1k, m-n). Finally, the residues that flank the PGGG motif are buried and may be more sensitive to mutations of valine to glutamine at 300 and serine to lysine at 305 from R2 to R1 (Supplementary Fig. 1k, m-n). From these analyses, we predict that for the CBD and AGD tauopathy conformations, the 3R sequences would be incompatible with the beta-arch conformations. At the same time, they should be compatible with AD, CTE, PiD, and likely also PSP, GGT, and GPT. This partitioning of sequence compatibility follows disease classification based on tau isoform pathology that is mixed 3R/4R for AD and CTE, 3R only for PiD, and 4R only for the others. Thus, the sequence upstream of VQIVYK in 3R tau (i.e., R1R3) may not be able to adopt specific beta-arch conformations, yet it is largely compatible with the available tau fibril structures. Therefore, understanding how the changes in local dynamics of this region in the monomeric state may uncover the role of the pre-fibril state of this region in defining how to regulate tau folding into amyloids.

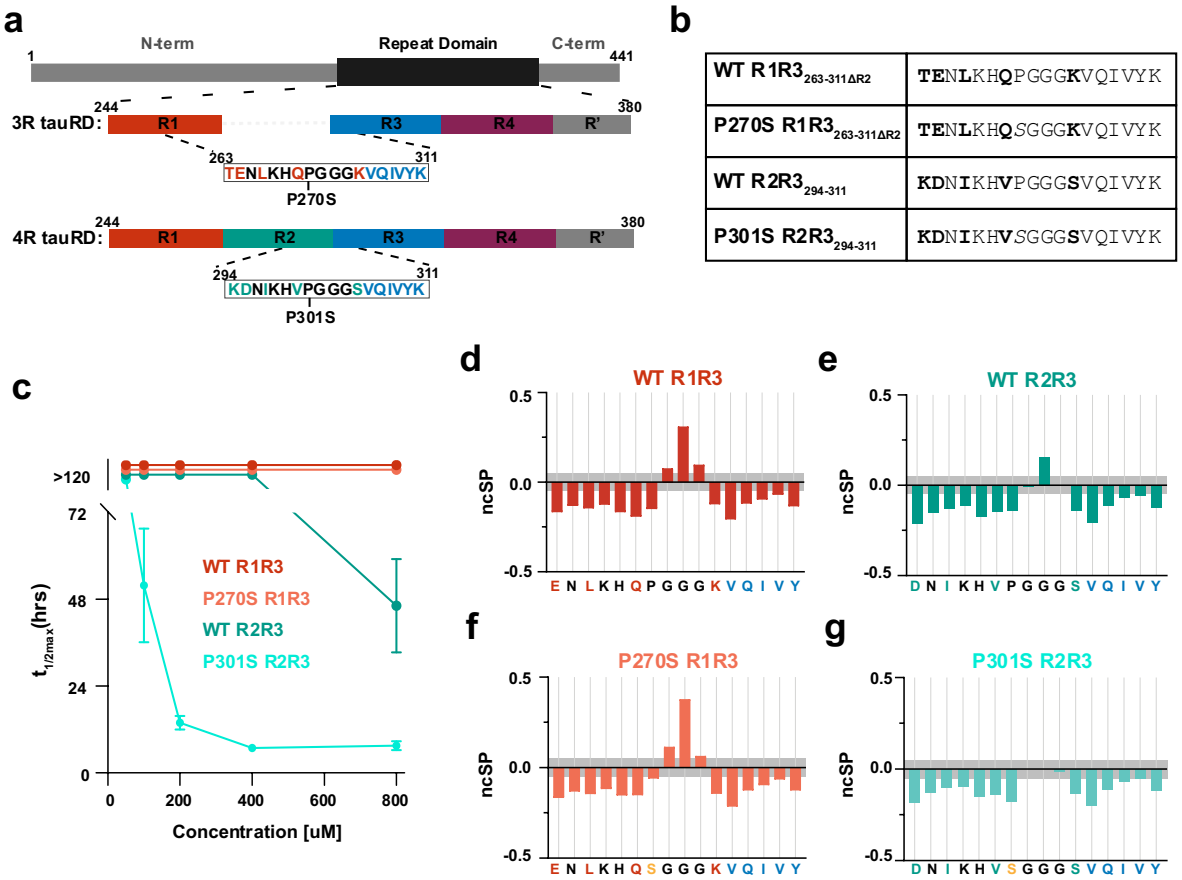


Figure 1. Proline to serine mutations in a 4R-like fragment leads to a more unstructured turn. a. Cartoon schematic of 3R and 4R-like tau fragments. Repeat domains are colored red, green, blue, and purple of repeats 1, 2, 3, and 4. Inset shows WT minimal fragments that span repeats 1 and 3 and repeats 2 and 3. The site of the proline to serine mutation is indicated below the sequence. **b**. Table listing the four tested peptide sequences. **c**. ThT fluorescence aggregation assay of WT R1R3$_{263\text{-}311\Delta R2}$ (red), P270S R1R3$_{263\text{-}311\Delta R2}$ (orange), WT R2R3$_{295\text{-}311}$ (green), and P301S R2R3$_{295\text{-}311}$ (cyan) peptides carried out at 5 concentrations: 50, 100, 200, 400, and 800 μM. Data is shown as the $t_{1/2max}$ fit to each aggregation curve. The data is plotted as an average with a standard deviation. **d-g**. neighbor corrected structure propensity (ncSP) calculated from peptide 25˚C Hα chemical shift assignments using ncSPC calculator for WT R1R3$_{263\text{-}311\Delta R2}$ (d, red), P270S R1R3$_{263\text{-}311\Delta R2}$ (e, green), WT R2R3$_{295\text{-}311}$ (f, orange), and P301S R2R3$_{295\text{-}311}$ (f, cyan). Negative/positive values correspond to β-sheet/α-helical propensity. Gray bars highlight the deviations expected for random coil structures.

158

159 **Molecular Dynamics ensembles differentiate the 'SGGG' structure of the aggregation-prone**
160 **sequence.**

161 We employed Molecular Dynamics (MD) simulations of these peptide model systems to assess
162 differences in intramolecular interactions and accessible states. The MD simulations for the four
163 peptides (Fig 1b.) were performed using a two-stage production scheme followed by analysis with
164 a Markov State Model (Supplementary Note 1 and Supplementary Fig. 3). Simulations were
165 generated using canonical sampling schemes with the Amber-991lsb ff which produces ensembles
166 suitable for disordered sequences, but with a bias towards increased β-sheet propensity[38,39] (Fig.
167 2a). From the resultant 15μs of final production simulations we calculated the distance between
168 the backbones of the C-terminal amyloid motif 'VQIVYK' and the N-terminal residues

169 'TENLKH' or 'KDNIKH' and observed wider distributions for the proline-to-serine peptides with
170 higher sampling of states with backbone distances between 0.5-1.0 nm for P301S R2R3$_{295\text{-}311}$ (Fig.
171 2b). Furthermore, we assessed the relative flexibility for each residue across the simulation using
172 the root mean squared fluctuation (RMSF). We observed that the four peptides had relatively
173 similar fluctuation, with only a significant increase in flexibility for residues in the 'SGGG' motif
174 of P301S R2R3$_{295\text{-}311}$.

175 We built and validated a Markov State Model (MSM) with pyEMMA[40] (see Supplementary Note
176 1 and Supplementary Fig. 3). From the MSM, we visualized a free energy surface and mapped the
177 6 metastable states of the landscape (Fig. 2d-e). For each state, we calculated the relative frequency
178 for each peptide and generated heatmaps for the mean distance (rainbow) and deviation (grayscale)
179 using the 4R isoform numbering (Fig. 2f-g, states 1-6). We observed that the lowest energy state
180 (Fig. 2e-g, state 1, yellow) was populated only by the 'PGGG' turn-containing peptides (WT
181 R1R3$_{295\text{-}311\Delta R2}$ and WT R2R3$_{295\text{-}311}$) and is composed of defined β-hairpin structures with contacts
182 above the diagonal between residues 295-299/305-309 (Fig. 2g, state 1). The next lowest energy
183 state (Fig. 2e-g, state 2, green) was sampled by all four peptides and represents a collection of local
184 minima with structures forming a less defined turn and a shifted register with contacts below the
185 diagonal (Fig. 2g, state 2). P270S R1R3$_{295\text{-}311\Delta R2}$ primarily populated state 3 and had contacts
186 between 297-299/305-309 (Fig. 2e-g, state 3, turquoise). State 4 (Fig. 2e-g, state 4, teal) was
187 primarily populated by WT R1R3$_{295\text{-}311\Delta R2}$ and WT R2R3$_{295\text{-}311}$ and represents an interchange from
188 the local minima state 1 and the rapidly exchanging state 2 with interactions below the diagonal
189 (Fig. 2g, state 4). Higher energy state 5 (Fig. 2e-g, state 5, indigo) is also accessible to all four
190 peptides but is preferentially sampled by the P301S R2R3$_{295\text{-}311}$ sequence representing a more
191 expanded structure with medium range interactions below the diagonal (Fig. 2g, state 5). By
192 contrast, state 6 was uniquely populated by the P301S R2R3$_{295\text{-}311}$ defined by transient interactions
193 above the diagonal, leaving reduced contacts to ^{306}VQIVYK^{311} (Fig. 2e-g, state 6, purple).

194 These observations reveal a unique sampling of states across the four peptides. Sequences with
195 P301 at the '**PG**GG' motif promote sampling of a β-hairpin-like structure with transient
196 interactions across the peptide sequence. The proline to serine mutation disrupts the turn
197 localization. However, the sequence properties of 'TENLKHQ-K' in P270S R1R3$_{295\text{-}311\Delta R2}$
198 increase the sampling of states with a left-shifted turn 'SG**GG**' that still contains medium-distance
199 contacts to VQIVYK. In the P301S R2R3$_{295\text{-}311}$ peptide, the 'DNIKHV-S' sequence shifts
200 sampling to states with a turn centered before the 'SGGG' motif, and the induction of a
201 hydrophobic residue at position 300 increases the sampling of interactions to residues 305-307
202 resulting in reduced interactions to the latter part of the VQ**IVYK** motif.

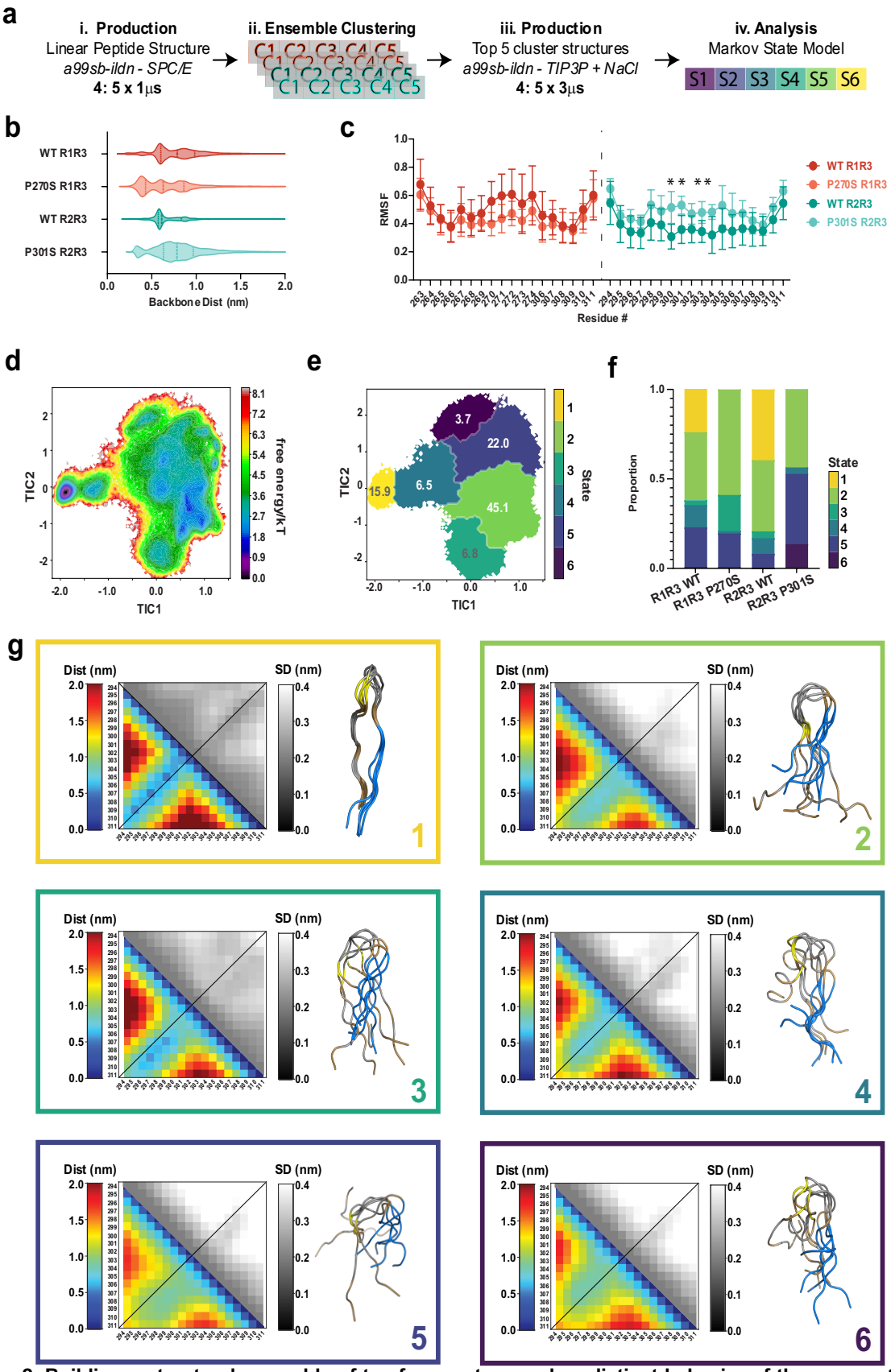


Figure 2. Building a structural ensemble of tau fragments reveals a distinct behavior of the aggregation-prone 4R-derived peptide.

a. Two stage MD simulation workflow for: WT R1R3$_{263\text{-}311\Delta R2}$, P270S R1R3$_{263\text{-}311\Delta R2}$, WT R2R3$_{295\text{-}311}$, and P301S R2R3$_{295\text{-}311}$ **b.** backbone distance between the center of mass of the N-term residues 1-6 (TENLKH/KDNIKH) and C-term amyloid motif residues (VQIVYK) second-stage MD ensembles. The data is shown as violin plots with median and 25/75% quartiles labeled in dash lines. **c.** Per-residue RMSF of 3R-derived (left) and 4R-derived (right) peptides from second-stage MD ensembles. The data is shown as average +/-SD across the 5 repeat simulations. **d.** Free Energy landscape mapped onto the first two TICA components calculated from MD ensemble MSM for the combined simulations of WT R1R3$_{263\text{-}311\Delta R2}$, WT R2R3$_{295\text{-}311}$, P270S R1R3$_{263\text{-}311\Delta R2}$, and P301S R2R3$_{295\text{-}311}$ for a total of 60µs. Data is colored based on normalized free energy from black to red. **e.** MSM metastable state map, states 1-6 mapped from yellow to purple on the free energy landscape. **f.** Bar plot illustrating the proportion of sampling of each state per peptide. The states sampled for each peptide are colored in c. **g**. State-specific Average(bottom) and Standard deviation (top) maps of pairwise distance calculated for all timepoints in each state in boxes colored in c. Average (bottom) colored from blue to red, 0 nm to 2.0 nm. Standard deviation (top) colored from black to white, 0 nm to 0.4 nm. The first 5 structures per state are shown in cartoon representation with 'VQIVYK' colored in blue, unique residues colored in brows, Proline/Serine position colored in yellow, and similar residues colored in gray.

Ensemble-based design of 4R tauRD mutants reduces tau aggregation propensity.

To design aggregation-resistant 4R tauRD, we based our strategy on the sequence differences from the 3R isoform and incorporated them in a 4R context. The first construct uses the sequence from R1 with V300Q and S305K mutations (herein, tauRD_QK). The second mutant incorporates the full R1 element with four mutations: D295E, I297L, V300Q, and S305K (herein, tauRD_EL_QK). In addition to testing the constructs with the wild-type proline (i.e., 'PGGG' motif), we tested our designs in the context of the P301S mutation (i.e., 'SGGG' motif) (Fig. 3a). We evaluated the aggregation propensity of the constructs in a ThT fluorescence aggregation assay induced with heparin. We fit the resultant curves to estimate a $t_{1/2max}$ (Supplementary Fig. 4a-f). The presence of fibrils at the endpoint of the aggregation reactions for the 10 µM and 40 µM reactions was confirmed by TEM (Supplementary Fig. 4g, h). We observed

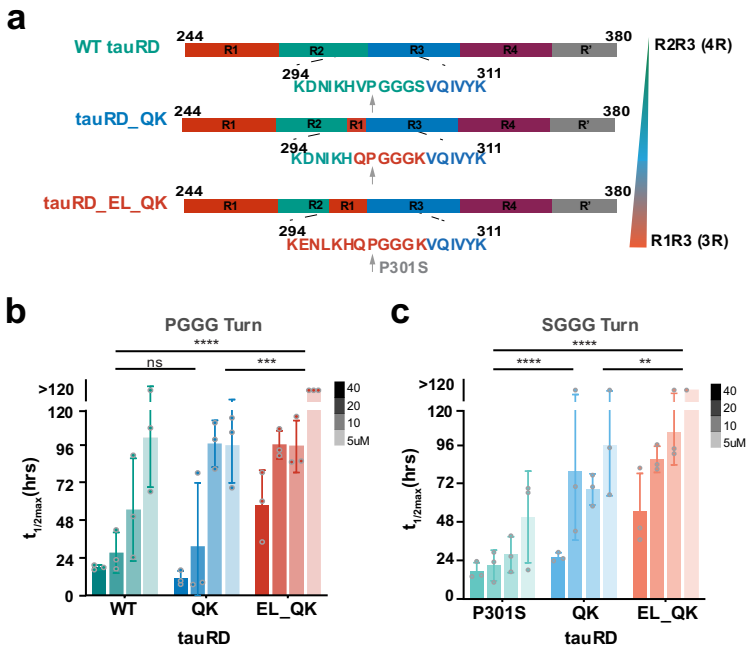


Figure 3. Designed tau sequences reduce aggregation of tau in vitro. a. Cartoon schematic of 4R tauRD and designed tauRD_QK and tauRD_EL_QK. Repeat domains are colored as in Figure 1, and the sequence inset of designed mutations below each schematic with residues is colored based on repeat. **b-c.** ThT fluorescence aggregation assay of tauRD constructs induced with 2X heparin carried out at 4 concentrations: 5, 10, 20, 40 µM (light to dark) for 'PGGG' containing constructs WT tauRD (green), tauRD_QK (blue), and tauRD_EL_QK (red) in b, and 'SGGG' containing constructs P301S tauRD (cyan), P301S tauRD_QK (sky blue), and P301S tauRD_EL_QK (orange) in c. Data is shown as the $t_{1/2max}$ fit to each aggregation curve. The data is plotted as an average with a standard deviation, significance calculated from column effects using 2way ANOVA (*=p < 0.05, **=p < 0.01, ***=p < 0.001, ****=p < 0.0001).

237 that the tauRD_QK mutant aggregated with similar rates to WT tauRD (Fig. 3b and Supplementary
238 Fig. 4a, b). By contrast, the tauRD_EL_QK showed values that were 2- and 3-fold slower than
239 WT tauRD (Fig. 3b and Supplementary Fig. 4c). The P301S tauRD_QK aggregated between 2-
240 and 4-fold more slowly compared to P301S tauRD while the P301S tauRD_EL_QK further
241 reduced aggregation leading to $t_{1/2max}$ values between 3- and 5-fold slower than P301S tauRD (Fig.
242 3c and Supplementary Fig. 4d-f).

243 Additionally, we evaluated our selected mutations in the context of 3R tauRD (244-380Δ275-305)
244 using the reverse mutations that include E264D, L266I, Q269V, and K274S (herein, 3R
245 tauRD_DI_VS) in constructs encompassing the full sequence observed in the ex-vivo Pick's
246 disease fibril structure[41]. ThT fluorescence intensity endpoints were used to evaluate the heparin-
247 induced aggregation propensity of the constructs, and the fibrils were visualized by TEM
248 (Supplementary Fig. 4i-k). 3R WT tauRD and 3R tauRD_DI_VS showed baseline ThT signal and
249 no fibrils by TEM. By contrast, the 3R P270S tauRD had a reduced ThT signal with minimal fibrils
250 observed by TEM relative to the 4R P301S tauRD construct (Supplementary Fig. 4j). However,
251 3R P270S tauRD_DI_VS increases fibril formation and produces ThT signal at about 70% of the
252 4R tauRD P301S (Supplementary Fig. 4j). Therefore, we found we could increase the aggregation
253 sensitivity of the 3R tauRD by incorporating the sequence from repeat 2 but in the context of a
254 destabilized 'SGGG' turn. Taken together, we observe that introducing the 3R-like mutations into
255 4R tauRD near the VQVIYK amyloid motif reduces the aggregation propensity of tau.

256 **Balance of intermediate and long-range contacts correlates with inhibition of tau**
257 **aggregation.**

258 To understand how the highly dynamic intramolecular interactions in tauRD change in response
259 to the designed mutations, we turned to XL-MS. This approach enables the capture of interactions
260 in complexes to define the architectures of assemblies[42-44]. More recently, our lab has established
261 ways to probe changes in the conformation of IDPs using XL-MS[12,30,45]. Previously, we have used
262 this approach to probe the effect of P301S on tauRD dynamics, uncovering how this mutant is
263 more sensitive to unfolding and, thus, more competent to self-associate[30]. Here, we extend this
264 approach to understand how the designed mutations change the distribution of local contacts in
265 proximity to the mutation and longer-range contacts to relate these changes to the reduction of tau
266 aggregation. The 6 purified tauRD constructs (Fig. 4a) were reacted with the lysine-reactive DSS
267 cross-linker (DSS, spacer arm 11.4A). The samples were analyzed by mass spectrometry, and the
268 data was processed using the xQuest pipeline[46]. Each sample was processed as 5 independent
269 replicates, and only consensus contacts were considered for further analysis (Fig. 4a).

270 To look at the relative accessibility of regions in tauRD, we determined the normalized frequency
271 of mono-link modifications, defined as DSS modifications on a single lysine, mapped onto each
272 repeat. WT and P301S tauRD show significantly higher frequencies of monolinks in repeat 3
273 despite a similar frequency of modification across the rest of the sequence (Fig. 4b and
274 Supplementary Fig. 5a). Next, we mapped the loop-links/cross-links (XL) frequency in these
275 samples, defined as intramolecular covalent contacts between pairs of lysines located on one or
276 two peptides, showing unique cross-link pair colored by the normalized frequency (Supplementary
277 Fig. 5b). We grouped and clustered these XL by contact order: contacts within repeats (local; ≤ 10
278 residues), between adjacent repeats (intermediate; 11-39 residues), and further (long-range; ≥ 40
279 residues) (Supplementary Fig. 5c). First, a comparison of the WT and P301S tauRD analysis

280 reaffirms our previous findings that the P301S mutation appears to destabilize intermediate and
281 long-range contacts within the repeat domain of tau[30], while the relative frequency of local XL
282 increases (Fig. 4c). For our four designed mutants, we observe an increase in intermediate contacts
283 coupled with a decrease in long-range XL and a decrease in local XL. (Fig. 4c and Supplementary
284 Fig. 5c).

285 Mapping these cross-links by repeat domain (i.e., R1-R4), we observe that the shift from long-
286 range to intermediate interactions corresponds to a shift from N-terminal interactions in Repeat 1-
287 2 in WT and P301S tauRD to more central interactions in repeats 2-4 for the designed constructs
288 (Fig. 4d). Finally, we relate the abundance of contacts to K311 within the VQIVYK amyloid motif
289 and find that the designed constructs have more abundance of unique interactions at this position.
290 We see interactions in our region of interest through XL at lysine positions K280, K281, and K311,
291 which are not identified in either WT or P301S tauRD (Supplementary Fig. 5d).

292 To quantify residue level changes in turn dynamics proximal to the amyloid motif in WT tauRD,
293 P301S tauRD, tauRD_EL_QK, and P301S tauRD_EL_QK, we employed ^{15}N-^{1}H heteronuclear
294 NOE (hetNOE) relaxation experiments[47]. Assignments were transferred from published spectra
295 (bmrb 19253)[48] and supplemented with ^{15}N-NOESY-HSQC and ^{15}N-TOCSY-HSQC experiments
296 to assign the mutant sites D295E, I297L, and S305K, assignment for V300Q not made
297 (Supplementary Fig. 5e and Source Data 6). Overall, we observe high levels of dynamics across
298 the entirety of each sequence with no significant difference in rigidity for the 295-311 region
299 (Supplementary 6a-e). Interestingly, we observe lower signal-to-noise for P301S tauRD and P301S
300 tauRD_EL_QK and no signals above the noise for the G302 position for P301S tauRD
301 (Supplementary Figure 6d). Higher concentration experiments for P301S tauRD and P301S
302 tauRD_EL_QK improve signal-to-noise for this region (Supplementary Fig. 6f), but peak
303 intensities above the noise for position 302 in the saturation spectra of P301S tauRD were not
304 recovered consistent with higher dynamics (Fig. 4e). The consistency in dynamics across the
305 tauRD constructs suggests the differences that confer distinct aggregation behavior do not alter the
306 protein dynamics globally suggesting similar ensemble diversity across the tauRD constructs.

307 Across the designed constructs, the most significant changes localize to intermediate-range
308 interaction between repeats 2, 3, and 4. Overall, we find that a balance of long and intermediate
309 contacts is correlated with aggregation inhibition, and the loss of these combined with the
310 enhancement of localized 'SGGG' turn dynamics increases the unfolding and exposure of the
311 amyloid motif.

312

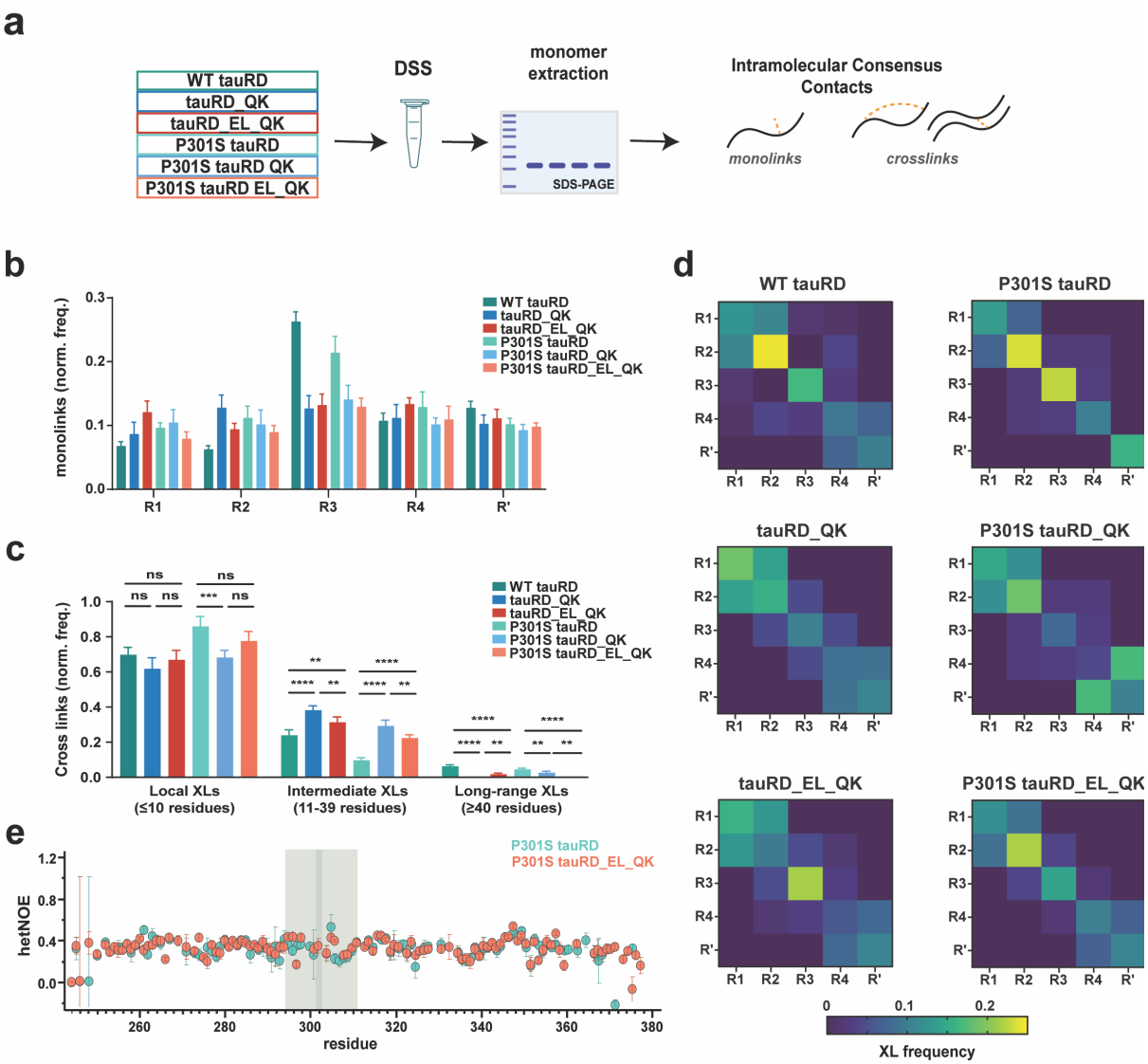


Figure 4. 3R-based mutations increase intermediate-range interactions. A. Schematic of sample preparation for XL-MS workflow. Tau constructs: WT tauRD (green), tauRD_QK (blue), and tauRD_EL_QK (red), P301S tauRD, P301S tauRD (cyan), P301S tauRD_QK (sky blue), and P301S tauRD_EL_QK (orange) were cross-linked with disuccinimidyl suberate (DSS) then ran on an SDS-PAGE gel for monomer extraction, digestion, and purification to be then processed through our XLMS workflow for identification of consensus mono-links, and loop-link/cross-links. **b.** Quantification of identified mono-links grouped by repeat: R1 (244-274), R2 (275-305), R3 (306-336), R4 (337-368), and R' (369-380); for tauRD constructs (colored as in a). Data is shown as the sum of the normalized mean and standard deviation across 5 replicated of the nseen for each construct. **c**. Quantification of local (< 10 residues), intermediate (11-39 residues), and long-range loop-links/cross-links (> 40 residues) by sample (colored as in a). Data is plotted as an average with a standard deviation, significance calculated one way ANOVA (*=p < 0.05, **=p < 0.01, ***=p < 0.001, ****=p < 0.0001). **d**. Heatmap of distribution of loop-links/cross-links for each construct grouped by repeat as in c. Data is shown as the average normalized frequency of nseen for each construct colored from navy to yellow. **e.** ^{15}N-^{1}H heteronuclear NOE experiments for P301S tauRD (cyan) and P301S tauRD_EL_QK (orange). The ratio is calculated by comparing the integrated peak volume between the saturation and reference spectra (see Supplementary Fig. 6f). Peak errors were determined from the standard deviation of the intensity distribution in signal-less spectral regions. 295-311 region marked by the green bar, G302 residue marked with a gray bar.

313

314 **Aggregation-reducing tau mutations maintain microtubule-binding activity.**

315 The endogenous function of tau consists of stabilizing microtubules (MT)[17]. We performed a
316 fluorescent MT polymerization assay coupled with molecular modeling of the MT: tau complex
317 to determine how our engineered tau constructs alter this function. Purified tubulin dimer was
318 incubated with each of the 6 tau constructs (i.e. Fig 3a). The rate and extent of polymerization
319 relative to tubulin alone were determined by fitting the resultant curves (Supplementary Fig. 7a,
320 b). For the 'PGGG' turn constructs, we observe that tauRD_EL_QK had increased polymerization
321 by 1.5-fold relative to WT. In contrast, tauRD_QK polymerization amplitude was significantly
322 reduced by 1.5-fold (Fig. 5a and Supplementary Fig. 7a). However, all three constructs had no
323 significant difference in the $t_{1/2max}$ time to endpoint polymerization (Fig. 5b and Supplementary
324 Fig. 7a). Mutation of the proline to serine at position 301 has previously been shown to alter
325 binding to MT[49,50]. Consistent with this observation, we see a nearly 10-fold reduction of
326 polymerization for P301S tauRD compared to WT tauRD (Fig. 5a and Supplementary Fig. 7b).
327 P301S tauRD_QK and P301S tauRD_EL_QK constructs do not significantly recover MT binding
328 (Fig. 5a); however, the above baseline polymerization of the P301S tauRD_QK had a $t_{1/2max}$ time
329 approaching that of the 'PGGG' containing tauRDs (Fig. 5b and Supplementary Fig. 7b).

330 We similarly tested the 3R tauRD constructs with the polymerization assay at 10 μM and observed
331 similar behavior. Proline to serine constructs had low polymerization, while the 3R-tauRD_DI_VS
332 had deceased polymerization relative to the 3R WT tauRD construct (Supplementary Fig. 7c). Our
333 data recapitulate the observation that the P301S mutations significantly alter MT binding. In the
334 context of the 'PGGG' turns, the R1 sequence in the tauRD_EL_QK and 3R WT tauRD appears
335 to increase the polymerization extent without changing the polymerization rate.

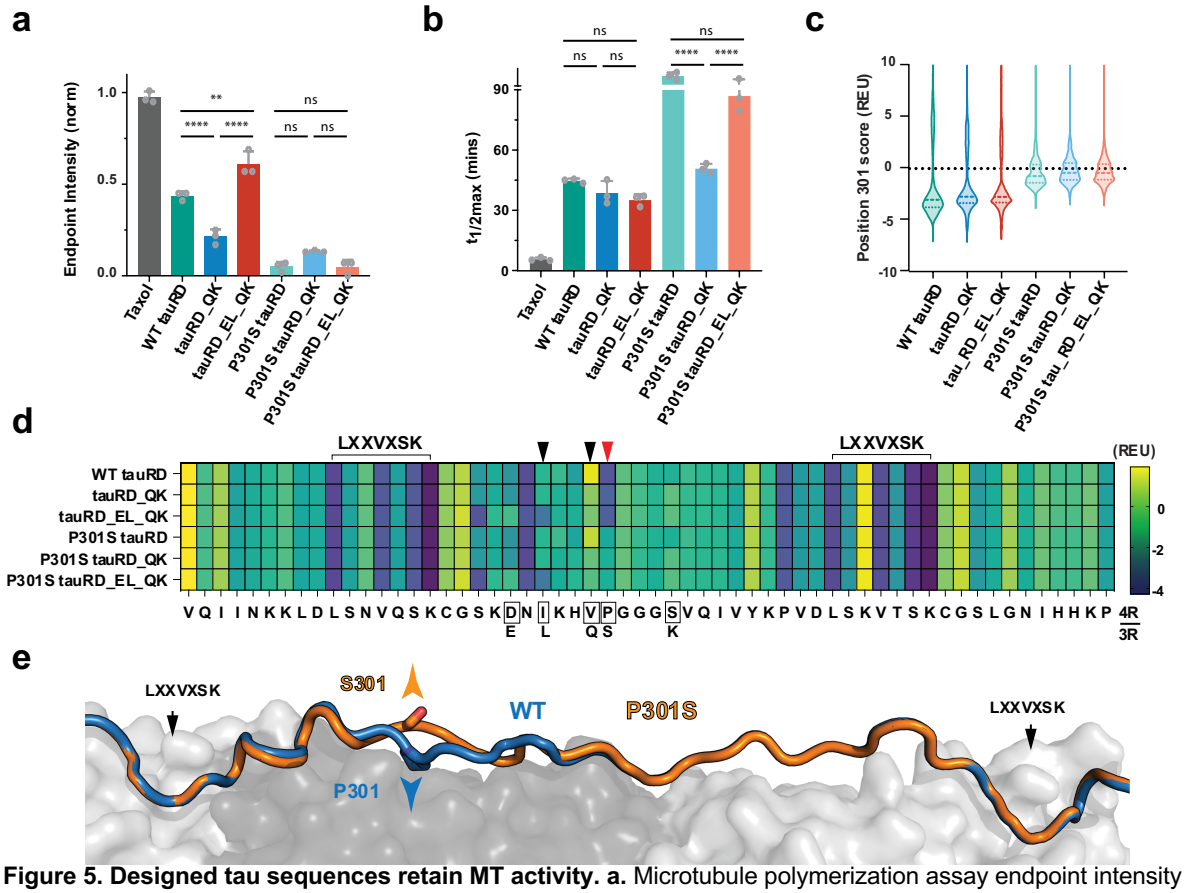


Figure 5. Designed tau sequences retain MT activity. a. Microtubule polymerization assay endpoint intensity of tauRD constructs at 10 μM for 'PGGG' containing constructs WT tauRD (green), tauRD_QK (blue), tauRD_EL_QK (red), and 'SGGG' containing constructs P301S tauRD (cyan), P301S tauRD_QK (sky blue), and P301S tauRD_EL_QK (orange). Data is shown as intensity mean and standard deviation, significance calculated from column effects using 2way ANOVA (*=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001). **b.** Microtubule polymerization assay $t_{1/2max}$ fit to each polymerization curve plotted as an average and standard deviation, significance calculated as in a. **c.** Energetics of residue 301 computed across 2500 MT:tau models for six constructs, colored as in a. The data is shown as violin plots with median and 25/75% quartiles labeled in dashed lines. **d.** Residue energy for tau 275-331 from MT:tau modeled complex for designed constructs. Energies are shown as the mean REU of 2,000 structures colored from high to low energy (yellow to navy)—fragment sequence below with boxed residues signifying the 4R tauRD sequence. Labels above the heatmap denote MT:MAP anchor motifs. Red arrows denote position 301, and blank arrows denote positions 297 and 300, which have lower energy for designed mutations. **e.** Lowest energy structure of modeled sequence for WT tau275-331 (blue, cartoon) and P301S tau275-331 (orange, cartoon) in complex with MT (gray, surface). Black labels denoting MT:MAP anchor points, position 301 labeled for WT pointing down (blue, sticks), and P301S pointing up (orange, sticks).

Modeling the MT: tau complex reveals key stabilizing interactions.

Local structure in tau has been implicated to be important for its binding to MTs, with the addition of the P301S mutation resulting in an almost complete loss of binding to tau[50]. Available structural models of tau in the context of microtubules include tau in a fully extended conformation determined by cryo-EM[51] and a more compacted conformation visualized by NMR[52]. To explore how these turn sequence changes may affect the bound state of the MT: tau complex, we modeled the interaction energetics of different tau constructs bound to MTs using an available cryo-EM structure[51]. First, we built WT and mutant 4R tauRD using a threading protocol into the 4xR2 tau

345 from the cryo-EM model (PDB ID 6CVN) to rebuild the missing sequence (residues 300-303),
346 then trimmed the tau model to contain only repeats 2 and 3 (residues 275-331) and minimized the
347 resulting low energy structure in the context of tubulin □/□/□ trimer. The energetics and
348 conformational heterogeneity were compared for all 6 constructs and we visualized the 10 lowest
349 energy complexes (Supplementary Fig. 7d-f). We compared the energetics of our models by
350 scoring the 301 positions of our MT: tau complex and observed a difference in energetics between
351 the 'PGGG' and 'SGGG' models (Fig. 5c).

352 Indeed, plotting the per residue energy contributions of the fragment, we find that the proline to
353 serine yields the most dramatic shift (Fig. 5d, red arrow) and sits firmly between two well-
354 established tubulin binding motifs (Fig. 5d, LXXVXSK motifs)[52,53]. The changes in the energy
355 terms that dictate this difference are attractive and solvation energies suggestive of positive
356 contributions to the burial of the proline ring (Supplementary Fig. 7g). While the serine has the
357 potential to form an additional hydrogen bond, the energy contribution from that interaction does
358 not overcome the increased van der Waals forces (Supplementary Fig. 7g). At the same time, the
359 serine residue points away from the MT (Fig. 5e). Interestingly, we also find that valine to
360 glutamine mutations at position 300 (from tauRD_QK), as well as isoleucine to leucine mutations
361 at position 297 (from tauRD_EL_QK), have lower energy and contribute to complex stability (Fig.
362 5e, black arrows). This boost in stabilization may explain why the tauRD_EL_QK may yield
363 higher MT binding compared to the WT tauRD (Fig. 5a). By modeling the 'PGGG' loop within
364 the available MT: tau structure, there is clear importance of the proline at position 301, suggesting
365 that the proline position not only affect the propensity to aggregate but impact the MT bound state
366 energetics.

367 **Stabilizing mutations render tau less sensitive to exogenous seeds in cells.**

368 Finally, to evaluate whether our designed tau sequences influence response to pathogenic seeds,
369 we developed a cellular assay leveraging mEOS3.2 fusions of tauRD harboring our designed
370 mutants in the context of the pathogenic P301S mutation. To assess the behavior of these mutants,
371 we employed a HEK293T cell model of tauRD aggregation leveraging the FRET-compatible
372 photoconvertible mEOS3.2 system. We produced cell lines stably expressing P301S, P301S
373 tauRD_QK, and P301S tauRD_EL_QK but also single mutants P301S tauRD encoding V300Q
374 (herein P301S tauRD_Q) and P301S tauRD encoding S305K (herein P301S tauRD_K), as fusions
375 to mEOS3.2, which were sorted to maintain similar high expression levels across cell lines. The
376 cells were seeded with a concentration series of recombinant tau fibrils, photoconverted, and then
377 analyzed by flow cytometry (Fig. 6a). To ensure optimal FRET efficiency, the photoconversion
378 time was determined empirically (Supplementary Fig. 8a). Then, we employed a gating strategy
379 to observe the proportion of cells that were FRET-positive (Supplementary Fig 8b). As observed
380 previously, P301S tauRD-mEOS3.2 yielded high percentages of cells, with 74±4% of FRET-
381 positive cells seeded with 100 nM recombinant tau fibrils (Fig. 6b, green). The introduction of
382 V300Q and S305K mutations alone reduced seeding 1.2-fold (Fig. 6b, yellow and orange) while
383 combining these mutations in the P301S tauRD_QK reduced seeding activity nearly 2-fold (Fig.
384 6b, blue). Finally, introducing the P301S tauRD_EL_QK reduces seeding activity by 2-fold
385 compared to the P301S tauRD control (Fig. 6b, red). Representative images of cells expressing
386 each of the 6 tauRD constructs controlled or seeded with 100 nM recombinant tau fibrils show a
387 reduction in puncta frequency (Fig. 6c). This data reveals that modification of the turn residues
388 V300Q and S305K is sufficient to reduce P301S tauRD sensitivity to pathogenic tau seeds in cells.

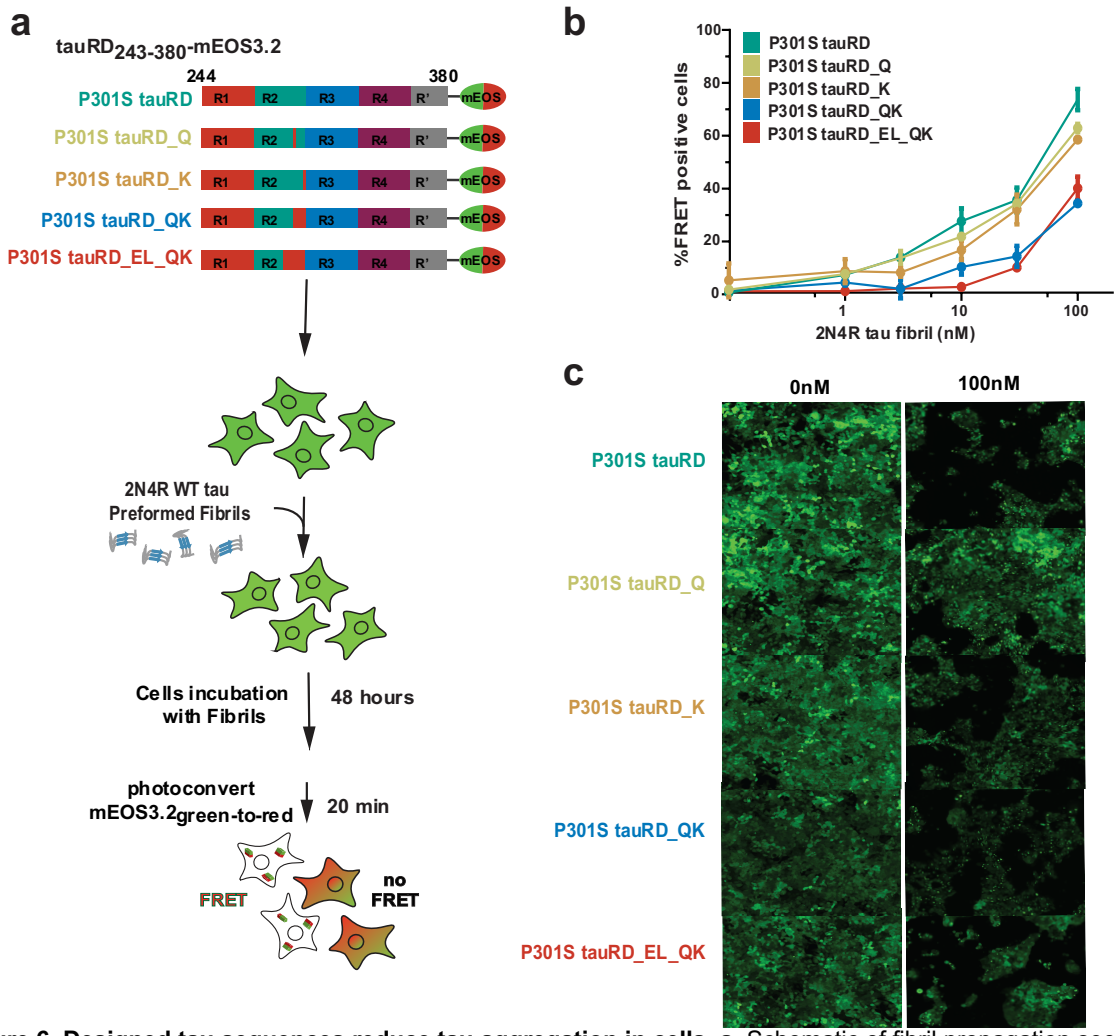


Figure 6. Designed tau sequences reduce tau aggregation in cells. a. Schematic of fibril propagation seeding workflow. Cartoon of 4R tauRD P301S designed constructs encoding 3R-like mutants as fusions to mEOS3.2, used for lentiviral transfection into HEK 293T cell and sorted for high expression, treated with lipofectamine mediated seeding from 2N4R WT tau preformed fibrils, followed by incubation time, photoconversion and analysis by percent FRET positivity. **b.** Percent FRET positivity (FRET pairs: tauRD-mEOS3.2green/tauRD mEOS3.2red) was measured by flow cytometry on n = 3 biological replicates of at least 10,000 cells per fibril concentrations and expression construct. Data are shown as averages with standard deviation. **c.** Imaging of HEK 293T cells expressing designed constructs 48hrs after no treatment control and max treatment of 100 nM fibril concentration.

DISCUSSION

The conversion of tau into insoluble aggregates is linked to several neurodegenerative diseases. Tau isoform conversion between 3R and 4R leads to different aggregation propensities; nevertheless, minor sequence changes surrounding amyloidogenic motifs provide mechanistic hints toward the intrinsic regulation of tau aggregation. We find that the conserved proline in the 'PGGG' motif, coupled with the residues that flank it, plays a vital role in limiting the self-assembly of the amyloidogenic ^{306}VQIVYK^{311}. This region adopts multiple conformations in the end-state of the tau fibril fold, providing multiple residues that can form stabilizing interactions with the amyloid motif. We observe that the local rearrangements driven by P301S mediate the

399 exposure of the amyloid motif in the 4R context, which the 3R-based designed sequence can
400 reverse. Introducing the P301S mutation in the 4R-derived fragment yields more disordered and
401 dynamic states for glycine 302, reducing local and long-range interactions and increasing the
402 exposure of the amyloid motif. This P301S destabilizing effect is minimized in the context of only
403 two residue swaps, V300Q and S305K, with minimal changes in the dynamics. With these
404 mutations, we see a rearrangement at the 'PGGG'/'SGGG' region, increasing intermediate-range
405 interactions. We propose a structural model for how the proline to serine mutation is the main
406 driver of alterations in binding to MTs. Yet, we identify differential interactions for the 294-300
407 residues to MTs between 3R and 4R tau that increased microtubule polymerization when
408 containing a 'PGGG' motif. We finally detect decreased sensitivity to pathogenic seeds through
409 our paired mutations. Together, our data suggest an avenue for tuning tau aggregation through
410 small local structure rearrangements that alter turn dynamics which retain biological activity.

411 FTD-linked pathogenic mutations largely localize to the repeat domain of MAPT and, more
412 specifically, cluster at inter-repeat elements[54], including specific mutations in the variable exon 10
413 of *MAPT* associated with altering splicing[54]. A unifying mechanism for how a diversity of FTD-
414 linked mutations drive the same disease is unclear; cryo-em structures of tau amyloids isolated
415 from FTD-tau patients may provide hints in the future. It remains unknown how these mutations
416 lead to tau-based dysfunction. Still, the specificity of the pattern of mutations in MAPT may
417 suggest that loss-of-function (LOF) and gain-of-function (GOF) mechanisms contribute to
418 pathological fibrilization and impaired microtubule binding, ultimately leading to FTD-tau. The
419 mutation sites often localize in proximity to amyloid motifs, suggesting that mutations may disrupt
420 biological function and promote amyloid-motif-dependent tau self-assembly. Indeed, analyzing
421 disease-associated mutations in extensive repositories of globular proteins (using AlphaFold
422 models) reveals distinct patterns: LOF mutations are randomly distributed across the protein
423 structure. In contrast, GOF mutations cluster at sites critical to biological function, where they
424 cause disruption[15]. Applying this concept to MAPT may hint at a unifying mechanism of FTD-tau
425 disease. Developing new strategies to modulate tau is essential to better understand how to mitigate
426 GOF and LOF effects.

427 The field has focused on understanding the pathological effects of endpoint aggregates to guide
428 potential ways to diagnose or treat disease. How tau interacts with MTs has garnered less interest.
429 It remains unclear what features of microtubule-associated proteins (MAPs) are required for
430 binding to MTs. The *MAPT* gene is highly conserved across evolution[53]. Similarly, the related
431 *MAP2* and *MAP4* genes are conserved but do not encode aggregation-prone sequences[53]. These
432 genes are processed similarly during splicing, producing both 3R and 4R isoforms of the sequence
433 that retain binding to MTs but are not implicated in misfolding diseases. Two amino acids in the
434 tau amyloid motif, **V**Q**I**VYK, are conserved across MAPs. Because this site is essential for
435 splicing, and even codon changes can alter splicing, part of the amyloid motif may be preserved
436 to retain efficient splicing. However, it is still being determined why the latter part of the amyloid
437 motif is conserved. Our data demonstrate that substituting two residues flanking the turn preceding
438 the VQIVYK motif in tau to mimic the 3R localized sequence reduces 4R tau aggregation and
439 maintains tau's ability to stabilize MTs. Prior cryo-EM structural data on tau bound to MTs showed
440 a fully extended conformation, highly disordered at the inter-repeat regions surrounding the
441 'PGGG' motif. By contrast, NMR studies showed a relatively collapsed conformation of this
442 element bound to MTs, adopting a turn conformation around the 'PGGG' motif. Our tau: MT
443 models of the bound state suggest a role for the wild-type proline the interaction complex.

Future experiments and analysis must more closely interrogate the relationship between the tau monomer ensemble in solution, MT binding states, and the early misfolding steps that lead to disease. We suspect these intermediate and long-range interactions may be important in forming the MT: tau encounter complex. Interestingly, XL-MS experiments on tauRD encoding mutations that convert the local sequence from 4R-like to 3R-like alter the frequency of global long-range contacts but increase intermediate-range contacts. This shift in frequency of interactions correlates with stabilization of the 'PGGG' turn to reduce aggregation and maintain or even increase MT activity. Future experiments may explore how the computational design of tau may lead to mutants that can replace wild-type MAPT to reduce disease. A deeper comparison of naturally occurring MAP sequences may offer clues for the design of tau sequences that retain MT binding but aggregate less. Thus, more focus must be brought onto the effect of mutations on the multi-state problem of tau conformation in solution (aggregation-resistant), tau bound to MT (functional state), and tau misfolding towards pathological aggregates to begin to understand the origins of tau in function and disease.

A deeper interpretation of the evolvability of tau to tune these two properties may be inferred from the evolutionary sequence history of MAPs. Future studies must test the biological consequences of designer MAPT sequences in animal models as ongoing therapeutic approaches to reduce tau levels may lead to unknown consequences in humans.

METHODS

Peptide synthesis and disaggregation

All peptides were synthesized as ordered by Genscript with N-terminal acetylation and C-terminal amidation modifications at >95% purity. Peptides were disaggregated as previously described[55]. In brief, peptides were resuspended in a 250 µL Trifluoroacetic acid (TFA) (Pierce) incubated at room temperature (RT) for 1-16 hours in a chemical fume hood. The peptide solution was dried under a stream of nitrogen gas, and then 100 µL of H_2O was added, followed by flash freezing, and immediately placed under vacuum to remove any residual solvents. The peptide residue was resuspended in 200 µL of H_2O to determine the concentration and, with 10X PBS pH 7.4, adjusted with NaOH and H_2O to achieve the desired pH and concentration.

ThT fluorescence aggregation assays

Peptide concentrations were adjusted with 1X PBS (136.5 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and 1 mM Dithiothreitol (DTT) to a desired concentration between 800 µM and 50 µM. Thioflavin T (ThT) was added to the samples at a final concentration of 25 µM. Wild-type or mutant 4R or 3R tauRD protein was diluted in 1X PBS with 10 mM DTT to the desired concentration of 40, 20, 10, and 5 µM. A 2X heparin concentration (Amsbio) was added to the tauRD protein. All samples were aliquoted at 45 µL per replicate in a 384-well clear bottom plate. All conditions were performed as three independent replicates at 37 □C. ThT kinetic scans were run every 30 min, with 10 seconds of orbital shaking before each read, on a Tecan Spark plate reader at 446 nm Ex (10 nm bandwidth) and 482 nm Em (10 nm bandwidth). Blank wells containing buffer and ThT were subtracted from experimental values. The data were plotted,

485 and the $t_{1/2}$ values were estimated using a non-linear regression model fitting in GraphPad Prism
486 and reported as averages with a standard deviation.

487 **Peptide NMR and secondary structure calculation from chemical shifts**

488 NMR spectra were acquired on Agilent DD2 spectrometers operating at 600 MHz at the UTSW
489 Biomolecular NMR Facility. ^{1}H-^{1}H TOCSY spectra[56] and ^{1}H-^{1}H NOESY spectra[56] were acquired
490 at 25 □C. Peptides were adjusted to 500 µM in 1X PBS, 1mM DTT with 10% D_2O or 99.998%
491 D_2O, and the total acquisition time for ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY spectra ranged from 8-
492 11 hours. Five ^{1}H-^{1}H TOCSY spectra were collected in series for the aggregation-prone P301S
493 R2R3$_{294-311}$ peptide, each 10 hours long. All data were processed with NMRpipe[57] and analyzed
494 with CARA (http://cara.nmr.ch/doku.php) and SPARKY[58]. Secondary structure propensity was
495 calculated using the neighbor-corrected secondary structure propensity calculator (ncSPC) from
496 the Ha chemical shift assignments. The calculator used the deviations from random coil chemical
497 shifts to estimate the proportion of secondary structure per residue corrected by a specified
498 neighbor window. For our analysis, we used a neighboring residue window of 3[36,37]. The data were
499 plotted in GraphPad Prism.

500 **Tau amyloid fibril structure comparison**

501 Evaluated fibril structures were taken from previous analyses of structural stability (PDB accession
502 codes: 5o3l, 5o3t, 6npw, 7p66, 6tjo, 7p6d, 7p6a, 7p65, 6gx5, 8q92). Folds categorized as described
503 Alzheimer's disease (AD), chronic traumatic encephalopathy (CTE), Globular glial tauopathy
504 (GGT), corticobasal degeneration (CBD), Argyrophilic grain disease (AGD), Progressive
505 supranuclear palsy (PSP), GGT/PSP-fold (GPT), Picks disease (PiD), and PS19 AD mouse model
506 with human tau derived fibrils (P301Sm). All structures were aligned based on Cß positions of the
507 ^{295}VQIVYK^{311} motif and images created using Pymol. Solvent accessible surface area (SASA)
508 was calculated for the middle chain of the fibril. Then, the per residue normalized SASA value
509 was calculated as a ratio to the maximum estimated SASA of each residue as described in Tien et
510 al. 2013[59]—normalized DDG values obtained from Mullapudi et al. 2023{Mullapudi, 2023 #4}.
511 New values were calculated for P301Sm fibril (PDB accession code 8q92) following the
512 previously described protocol {Mullapudi, 2023 #4}. Briefly, the PDB structure was expanded to
513 obtain 9 layers of tau monomer. The flex-DDG fibril alanine scan method was implemented, and
514 the change in ΔG between the alanine mutant and the WT structure at each position was calculated.
515 All DDG values were min-max normalized to each structure's 295-311 region.

516 **Molecular dynamics simulations production**

517 Molecular Dynamics simulations were performed using PLUMED patched GROMACS v2018.4.
518 Each peptide was simulated in a dodecahedron box with SPC/E water starting from a fully
519 extended conformation built-in Pymol for a total of 5 µs followed by 15 µs in TIP3 water with 150
520 mM NaCl ions. The initial dodecahedron simulation box was constructed using the extended
521 peptide structure with a minimum 1.5 nm boundary condition. The AMBER99sb-ildn
522 forcefield[60] was used for all simulations. After an initial 1000 steepest descent steps of converged
523 energy minimization, 10 ns of NVT and 20 ns of NPT (first 10 with Berendsen[61] and the last 10
524 with Parrinello-Rahman[62] barostats) equilibrations were performed.

525 The subsequent production level trajectories are based on 5 fs time steps using hydrogen-only
526 virtual sites[63]. Production level trajectories were obtained for an NPT ensemble with Parrinello-
527 Rahman barostat and periodic boundary conditions with Particle Mesh Ewald (PME)[64] summation
528 for long-range electrostatics. A total of 20 µs trajectories were generated by producing five 1 µs
529 simulations per peptide. These trajectories were analyzed using the GROMACS clustering
530 function with a cluster cutoff of 0.54 nm. The cluster cutoff was chosen empirically to produce a
531 maximum of 15 clusters for all four peptides. The center structure of the top 5 clusters was then
532 used as the starting structure for simulations with Na+ and Cl- ions included in the solvent. These
533 second-stage simulations were performed with the same minimization, equilibration, and
534 production workflow with only a few modifications. A dodecahedron simulation box was used,
535 using the new starting structures with an increase in the minimum boundary cutoff to 1.6 nm. The
536 same AMBER99sb-ildn forcefield[60] was used with an update to a TIP3P water model with 150
537 mM ionic strength by including Na+ and Cl- ions. The energy minimization step was performed
538 with the steepest decent algorithm to obtain a maximum force <1000.0 kJ/mol/nm. Equilibration
539 was run aligned with the previous step. The subsequent production level trajectories were
540 performed as the initial simulations, with an updated 2 fs timestep to account for the ionic solvent.
541 A total of 60 µs trajectories were generated by producing five 3 µs trajectories per peptide. All
542 simulations were done on UTSW's BioHPC computing cluster.

543 **MD analysis: clustering, RMSD, and distance calculations**

544 All analysis was done using GROMACS 5.0.4[70]. The analysis framework was done as previously
545 described[55]. Specifically, backbone distance in Fig. 2b and Supplemental Fig. 3a was calculated
546 using the 'gmx pairdist' command between the center of mass of backbone atoms for residues 1-
547 6 (i.e., XXNXKH) vs. residues 12-18 (i.e., XVQIVYK) from the 15 µs ensemble of the second
548 production stage and the 5 µs of the initial replicate simulations. Data is shown as violin plots (e.g.,
549 density distributions) with mean and 25% and 75% quartile marked. For the second step of the
550 simulation, the 5 independent trajectories were concatenated into a 15 µs ensemble, and the 2-
551 dimensional RMSD was calculated every 10 ns using the 'gmx rms' command. The 15 µs
552 ensembles were also compared to the published structure of tau. The RMSD of the ensemble
553 referenced the P301S fibril structure (PDB ID 8Q96), which was reported as a distribution violin
554 plot in Supplemental Fig. 3e. All graphs were plotted directly from gmx output or using GraphPad
555 Prism. Images were created using Pymol for cluster structures.

556

557 **MD state analysis with Markov State Model (MSM)**

558 MSM of the peptide simulations was built from the combined ensemble of all four peptides and
559 the individual peptide ensembles, concatenating each peptide 15 µs simulations in ionic solvent
560 using the pyEMMA software. We constructed the MSM of these trajectories to capture the kinetic
561 behavior of the system with the PyEMMA 2.5.12 library[40] in Python 3.10.9. After the second step
562 of MD simulation WT R1R3$_{263\text{-}311\Delta R2}$, P270S R1R3$_{263\text{-}311\Delta R2}$, WT R2R3$_{295\text{-}311}$, and P301S R2R3$_{295\text{-}}$
563 $_{311}$, we trimmed the 50 µs trajectory at 100 ps timestep, and each resulting trajectory (5
564 concatenated trajectories for 4 systems) contains 30001 frames. These trajectories were trimmed
565 again to contain only Cα atoms and concatenated to generate a 60,000,001-frame trajectory. The
566 MSM featurization step was performed based on the minimum pairwise distance between residues
567 1-7 (i.e., TENLKHQ vs. KDNIKHV) against amyloid motif residues 12-18 (i.e., XVQIVYK)
568 with a total 49 distance pairs with 150001 frames for a single system. We considered the four

569 featurization techniques for model construction: torsion angle, backbone atom positions, backbone
570 atom intra-molecular distances, and minimum pairwise distance between amyloid motifs. The
571 featurization criteria were chosen based on VAMP-2 scores (Variational Approach for Markov
572 Process).

573 We selected the minimum pairwise residue distance over the torsion angles because these motifs
574 covered functional conformational change and reported on interactions across the peptide region
575 that differ in sequence. Further dimensionality reduction was performed from a total of 49
576 pair distances (600000001,49; 150001,49) to 2 dimensions (600000001,49; 150001,2) using the
577 time-lagged Independent Component Analysis (tICA). TICA identifies the slowest mode of feature
578 space and maximizes the autocorrelation of reduced coordinates for state decomposition. After
579 that, we used the k-means clustering algorithm for the decomposition of reduced conformational
580 space into 200 microstates. Those conformations that belong to the same energy basin in the TICA
581 subspace are clustered into each microstate, and transition occurs quickly among these states.
582 Further, we lumped the microstates into 6 macrostates using PCCA+ to predict the major
583 conformational changes in peptides. The MSM was determined on an implied time scale plot, and
584 the lag time 200 (time/step) was chosen. The Chapman-Kolmogorov test was applied to validate
585 the consistency of probabilities predicted from MSM and MD simulations. The calculated MSM
586 was further coarse-grained into hidden Markov models to calculate the metastable macrostates and
587 transition probabilities. Two hundred representative structures were computed for each state
588 extracted from the centroid of the most populated microstate. We calculated the pairwise distance
589 for all residue pairs for the 200 structures of each cluster using 'gmx pairdist' in gromacs and
590 plotted the mean distance and SD in heatmaps using MATLAB plotting software. The first 10
591 structures were used as representative models for each state. Images were created using PyEMMA
592 graphics output options or Pymol for representative structures.

593 **Recombinant tauRD and full-length tau production**

594 We utilized several forms of recombinant tau. Recombinant repeat domain (tauRD) 4R isoform
595 (4R tauRD, 4R P301S tauRD, 4R tauRD_EL_QK, 4R tauRD_EL_QK, and 4R P301S tauRD_QK)
596 and 3R isoform (3R tauRD, 3R P270S tauRD, 3R tauRD tauRD_VS_DI and 3R P301S tauRD
597 tauRD_VS_DI) constructs were purchased from Twist Bioscience encoded into pET-29b(+) vector
598 without tags. Each plasmid was transformed into BL21-Gold (DE3) cells. Cells were grown in 1X
599 Terrific Broth media with 1 µg/ml kanamycin to OD600 of 1.4 and induced with 1 mM isopropyl
600 β-D-1-thiogalactopyranoside (IPTG) for 3 h at 37 °C. To produce N^{15} isotope-labeled samples of
601 4R tauRD, 4R tauRD P301S, 4R tauRD_EL_QK, and 4R tauRD P301S tauRD_EL_QK, a single
602 colony of freshly transformed plasmid into BL21(DE3) was inoculated into 10mL 2xLB
603 supplemented with 0.05mg/mL Kanamycin and incubated for 7-8 hours at 37°C while shaking at
604 220rpm. The 10mL culture was then mixed into 100mL of M9 minimal media (42mM Na_2HPO_4,
605 22mM KH_2PO_4, 8.5mM NaCl, 0.1mM CaCl, 2mM Mg_2SO_4, 1E-4% Thiamine, 0.4% Glucose,
606 187mM NH_4Cl, 0.05mg/mL Kanamycin) and incubated overnight at 37°C while shaking at
607 220rpm. In the morning of the following day, the cells were spun down at 2,000g for 10 minutes
608 and resuspended in 20 mL of M9 minimal media containing 15N labeled NH_4Cl in place of the
609 unlabeled molecule. This resuspension was immediately added to 1L of M9 minimal media with
610 ^{15}N labeled NH_4Cl and incubated at 37°C shaking at 220rpm. Once OD600 reached 0.6-0.8 A.U.,
611 1mL of 1M IPTG was added to induce protein expression for 3 hrs at 37ºC. The cells were
612 harvested and lysed in 20 mM 2-(N-morpholino) ethane sulfonic acid (MES), 1 mM
613 Ethylenediaminetetraacetic acid (EDTA), 1 mM MgCl, 5 mM β-mercaptoethanol (BME), 1 mM

614 PMSF, pH 6.8, using an Omni Sonic Ruptor 400 with lysate on ice. After sonication, 500 mM
615 NaCl was added to the lysates, which were boiled for 20 minutes before clarifying by
616 centrifugation. The supernatant of the clarified lysate was exchanged into 20 mM MES, 50 mM
617 NaCl, 5 mM BME, and pH 6.8 by dialysis. The dialyzed sample was filtered, applied to a HiTrap
618 SP HP (GE), and eluted with a 50 mM–1 M NaCl gradient. Tau-containing fractions were
619 concentrated on an Amicon-15 concentrator (5KD cutoff) and applied to HiLoad 16/600 Superdex
620 75 pg (GE) and eluted into 1X PBS (136.5 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM
621 KH_2PO_4, pH 7.4), 10 mM DTT. Aliquots were all stored at −80 °C in 1 X PBS.

622 The pET-28b-tau plasmid encoding full-length 2N4R WT tau was a gift from Dr. David Eisenberg
623 (UCLA). The plasmid was transformed into BL21-Gold (DE3) cells. Cells were grown in 1X
624 Terrific Broth media to OD600 of 1.4 and induced with 1 mM isopropyl β-D-1-
625 thiogalactopyranoside for 3 h at 37 °C. The cells were harvested and lysed in 50 mM Tris, 500 mM
626 NaCl, 1 mM β-mercaptoethanol, 20 mM imidazole, 1 mM phenylmethylsulfonyl fluoride (PMSF),
627 pH 7.5, using an Omni Sonic Ruptor 400 at 4 °C. The lysates were centrifuged at 13,000 x g for
628 20 minutes at 4 °C, and the supernatant was applied to a Ni-NTA gravity column and eluted with
629 50 mM Tris, 250 mM NaCl, 1 mM β-mercaptoethanol, and 300 mM imidazole. Eluting fractions
630 containing tau were desalted into 50 mM MES, 50 mM NaCl, and 1 mM β-mercaptoethanol (pH
631 6.0) by PD-10 column (GE). Exchanged fractions were applied to a HiTrap SP HP (GE) and eluted
632 with a 50 mM–1 M NaCl gradient. Tau-containing fractions were concentrated on an Amicon-15
633 concentrator (10kDa cutoff) and applied to a HiLoad 16/600 Superdex 200 pg (GE) and eluted into
634 1X PBS (136.5 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). Aliquots
635 were all stored at −80 °C in 1X PBS.

636 **4R tauRD NMR experiments**

637 ^{1}H-^{15}N HSQC spectra, ^{1}H-^{15}N NOESY-HSQC, and ^{1}H-^{15}N TOCSY-HSQC spectra were acquired
638 at 10 □C. NMR spectra were acquired on Agilent DD2 spectrometers operating at 800 MHz at the
639 UTSW Biomolecular NMR Facility. tauRD constructs were concentrated to 200-300μM in tauRD
640 NMR buffer (10mM phosphate buffer, 100mM NaCl, and 4mM DTT with 10% D_2O), replicating
641 buffer conditions of previously published data. The concentration of tauRD constructs was
642 estimated using 215nm absorbance and verified to have a comparable signal-to-noise for NMR
643 spectra. The total acquisition time ranged from 64-80 hours for ^{1}H-^{15}N NOESY-HSQC and ^{1}H-
644 ^{15}N TOCSY-HSQC spectra. The four tauRD constructs were then diluted to 50uM, and ^{1}H-^{15}N
645 HSQC spectra were measured for 4 hours. All data was processed with NMRpipe[57] and analyzed
646 with CARA (http://cara.nmr.ch/doku.php), SPARKY[58], and ccpNMR[65]. Peak assignments were
647 transferred from deposited data for tau K18 (bmrb 19253)[66], and additional N-termini peaks were
648 assigned from htau4p (bmrb 50701)[67]. Mutations in tauRD_EL_QK and P301S tauRD_EL_QK
649 were assigned based on sidechain assignments in ^{1}H-^{15}N NOESY-HSQC and ^{1}H-^{15}N TOCSY-
650 HSQC and assignments of K19 tau made from 3R tauRD fragment[68] (Visualized Source Data 6).
651 No assignments were made for the V300Q mutation.

652 (^{1}H)-^{15}N heteronuclear NOE relaxation data (hetNOE) for 150-300 μM of 4R tauRD, P301S
653 tauRD, tauRD_EL_QK, and P301S tauRD_EL_QK samples in tauRD NMR buffer 10mM were
654 obtained using standard Bruker pulse sequences acquired at 10 °C in an 800MHz (^{1}H) Bruker
655 AVNEO spectrometer equipped with a Z-gradient cryoprobe. Protein concentrations were
656 estimated using a 215nm absorbance signal; as stated previously, peak intensities were compared
657 across samples to ensure a 20% difference in intensities. Similar experiments were performed at

658 25 °C for 300uM P301S tauRD and P301S tauRD_EL_QK, and transferring assignments was not
659 possible as published data is at 10 °C and peak intensities were lost or shifted for most peaks,
660 interpreting the data impossible. Limited assigned peaks and intensity differences are available in
661 Source Data 5. In all relaxation experiments, the spectral width was 7812.5 Hz for ^{1}H and 1824.25
662 Hz for ^{15}N dimensions. HetNOEs were calculated from the ratio of cross-peak intensities in spectra
663 collected with and without amide proton saturation during the recycle delay. Uncertainties in peak
664 heights were determined from the standard deviation of the intensity distribution in signal-less
665 spectral regions. All NMR spectra were processed in Topspin 4.1.1 and analyzed in Sparky
666 3.1969[58]. The data were plotted in GraphPad Prism.

667 **Transmission electron microscopy**

668 An aliquot of 5 µL of the sample was placed onto a glow-discharged Carbon/Formvar-coated 400-
669 mesh copper grid for 1 minute, washed with distilled water, and then negatively stained with 2 µL
670 of 2% uranyl acetate for 1 minute. Images were acquired on a Tecnai G^{2} spirit transmission
671 electron microscope (FEI, Hillsboro, OR), serial number D1067, equipped with a LaB_6 source at
672 120 kV using a Gatan ultrascan CCD camera.

673 **Gel-extracted sample preparation for cross-linking mass spectrometry analysis**

674 174 µg of each construct of 4R tauRD (WT, QK, EL_QK, P301S, P301S QK, and P301S_EL_QK)
675 was diluted in 1X PBS, 10 mM DTT to a final concentration of 58 µM. The cross-linking reaction
676 was conducted by adding disuccinimidyl suberate (DSS, DSS-d0 and -d12, Creative Molecules,
677 dissolved in DMF) to a final concentration of 1 mM for three minutes at 37 °C while shaking at
678 350 RPM. The reaction was then quenched by adding ammonium bicarbonate to 50 mM final
679 concentration and incubating for 30 minutes at 37 °C while shaking at 350 RPM. Cross-linked
680 samples were resolved on SDS-PAGE gels (NuPAGE™, 4 to 12%, Bis-Tris, 1.5 mm), and the
681 monomer band was extracted from the gel for XL-MS analysis using the following protocol. Bands
682 were sliced into 1mm^{3} pieces and washed twice with MQ water, discarding supernatant after each
683 time. Gel pieces were then covered in acetonitrile/50 mM ammonium bicarbonate (mixed in ratio
684 2:3, v/v) and sonicated for 5 minutes, and the supernatants were discarded each time. This step
685 was repeated twice. Next, gel pieces were washed in 100% acetonitrile until they became white,
686 then the liquid was removed, and gel pieces were dried by lyophilization. Subsequently, gel pieces
687 were covered in 25 mM ammonium bicarbonate with 10 mM DTT and incubated for 1 hour at
688 56 °C while shaking at 500 RPM. Next, samples were cooled down, supernatants removed, and a
689 freshly prepared 25 mM ammonium bicarbonate solution with 55 mM iodoacetamide was added.
690 Gel pieces were incubated for 45 minutes in the dark at room temperature. Then, gel pieces were
691 sequentially washed with 25 mM ammonium bicarbonate, 50% acetonitrile, and 100% acetonitrile
692 with quick vertexing after each wash. The liquid was removed, and gel pieces were dried by
693 lyophilization. Trypsin digestion was initiated by adding 1:10 (trypsin to protein, m/m) mass
694 spectrometry grade trypsin (New England Biolabs #P8101S) in 50 mM ammonium bicarbonate.
695 Then, the reaction was incubated overnight at 37 °C while shaking at 500 RPM. Digestion was
696 stopped by adding 2% (v/v) formic acid, and supernatants were collected. Peptides were
697 additionally extracted by covering gel pieces in MQ water/acetonitrile/formic acid (50:50:0.1,
698 v/v/v), incubated for 15 minutes at 37 °C while shaking at 500 RPM, and then sonicated for 5
699 minutes. This step was repeated twice, and collected supernatants were pooled, snap-frozen, and
700 lyophilized. Peptides were resuspended in 50 mM ammonium bicarbonate/acetonitrile/formic acid
701 (93:5:2), further purified by solid phase extraction using Sep-Pak tC18 cartridges (Waters

702 #WAT054960), snap frozen and lyophilized. Samples were processed as 5 technical replicates on
703 a Thermo Orbitrap Fusion Lumos at UTSW Proteomics core. The mass spectrometer was operated
704 in data-dependent mode by selecting the five most abundant precursor ions (m/z 350–1600, charge
705 state 3+ and above) from a preview scan and subjecting them to collision-induced dissociation
706 (normalized collision energy = 35%, 30 ms activation). Fragment ions were detected at low
707 resolution in the linear ion trap. Dynamic exclusion was enabled (repeat count 1, exclusion
708 duration 30 s). Data was processed to identify unmodified, 1 mono-link, 2 mono-links, and loop-
709 links by peak areas. Data was reported as fraction modification normalized to the total peak area
710 per replicate.

711 **Analysis of mass spectrometry data**

712 Thermo.raw files were converted to the open.mzXML format using msconvert
713 (proteowizard.sourceforge.net) and analyzed using an in-house version of xQuest[46]. Spectral pairs
714 with a precursor mass difference of 12.075321 Da were extracted and searched against the
715 respective FASTA databases containing tau (TAU_HUMAN P10636-8) or designed constructs
716 with and without P301S mutations. xQuest settings were as follows: Maximum number of missed
717 cleavages (excluding the cross-linking site) = 2, peptide length = 5–50 aa, fixed
718 modifications = carbamidomethyl-Cys (mass shift = 57.021460 Da), variable
719 modification = oxidation of methionine (mass shift = 15.99491 Da), mass shift of the light
720 crosslinker = 138.0680796 Da, mass shift of mono-links = 156.0786442 and 155.0964278 Da,
721 MS1 tolerance = 10 ppm, MS2 tolerance = 0.2 Da for common ions and 0.3 Da for cross-link ions,
722 search in ion-tag mode. Post-search manual validation and filtering were performed using the
723 following criteria: xQuest score >25, FDR <0.05, mass error between − 2.2 and + 3.8 ppm,
724 %TIC >10, and a minimum peptide length of six aa. FDRs for the identified cross-links were
725 estimated using xProphet[69]. The five replicate data sets were compared, and cross-links in five out
726 of five data sets were used to generate a consensus data set. Normalized mono-links and cross-link
727 data with information on cross-linked residue positions and nseen (frequency) were visualized
728 using customized Matlab scripts. Hierarchical clustering was done using Matlab tools clustergram
729 and linkage to build linkage dendrogram trees across identified loop-links/cross-links for all 6
730 constructs.

731 **Tau aggregation in cells using mEOS biosensor cell lines**

732 FM5-CMV constructs of 4R P301S, P301S_Q, P301S_K, P301S_QK, and P301S EL_QK fused
733 to mEOS3.2 at the C-terminal were used for cell expression. The virus of each construct was
734 produced in Lenti-X™ 293 T Cell Line (Takara, Cat. #632180) and previously described[70].
735 Specifically, 300 ng PSP helper plasmid, 100 ng VSVG, and 100 ng of the plasmid of interest were
736 transiently co-transfected using 1.875 µL TransIT (Mirus Bio), and Opti-MEM (Gibco) were
737 mixed to a final volume of ~30 µL. Media (10% FBS, 1% Pen/Strep, 1% GlutaMax in Dulbecco's
738 modified Eagle's medium) containing the virus was collected after 48 h and centrifuged at 1000
739 RPM for 5 min to remove debris and dead cells. For transduction, 10-50 µL of virus suspension
740 was added to HEK293T (ATCC CRL-1268), and cells were grown in virus-containing media for
741 72 h before expanding to the most confluent condition. Cells were harvested from a 10 cm dish
742 with 0.05% trypsin, resuspended in flow cytometry buffer (HBSS plus 1% FBS and 1 mM EDTA),
743 and subjected to FACS (Sony Biotechnology). Populations with FITC-A signal above 10^4 were
744 collected for high-expressing cells. Following FACS and expansion, cells were maintained as a
745 polyclonal line. All stable cell lines were amplified, frozen, and stored in liquid nitrogen until use.

746 For seeding experiments, cells were plated in 96-well plates at 20,000 cells/well in 130 µL of
747 media. 24 h later, the cells were treated with 20 µL of fibril treatment with a dilution series in
748 technical triplicates. Briefly, treatments were composed of heparin-induced 2N4R recombinant tau
749 fibrils (8 µM monomer equivalent), sonicated for 5 minutes at an amplitude of 65 and 30 s off/on
750 a Q700 Sonicator (QSonica). Sonicated fibrils were diluted to the desired concentration at 10 µL,
751 mixed with 1 µL lipofectamine and 9 µL Opti-MEM, incubated at RT for 30 min per condition,
752 and replicated. After incubation, samples were added to cell media for final concentration in the
753 well of 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, and media control. After 48 hours, cells
754 were harvested with 0.05% trypsin and then fixed in 2% paraformaldehyde (Electron Microscopy
755 Services) for 10 min at room temperature. PFA was removed, and cells were resuspended in 1X
756 PBS. For the initial experiment, the optimal time of photoconversion was determined by exposing
757 cells to UV for 0-60 minutes and measuring the mEOS red-to-green ratio, which determined 20
758 minutes as the optimal Red/Green ratio used for all subsequent experiments.

759 **Flow cytometry**

760 A BD LSRFortessa was used to perform FRET flow cytometry. The mCerulean3 and FRET signal
761 was measured by exciting with the 405 nm laser, and fluorescence was captured with a 405/50 nm
762 and 525/50 nm filter, respectively. mClover3 signal was measured by exciting cells with a 488 nm
763 laser, and fluorescence was captured with a 525/50 nm filter. To quantify FRET, we used a gating
764 strategy where mCerulean3 bleed through into the mClover3, and FRET channels were
765 compensated using FlowJo analysis software, as described previously[71]. Subsequently, we created
766 a final bivariate plot of FRET vs. Cerulean and introduced a triangular gate to assess the number
767 of FRET-positive cells. This FRET gate was adjusted to biosensor cells that received lipofectamine
768 alone and are thus FRET-negative. FRET signal is defined as the percentage of FRET-positive
769 cells in all analyses. For each experiment, 10,000 mClover3/mCerulean3 double-positive cells per
770 replicate were analyzed, and each condition was analyzed in triplicate. Data analysis was
771 performed using FlowJo v10 software (Treestar).

772 For the mEOS3.2 cell expression system, tauRD-mEOS biosensor cells were first photoconverted
773 under UV for 20 min. 10,000 singlet events corresponding to the donor (non-photoconverted
774 mEOS3.2) and acceptor (photoconverted mEOS3.2) positive cells were collected for each sample.
775 Non-photoconverted mEOS3.2 was collected at the Alexa Flour 488 channel. FCS files were
776 exported from the BD FACSDiva data collection software and analyzed using FlowJo v10
777 software (Treestar). Compensation was manually applied to correct donor bleed-through into the
778 FRET channel guided by a sample with non-aggregated and photoconverted tauRD-mEOS.
779 Samples were gated on the acceptor intensity such that cells with similar concentrations of tauRD-
780 mEOS were analyzed to mitigate the contribution of differences in concentration leading to
781 apparent changes in the fraction of FRET-positive cells in each condition. FRET-positive cells
782 were quantified by gating double-positive singlet events with a ratio of FRET to donor signal
783 higher than that of a population of tauRD-mEOS photoconverted cells without aggregates.

784 **Microtubule Polymerization assay**

785 Experiments were carried out following the Cytoskeleton Tubulin Polymerization Fluorescence-
786 based Assay (BK011P) protocol. TauRD constructs in 20 mM MES and 10 mM DTT buffer were
787 mixed with 99.9% purified tubulin at 2 mg/ml in 80 mM PIPES, 2 mM MgCl, 0.5 mM EGTA, pH
788 6.9, 10 µM fluorescent reporter to a final concentration at 10 µM. 3 µM paclitaxel solution was
789 used as a positive control. All conditions were done in triplicates at 37 °C. ThT kinetic scans were

run every minute on a Tecan Spark plate reader at 350 nm Ex (10 nm bandwidth) and 435 nm Em (20 nm bandwidth). Blank wells containing buffer and tubulin were subtracted from experimental values. The data were min-max normalized to the sample with maximal signal (Taxol positive control), and the data was used for non-linear regression to determine the intensity max and the $t_{1/2}$max for assessing polymerization time. Column effects using 2way ANOVA were calculated for the intensity max and $t_{1/2}$max using GraphPad Prism v8.0-10.0.

MT:tau structure Analysis

The synthetic tau (R2X4) model bound to microtubules (MT) from PDB ID 6CVN was rebuilt to have the 4R tauRD designed sequence using the Rosetta threading protocol with bilateral loop building to include the missing residue from the structures[51]. For each sequence, 2,500 models of only tau without the MT were produced, constrained by the initial template of the tau structures. Models were then realigned with the α/β/α tubulin, and the tau chain was trimmed to contain only tau repeats 2 and 3 (275-331) with backbone atoms only (atoms NH, N, CA, CB, and C=O). Each structure was minimized using the Rosetta protocol with the `lbfgs_armijo_nonmonotone` algorithm and a minimum tolerance of 0.001 (https://www.rosettacommons.org/software). The resulting minimized structures were scored as a complex using the InterfaceAnalyzer protocol, pre-packing separate interfaces and tubulin chains comprising interface 1 and tau as interface 2. The resulting structures were compared to the lowest scoring structure to determine the built model RMSD using the Rosetta scoring function. All model building and analysis were performed on UTSW's BioHPC computing cluster. All plots were generated with GraphPad Prism. Images were created using Pymol.

Statistics and Reproducibility

All statistics were calculated using GraphPad Prism v8.0-10.0. Three independent ThT experiments were run for each condition. The data were normalized to the highest amplitude, and averages and standard deviations were plotted. Plots were fitted to a non-linear regression model, from which $t_{1/2}$ values were derived. $t_{1/2}$ error represents a 95% CI. TEM grids of endpoint ThT samples were screeded and imaged at least 4 times to obtain representative images. Flow cytometry cell aggregation was conducted in three independent experiments, whose values are plotted. Error bars represent 95% CI.

Competing Interests

The authors declare that they have no competing interests.

Data availability

All MD, NMR (peptide and tauRD), ThT, XL-MS, MT stabilization, MT: tau modeling, and cell-based aggregation data are available as Source Data 1, Source Data 2, Source Data 3 and 4, Source Data 5, Source Data 6, Source Data 7, and Source Data 8, respectively. Source Data are also available on zenodo under accession number 13821565 [https://doi.org/10.5281/zenodo.13821565]. PDB IDs used in the study include 2MZ7 [https://doi.org/10.2210/pdb2MZ7/pdb], 6CVN [https://doi.org/10.2210/pdb6CVN/pdb], and 8Q92 [https://doi.org/10.2210/pdb8q92/pdb]. Supplementary Data for raw MD trajectory files, structure files for MSM modeling and validation file, and MT:tau modeling is available on zenodo under accession number 13821565 [https://doi.org/10.5281/zenodo.13821565]. All mass spectrometry data has been deposited in MassIVE under accession number MSV000093880 and is assigned to ProteomeXchange under accession number PXD048689.

Code availability

All MD simulations were performed and analyzed using Gromacs v5.0.4 (available at http://www.gromacs.org). Markov State Model analysis was performed using pyEMMA v2.5.7 (http://emma-project.org). NMR spectra were processed, assigned, and visualized using NMRpipe (https://www.ibbr.umd.edu/nmrpipe/install.html), TopSpin (https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html), Cara v1.9.1 (http://cara.nmr.ch/doku.php), ccpNMR (https://ccpn.ac.uk/software/), and Sparky 3 (https://www.cgl.ucsf.edu/home/sparky/). XL-MS data was analyzed using xQuest v2.1.5 (https://gitlab.ethz.ch/leitner_lab/xquest_xprophet). All MT:tau simulations were performed using Rosetta v3.12 (available at https://www.rosettacommons.org).

Uncategorized References

1 Vaquer-Alicea, J., Diamond, M. I. & Joachimiak, L. A. Tau strains shape disease. *Acta Neuropathol* **142**, 57-71 (2021). https://doi.org/10.1007/s00401-021-02301-7

2 Gao, Y.-L. *et al.* Tau in neurodegenerative disease. *Annals of Translational Medicine* **6** (2018/05). https://doi.org/10.21037/atm.2018.04.23

3 Orr, M. E., Sullivan, A. C. & Frost, B. A Brief Overview of Tauopathy: Causes, Consequences, and Therapeutic Strategies. *Trends in Pharmacological Sciences* **38**, 637-648 (2017). https://doi.org/10.1016/j.tips.2017.03.011

4 Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185-190 (2017). https://doi.org/10.1038/nature23002

5 Falcon, B. *et al.* Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420-423 (2019). https://doi.org/10.1038/s41586-019-1026-5

6 Lövestam S, K. F., van Knippenberg B, Kotecha A, Murzin AG, Goedert M, Scheres SHW. . Assembly of recombinant tau into filaments identical to those of Alzheimer's disease and chronic traumatic encephalopathy. *eLife* **11** (2022). https://doi.org/10.7554/eLife.76494

7 Han, Z. Z., Kang, S. G., Arce, L. & Westaway, D. Prion-like strain effects in tauopathies. *Cell Tissue Res* **392**, 179-199 (2023). https://doi.org/10.1007/s00441-022-03620-1

8 Kaufman, S. K., Thomas, T. L., Del Tredici, K., Braak, H. & Diamond, M. I. Characterization of tau prion seeding activity and strains from formaldehyde-fixed tissue. *Acta Neuropathol Commun* **5**, 41 (2017). https://doi.org/10.1186/s40478-017-0442-8

9 Mirbaha, H. *et al.* Inert and seed-competent tau monomers suggest structural origins of aggregation. *Elife* **7** (2018). https://doi.org/10.7554/eLife.36584

10 Mirbaha, H. *et al.* Seed-competent tau monomer initiates pathology in a tauopathy mouse model. *J Biol Chem* **298**, 102163 (2022). https://doi.org/10.1016/j.jbc.2022.102163
11 Sharma, A. M., Thomas, T. L., Woodard, D. R., Kashmer, O. M. & Diamond, M. I. Tau monomer encodes strains. *eLife* **7** (2018). https://doi.org/10.7554/elife.37813
12 Hou, Z., Chen, D., Ryder, B. D. & Joachimiak, L. A. Biophysical properties of a tau seed. *Sci Rep* **11**, 13602 (2021). https://doi.org/10.1038/s41598-021-93093-z
13 Goedert, M. & Spillantini, M. G. Tau mutations in frontotemporal dementia FTDP-17 and their relevance for Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1502**, 110-121 (2000). https://doi.org/10.1016/s0925-4439(00)00037-5
14 van Swieten, J. & Spillantini, M. G. Hereditary frontotemporal dementia caused by Tau gene mutations. *Brain Pathol* **17**, 63-73 (2007). https://doi.org/10.1111/j.1750-3639.2007.00052.x
15 Kinoshita, J. & Clark, T. Alzforum. *Methods Mol Biol* **401**, 365-381 (2007). https://doi.org/10.1007/978-1-59745-520-6_19
16 Bugiani, O. *et al.* Frontotemporal dementia and corticobasal degeneration in a family with a P301S mutation in tau. *J Neuropathol Exp Neurol* **58**, 667-677 (1999). https://doi.org/10.1097/00005072-199906000-00011
17 Lee, G., Neve, R. L. & Kosik, K. S. The microtubule binding domain of tau protein. *Neuron* **2**, 1615-1624 (1989). https://doi.org/10.1016/0896-6273(89)90050-0
18 Liu, F. & Gong, C.-X. Tau exon 10 alternative splicing and tauopathies. *Molecular Neurodegeneration* **3** (2008). https://doi.org/10.1186/1750-1326-3-8
19 Schoch, K. M. *et al.* Increased 4R-Tau Induces Pathological Changes in a Human-Tau Mouse Model. *Neuron* **90**, 941-947 (2016). https://doi.org/10.1016/j.neuron.2016.04.042
20 Limorenko, G. & Lashuel, H. A. To target Tau pathologies, we must embrace and reconstruct their complexities. *Neurobiol Dis* **161**, 105536 (2021). https://doi.org/10.1016/j.nbd.2021.105536
21 Mummery, C. J. *et al.* Tau-targeting antisense oligonucleotide MAPT(Rx) in mild Alzheimer's disease: a phase 1b, randomized, placebo-controlled trial. *Nat Med* **29**, 1437-1447 (2023). https://doi.org/10.1038/s41591-023-02326-3
22 Yoshiyama, Y. *et al.* Synapse Loss and Microglial Activation Precede Tangles in a P301S Tauopathy Mouse Model. *Neuron* **53** (2007/02/01). https://doi.org/10.1016/j.neuron.2007.01.010
23 Holmes, B. B. *et al.* Proteopathic tau seeding predicts tauopathy in vivo. *Proc Natl Acad Sci U S A* **111**, E4376-4385 (2014). https://doi.org/10.1073/pnas.1411649111
24 Macdonald, J. A. *et al.* Assembly of transgenic human P301S Tau is necessary for neurodegeneration in murine spinal cord. *Acta Neuropathologica Communications 2019 7:1* **7** (2019-03-18). https://doi.org/10.1186/s40478-019-0695-5
25 von Bergen, M. *et al.* Mutations of tau protein in frontotemporal dementia promote aggregation of paired helical filaments by enhancing local beta-structure. *J Biol Chem* **276**, 48165-48174 (2001). https://doi.org/10.1074/jbc.M105196200
26 Sahayaraj, A. E., Viswanathan, R., Pinhero, F., Abdul Vahid, A. & Vijayan, V. Sequence-Dependent Conformational Properties of PGGG Motif in Tau Repeats: Insights from Molecular Dynamics Simulations of Narrow Pick Filament. *ACS Chem Neurosci* **14**, 136-147 (2023). https://doi.org/10.1021/acschemneuro.2c00602
27 Mullapudi, V. *et al.* Network of hotspot interactions cluster tau amyloid folds. *Nat Commun* **14**, 895 (2023). https://doi.org/10.1038/s41467-023-36572-3
28 van der Kant, R., Louros, N., Schymkowitz, J. & Rousseau, F. Thermodynamic analysis of amyloid fibril structures reveals a common framework for stability in amyloid polymorphs. *Structure* **30**, 1178-1189 e1173 (2022). https://doi.org/10.1016/j.str.2022.05.002

29 Lövestam, S., Li, D., Wagstaff, J. L., Kotecha, A., Kimanius, D., McLaughlin, S. H., Murzin, A. G., Freund, S. M., Goedert, M., & Scheres, S. H. Disease-specific tau filaments assemble via polymorphic intermediates. *Nature* (2023). https://doi.org/10.1038/s41586-023-06788-w

30 Chen, D. *et al.* Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nat Commun* **10**, 2493 (2019). https://doi.org/10.1038/s41467-019-10355-1

31 Stelzl, L. S. *et al.* Global Structure of the Intrinsically Disordered Protein Tau Emerges from Its Local Structure. *JACS Au* **2**, 673-686 (2022). https://doi.org/10.1021/jacsau.1c00536

32 Li L, N. B., Mullapudi V, Li Y, Saelices L, Joachimiak LA. . Disease-associated patterns of acetylation stabilize tau fibril formation. *Structure (London, England : 1993)* **31** (2023). https://doi.org/10.1016/j.str.2023.05.020

33 Chakraborty, P. *et al.* Acetylation discriminates disease-specific tau deposition. *Nat Commun* **14**, 5919 (2023). https://doi.org/10.1038/s41467-023-41672-1

34 Arakhamia, T. *et al.* Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell* **180**, 633-644 e612 (2020). https://doi.org/10.1016/j.cell.2020.01.027

35 Wesseling, H. *et al.* Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's Disease. *Cell* **183**, 1699-1713 e1613 (2020). https://doi.org/10.1016/j.cell.2020.10.029

36 Tamiola, K. & Mulder, F. A. Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. *Biochem Soc Trans* **40**, 1014-1020 (2012). https://doi.org/10.1042/BST20120171

37 Tamiola, K., Acar, B. & Mulder, F. A. A. Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. *Journal of the American Chemical Society* **132** (December 3, 2010). https://doi.org/10.1021/ja105656t

38 Robustelli, P., Piana, S. & Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **115**, E4758-E4766 (2018). https://doi.org/10.1073/pnas.1800690115

39 Jephthah, S., Pesce, F., Lindorff-Larsen, K. & Skepo, M. Force Field Effects in Simulations of Flexible Peptides with Varying Polyproline II Propensity. *J Chem Theory Comput* **17**, 6634-6646 (2021). https://doi.org/10.1021/acs.jctc.1c00408

40 Scherer, M. K. *et al.* PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J Chem Theory Comput* **11**, 5525-5542 (2015). https://doi.org/10.1021/acs.jctc.5b00743

41 Falcon, B. *et al.* Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137-140 (2018). https://doi.org/10.1038/s41586-018-0454-y

42 Leitner, A. *et al.* The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* **20**, 814-825 (2012). https://doi.org/10.1016/j.str.2012.03.007

43 Leitner, A. *et al.* Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc Natl Acad Sci U S A* **111**, 9455-9460 (2014). https://doi.org/10.1073/pnas.1320298111

44 Lasker, K. *et al.* Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc Natl Acad Sci U S A* **109**, 1380-1387 (2012). https://doi.org/10.1073/pnas.1120559109

45 Chlebowicz, J. *et al.* Saturation mutagenesis of alpha-synuclein reveals monomer fold that modulates aggregation. *Sci Adv* **9**, eadh3457 (2023). https://doi.org/10.1126/sciadv.adh3457

46 Rinner, O. *et al.* Identification of cross-linked peptides from large sequence databases. *Nature Methods 2008 5:4* **5** (2008-03-09). https://doi.org/10.1038/nmeth.1192

47 Farrow, N. A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984-6003 (1994). https://doi.org/10.1021/bi00185a040

48 Barré P, E. D. Structural transitions in tau k18 on micelle binding suggest a hierarchy in the efficacy of individual microtubule-binding repeats in filament nucleation. *Protein science : a publication of the Protein Society* **22** (2013). https://doi.org/10.1002/pro.2290

49 Lathuilière, A. *et al.* Motifs in the tau protein that control binding to microtubules and aggregation determine pathological effects. *Scientific Reports* **7** (2017). https://doi.org/10.1038/s41598-017-13786-2

50 Xia, Y. *et al.* Impaired tau–microtubule interactions are prevalent among pathogenic tau variants arising from missense mutations. *The Journal of Biological Chemistry* **294** (2019). https://doi.org/10.1074/jbc.RA119.010178

51 Kellogg, E. H. *et al.* Near-atomic model of microtubule-tau interactions. *Science* **360**, 1242-1246 (2018). https://doi.org/10.1126/science.aat1780

52 Kadavath, H. *et al.* Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proc Natl Acad Sci U S A* **112**, 7501-7506 (2015). https://doi.org/10.1073/pnas.1504081112

53 Sündermann, F., Fernandez, M.-P. & Morgan, R. O. An evolutionary roadmap to the microtubule-associated protein MAP Tau. *BMC Genomics* **17** (2016). https://doi.org/10.1186/s12864-016-2590-9

54 Gerasimavicius, L., Livesey, B. J. & Marsh, J. A. Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure. *Nature Communications* **13** (2022). https://doi.org/10.1038/s41467-022-31686-6

55 Bali, S. J., L. Modifying Amyloid Motif Aggregation Through Local Structure - PubMed. *Methods in molecular biology (Clifton, N.J.)* **2340** (2022). https://doi.org/10.1007/978-1-0716-1546-1_15

56 Hinds, M. G. & Norton, R. S. NMR Spectroscopy of Peptides and Proteins. (1994). https://doi.org/10.1385/0-89603-274-4:131

57 Delaglio F, G. S., Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **6** (1995). https://doi.org/10.1007/BF00197809

58 Lee W, T. M., Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics (Oxford, England)* **31** (2015). https://doi.org/10.1093/bioinformatics/btu830

59 Tien, M. Z., Meyer, A. G., Sydykova, D. K., Spielman, S. J. & Wilke, C. O. Maximum allowed solvent accessibilites of residues in proteins. *PLoS One* **8**, e80635 (2013). https://doi.org/10.1371/journal.pone.0080635

60 Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958 (2010). https://doi.org/10.1002/prot.22711

61 Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**, 3684-3690 (1984). https://doi.org/10.1063/1.448118

62 Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190 (1981). https://doi.org/10.1063/1.328693

63 Feenstra, K. A. & Hess, B. B., H. J. C. . Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *Journal of Computational Chemistry* (1999).

64 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98** (1993/06/15). https://doi.org/10.1063/1.464397

65 Skinner, S. P. *et al.* CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR* **66**, 111-124 (2016). https://doi.org/10.1007/s10858-016-0060-y

1009 66 Barre, P. & Eliezer, D. Structural transitions in tau k18 on micelle binding suggest a hierarchy in
1010 the efficacy of individual microtubule-binding repeats in filament nucleation. *Protein Sci* **22**,
1011 1037-1048 (2013). https://doi.org/10.1002/pro.2290
1012 67 Narayanan, R. L. *et al.* Automatic assignment of the intrinsically disordered protein Tau with
1013 441-residues. *J Am Chem Soc* **132**, 11906-11907 (2010). https://doi.org/10.1021/ja105657f
1014 68 Barre, P. & Eliezer, D. Folding of the repeat domain of tau upon binding to lipid surfaces. *J Mol*
1015 *Biol* **362**, 312-326 (2006). https://doi.org/10.1016/j.jmb.2006.07.018
1016 69 Walzthoeni, T. *et al.* False discovery rate estimation for cross-linked peptides identified by mass
1017 spectrometry. *Nat Methods* **9**, 901-903 (2012). https://doi.org/10.1038/nmeth.2103
1018 70 Stopschinski, B. E. *et al.* Specific glycosaminoglycan chain length and sulfation patterns are
1019 required for cell uptake of tau versus alpha-synuclein and beta-amyloid aggregates. *J Biol Chem*
1020 **293**, 10826-10840 (2018). https://doi.org/10.1074/jbc.RA117.000378
1021 71 Furman, J. L., Holmes, B. B. & Diamond, M. I. Sensitive Detection of Proteopathic Seeding
1022 Activity with FRET Flow Cytometry. *J Vis Exp*, e53205 (2015). https://doi.org/10.3791/53205

1023 **Acknowledgments**

1024 L.A.J. is supported by an Effie Marie Cain Scholarship in Medical Research and by grants from
1025 the NIH (R01AG076459), the Chan Zuckerberg Initiative (CZI) Collaborative Science Award
1026 (2018-191983), and the Welch Foundation (I-1928-20200401 to L.A.J and I-1304 to J.R). N.E.vN.
1027 was supported by TargetALS NI-2023-NAI-S6 S.B. is supported by an NIH F31 grant from the
1028 NINDS (F31NS12751301). NMR data was acquired at the Biomolecular NMR Facility at UTSW.
1029 Heteronuclear NOE experiments were performed in the "Manuel Rico" NMR Laboratory (LMR)
1030 of the Spanish National Research Council (CSIC), a node of the Spanish Large-Scale National
1031 Facility (ICTS R-LRB). Transmission electron microscopy was performed at the Electron
1032 Microscopy Core Facility at UTSW, supported by the NIH (1S10OD021685-01A1 and
1033 1S10OD020103-01). Computational resources were provided by the BioHPC cluster supported by
1034 the Lyda Hill Department of Bioinformatics at UTSW. We would like to thank Andrew Lemoff
1035 from the UTSW Proteomics Core for his valuable insights on troubleshooting procedures in
1036 crosslinking mass spectrometry. We would like to thank Qiong Wu from the UTSW BioNMR
1037 facility for her help in the optimization and collection of NMR data. We would also like to thank
1038 Javier Oroz from the IQF-CSIC for his guidance in the hetNOE relaxation experiments. We thank
1039 all members of the Joachimiak lab, particularly Dailu Chen, Pawel M. Wydorski, and Valerie A.
1040 Perez, and members of the UTSW CAND, specifically Jaime Vaquer-Alicea, for their discussions
1041 and input on the manuscript.

1042 **Author Contributions**

1043 S.B. and L.A.J. initiated the project. S.B. and A.W. performed MD calculations. S.B. performed
1044 NMR experiments with guidance from J.R. S.B., and N.E.V. performed NMR relaxation
1045 experiments. S.B. and R.S. performed MSM analysis. S.B. purified all proteins involved in the
1046 study. S.B. performed all peptide and protein aggregation experiments. S.B. performed TEM of
1047 tau fibrils. S.B. and D. C. performed full-length fibril formation. S.B. and P.M.W. performed the
1048 XL-MS analysis. S.B., V.A.P. and A.W. carried out the cell-based aggregation experiments. S.B.
1049 performed the Rosetta MT:tau structure calculations. Finally, S.B. and L.A.J. conceived of and
1050 directed the research and wrote the manuscript. All authors contributed to the revisions of the
1051 manuscript.