nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-022-30585-0 OPEN

Fluent molecular mixing of Tau isoforms in Alzheimer's disease neurofibrillary tangles

Aurelio J. Dregni[1,3], Pu Duan[1,3], Hong Xu[2], Lakshmi Changolkar[2], Nadia El Mammeri[1], Virginia M.-Y. Lee[2] & Mei Hong[1✉]

Alzheimer's disease (AD) is defined by intracellular neurofibrillary tangles formed by the microtubule-associated protein tau and extracellular plaques formed by the β-amyloid peptide. AD tau tangles contain a mixture of tau isoforms with either four (4R) or three (3R) microtubule-binding repeats. Here we use solid-state NMR to determine how 4R and 3R tau isoforms mix at the molecular level in AD tau aggregates. By seeding differentially isotopically labeled 4R and 3R tau monomers with AD brain-derived tau, we measured intermolecular contacts of the two isoforms. The NMR data indicate that 4R and 3R tau are well mixed in the AD-tau seeded fibrils, with a 60:40 incorporation ratio of 4R to 3R tau and a small homotypic preference. The AD-tau templated 4R tau, 3R tau, and mixed 4R and 3R tau fibrils exhibit no structural differences in the rigid β-sheet core or the mobile domains. Therefore, 4R and 3R tau are fluently recruited into the pathological fold of AD tau aggregates, which may explain the predominance of AD among neurodegenerative disorders.

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139, USA. [2] Department of Pathology and Laboratory Medicine, Institute on Aging and Center for Neurodegenerative Disease Research, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA. [3] These authors contributed equally: Aurelio J. Dregni, Pu Duan. ✉email: meihong@mit.edu

In the human brain, the microtubule-associated protein tau exists in six isoforms, which are mainly distinguished by the presence or absence of the second microtubule-binding repeat, R2 (ref. [1]). Tau isoforms that contain four microtubule-binding repeats (R1, R2, R3, and R4) are called four-repeat (4R) tau, whereas those that contain three microtubule-binding repeats (R1, R3, and R4) are called three-repeat (3R) tau. In a number of neurodegenerative diseases, these intrinsically disordered tau proteins become hyperphosphorylated and form intraneuronal aggregates. Alzheimer's disease (AD) is defined by these tau neurofibrillary tangles as well as the extracellular plaques formed by the β-amyloid peptide[2,3]. Interestingly, the AD tau tangles contain a mixture of 4R and 3R tau isoforms, whereas other tauopathies such as corticobasal degeneration and Pick's disease are characterized by either aggregated 4R tau or aggregated 3R tau. Recent cryo-electron microscopy (cryoEM) studies[4–6] revealed that there is a single well-defined molecular conformation for the β-sheet core of AD tau filaments, but did not determine how 4R tau and 3R tau are mixed in the aggregate. The rigid core of the AD tau filaments spans the R3, R4, and part of the R′ repeats, which are present in both 4R and 3R tau isoforms. Another mixed-tau disease, chronic traumatic encephalopathy, also exhibits a single rigid core structure for its tau filaments[7], with no discernible differentiation between 4R and 3R tau in the fibril. Understanding how different tau isoforms are mixed in these filaments is important, because extensive biochemical data have established that misfolded tau proteins in AD and other tauopathies spread by recruiting soluble tau protein into the aggregate, and distinct isoforms might be expected to interfere with the propagation of the amyloid structure[8–17]. To determine how 4R and 3R tau isoforms are mixed in the neurofibrillary tangles in AD brains, here we use solid-state nuclear magnetic resonance (NMR) spectroscopy to investigate AD-tau seeded ^{15}N-labeled and ^{13}C-labeled tau proteins. We show that the 4R and 3R tau isoforms are fluently mixed in AD tau filaments, with no detectable structural differences among pure 4R tau fibrils, pure 3R tau fibrils, and mixed 4R and 3R tau fibrils.

Results

To determine the mode of mixing of 4R and 3R tau in AD tau aggregates, we seeded ^{15}N or ^{13}C-labeled recombinant tau monomers with 10% AD brain-derived sarkosyl-insoluble tau (AD tau) (Fig. 1a) and characterized the resulting fibrils using solid-state NMR. AD tau was pooled from ten neuropathologically confirmed diagnosis of AD (Supplementary Tables 1 and 2). The use of multiple AD brains ensures that the molecular structural results we obtain reflect the average properties of AD brain tau filaments instead of the property of one AD patient. We chose brains with short postmortem intervals and high tau-burdens at late stages of the disease[18,19]. AD tau seeds were sequentially extracted from these brains. To these AD tau extracts we added ninefold excess of either a 1:1 mixture of 0N4R and 0N3R tau monomers or one of the two monomers. After 7 days of incubation under shaking, about half of the added monomers became insoluble (Fig. 1b and Supplementary Fig. 1), indicating a fivefold increase in insoluble mass[18,19]. Unseeded recombinant tau remained largely soluble. Transmission electron microscopy (TEM) images reveal that the monomer-added samples formed much longer filaments (200–300 nm) than the original AD tau seeds (40–70 nm) (Fig. 2 and Supplementary Fig. 2a). Most of these filaments have a twisted morphology whereby an ~10 nm width alternates with an ~25 nm width. A small number of relatively straight fibrils with ~16-nm width is also observed (Supplementary Table 3). These two morphologies are similar to the paired helical filaments (PHF) and straight filaments (SF) in AD tau[20], but differ from the long, straight and 14–23 nm wide fibrils assembled using heparin[21–23] (Fig. 2c and Supplementary Fig. 2b). The AD-tau templated fibrils can seed mouse tau in wild-type primary neurons into morphologically identical aggregates, as detected by immunofluorescence using a mouse-tau specific antibody (Fig. 1c). Recent studies showed that these AD-tau-seeded mouse tau aggregates are localized to dendrites and axons[19,24]. Quantification of the fluorescence intensities shows that the amount of aggregated mouse tau increased in the presence of templated fibrils compared to the 10% AD-tau control (Fig. 1d). Therefore, the templated recombinant tau fibrils are pathologically similar to AD tau; moreover, the templated fibrils have higher overall pathogenicity compared the original 10% AD tau seeds. A similar increase in pathogenicity were observed whether 4R tau alone, 3R tau alone, or a mixture of both isoforms were added to the AD tau seeds[25]. Together, these data indicate that the ultrastructural morphology and pathological activities of AD tau are faithfully propagated by recombinant tau, regardless of whether 4R tau, 3R tau or both isoforms were added.

To determine how 4R and 3R tau monomers are organized in AD-tau templated fibrils, we conducted a ^{15}N–^{13}C rotational-echo-double-resonance (REDOR) NMR experiment[26] (Fig. 3). This experiment measures the proportion of ^{13}C-labeled tau that is one strand away from an ^{15}N-labeled tau in the hydrogen-bonded cross-β fibril. Because two hydrogen-bonded β-strand backbones are separated by ~4.8 Å, the distance-dependent intermolecular ^{15}N–^{13}C dipolar coupling depends only on how two differentially labeled tau monomers are mixed (Fig. 3a, b). ^{13}C dipolar dephasing of an ^{15}N-labeled tau is manifested as intensity loss between a spectrum measured with the ^{13}C pulses off (S_0) and the ^{13}C pulses on (S). Greater ^{15}N–^{13}C dipolar dephasing, or lower S/S_0 values, indicates a higher probability of ^{13}C-labeled tau following an ^{15}N-labeled tau. To increase the spectral sensitivity, we observed this ^{15}N intensity loss through the signals of the directly bonded amide proton. No ^{1}H spectral resolution is required, as the REDOR experiment detects all rigid amides in the rigid β-sheet fibril core (Fig. 3c, d). To simplify the ^{1}H spectra and facilitate the comparison of 2D ^{15}N–^{1}H correlation spectra, we perdeuterated the protein monomers and back-exchanged them with water. The resulting recombinant tau contains only NH and OH protons.

We mixed 50% ^{15}N-labeled 4R tau and 50% ^{13}C-labeled 3R tau in two independently prepared AD-tau seeded samples ([AD] $4R^N3R^C$) (Supplementary Table 4). A second pair of AD-tau seeded samples has the reverse mixed labeling motif, containing 50% ^{15}N-labeled 3R tau and 50% ^{13}C-labeled 4R tau ([AD] $3R^N4R^C$). We first measured the REDOR dephasing (S/S_0) of heparin-fibrillized 0N4R tau samples[21,22] containing varying molar ratios of ^{15}N-labeled and ^{13}C-labeled monomers. These samples allow direct experimental calibration of the REDOR dephasing of different statistical mixtures of ^{15}N-labeled protein and ^{13}C-labeled protein. For these 0N4R tau samples, the probability of a ^{13}C-labeled monomer following an ^{15}N-labeled monomer ($p_{N\rightarrow C}$) in the fibril is equal to the fraction of ^{13}C-labeled monomers in the mixture. The fact that in vitro-fibrilized tau and AD-seeded tau do not have the same rigid core structure does not affect the use of intermolecular REDOR for determining the statistics of mixing, because the REDOR experiment only depends on the existence of parallel-in-register cross-β packing, which is true for both heparin-fibrillized tau and AD tau.

Figure 3e shows the REDOR dephasing of mixed ^{13}C-labeled and ^{15}N-labeled 0N4R tau samples, whose ^{13}C:^{15}N molar ratios range from 0:100% to 70%:30%. The dephasing curves show the expected trend: the higher the fraction of ^{13}C-labeled tau, the faster the dipolar dephasing, and the lower the S/S_0 values.

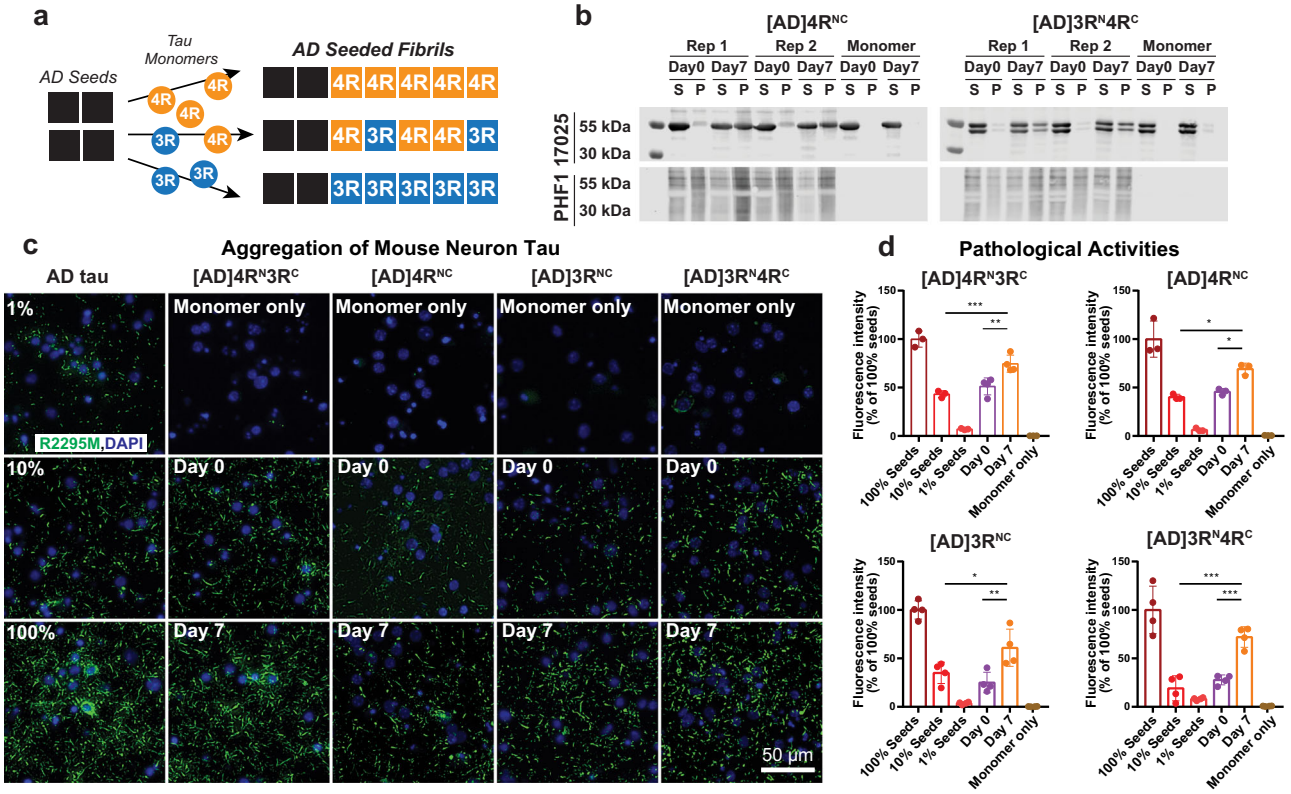


Fig. 1 Isotopically labeled recombinant tau seeded with AD tau filaments conserves the AD pathogenic activity in wild-type neurons. a Schematic of seeding recombinant 0N4R and 0N3R tau with AD tau extracts. **b** Sedimentation assay shows that recombinant tau monomers are incorporated into insoluble mass after 7 days of shaking with 10% AD tau seeds. Samples were fractionated into supernatant (S) and pellet (P) by ultracentrifugation. Tau species were revealed by the 17,025 antitotal tau antibody and PHF1 anti-phospho-tau antibody. Aliquots of replicate (Rep) samples 1 and 2 were collected for biochemical characterization at day 0 and day 7 (the end) of the fibrillization reaction. Reactions without AD tau seeds (monomer-only) were used as control to assess the extent of self-fibrillization. Full blots can be found in Supplementary Fig. 1. **c** Immunocytochemistry shows AD-tau-seeded recombinant tau fibrils recapitulate the pathogenicity of AD tau in primary neurons. Wild-type primary hippocampal neurons were transduced with AD tau or seeded tau fibrils. Aggregated tau pathology was revealed by R2295M anti-mouse tau antibody (green) and DAPI (blue) that stains cell nuclei. Seeded samples are AD-tau-seeded fibrils containing a 1:1 mixture of ^{15}N-labeled 4R tau and ^{13}C-labeled 3R tau, 4R-only tau, 3R-only tau, and a 1:1 mixture of ^{15}N-labeled 3R tau and ^{13}C-labeled 4R tau. **d** Quantification of fluorescence in **c** shows that the seeded fibrils at day 7 have increased pathogenicity compared to day 0. As expected, day 0 samples (10% AD tau seeds with monomers) showed similar pathogenicity levels to the 10% tau seeds without monomers. All values were normalized to the 100% AD tau seeds. Statistical analysis was performed using one-way ANOVA. Multiple comparison between groups was conducted using Tukey post hoc test. On the graph, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ in the multiple comparison tests. Error bars are presented as standard deviation (SD). Each data point stands for one biological repeat: [AD]4R^{N}3R^{C} n=3, [AD]4R^{NC} n=3, [AD]3R^{NC} n=4, [AD]3R^{N}4R^{C} n=4.

The measured REDOR decay curves can be simulated using the known intermolecular distances in a cross-β amyloid fibril (Supplementary Fig. 3). To do this, we conducted both explicit spin simulations[27] and second-moment analysis[28]. Both methods of simulation reproduce the trend of the mixing-ratio-dependent dipolar dephasing, but do not match quantitatively the measured REDOR dephasing of the calibration samples. This is not surprising, because each simulation method has shortcomings that limit its accuracy (see "Methods"). The explicit spin simulation is restricted to a maximum of about 10 spins, while the second-moment analysis neglects nuclear spin interactions other than ^{15}N–^{13}C dipolar coupling. Moreover, neither simulation captures the effects of natural abundance ^{13}C spins in the ^{15}N-labeled monomer. For these reasons, we rely on the more accurate experimental REDOR results of the calibration samples to extract the mixing probabilities of AD tau filaments.

The spectra of the [AD]4R^{N}3R^{C} samples yielded a REDOR dephasing curve that falls between the calibration curves for 30:70 and 40:60 ^{13}C:^{15}N labeled 0N4R tau (Fig. 3e). Interpolation of the measured S/S_0 values between the two calibration curves yielded a $p_{4\to3}$ value of 0.37 ± 0.03 (Supplementary Figs. 3 and 4). Thus, in AD-seeded tau filaments, a 4R tau is followed by a 3R tau with 37% probability, and hence by a 4R tau with 63% probability. The reversely mixed labeled [AD]3R^{N}4R^{C} samples showed faster REDOR dephasing: the measured REDOR curve falls between the calibration curves for 50:50 and 60:40 ^{13}C:^{15}N 0N4R tau. Interpolation of the two measured calibration curves to match the REDOR dephasing of the AD-seeded sample yielded a $p_{3\to4}$ value of 0.56 ± 0.02. Thus, in the AD-seeded tau filaments, a 3R tau is followed by a 4R tau with 56% probability, and hence by a 3R tau with 44% probability. These data together indicate that each 4R tau monomer is flanked by 63% 4R tau, while each 3R tau monomer is flanked by 56% 4R tau. Therefore, the AD tau filaments, when offered equal amounts of 4R tau and 3R tau monomers, neither completely separate the two isoforms, which would correspond to $p_{4\to3} = p_{3\to4} = 0$, nor strictly alternate the two isoforms, which would correspond to $p_{4\to3} = p_{3\to4} = 1$. Moreover, both 4R tau and 3R tau have a small preference for contacting 4R tau. A two-state Markov model (see "Methods")

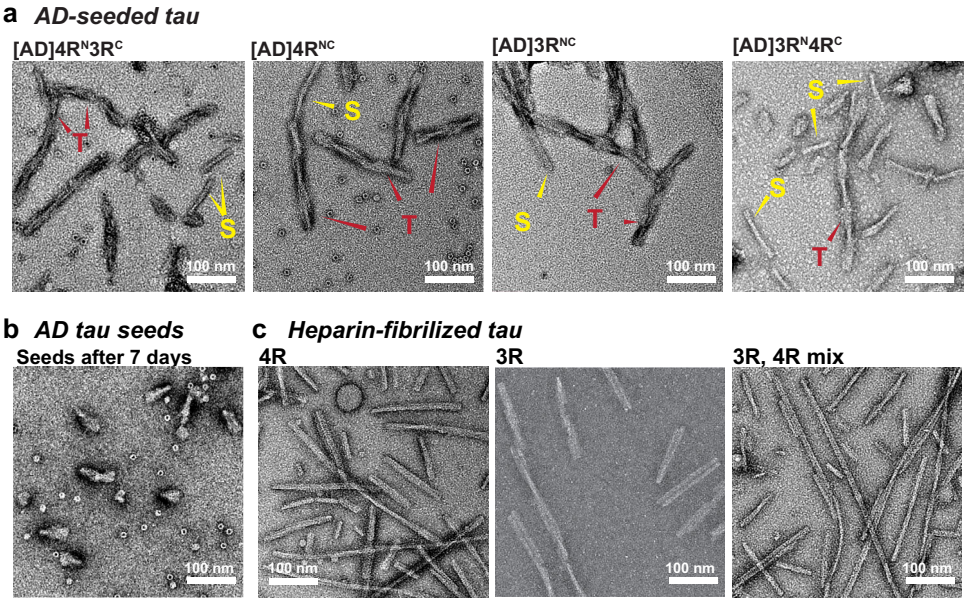


Fig. 2 Ultrastructural morphology of AD-tau seeded tau fibrils is distinct from that of heparin-fibrillized tau. a TEM images of AD-tau seeded fibrils containing a 1:1 mixture of ^{15}N-labeled 4R tau and ^{13}C-labeled 3R tau, 4R-only tau, 3R-only tau, and a 1:1 mixture of ^{15}N-labeled 3R tau and ^{13}C-labeled 4R tau. All samples show predominantly twisted (T) fibrils whose widths alternate between ~10 nm and ~25 nm, similar to the paired helical filaments of AD tau. In addition, a minority of straight (S) fibrils are observed, which are similar to the straight filaments of AD tau. **b** AD tau seeds after 7 days of incubation without added recombinant monomers. The seeds are fivefold shorter than the amplified tau fibrils. No long fibrils are observed. **c** Heparin-assembled recombinant tau fibrils. These fibrils are nearly exclusively straight, and are longer and more uniform than the straight filaments in AD-tau seeded filaments. At least five different images from different regions of each grid were obtained for each sample. A side-by-side comparison of these fibrils can be found in Supplementary Fig. 2.

shows that these probabilities can be converted to the mole fraction of 4R tau in the filament according to $\chi_4 = p_{3\to4}/(p_{3\to4} + p_{4\to3})$, giving a 4R tau mole fraction of 60 ± 2%. Thus, the AD tau filaments incorporate the two isoforms with a 60: 40 molar ratio in favor of 4R tau. This composition is consistent with the estimated 4R and 3R tau incorporation levels from the ^{15}N and ^{13}C spectral intensities (Supplementary Fig. 5 and Supplementary Table 4). Furthermore, because $p_{4\to4}$ (63%) is larger than the overall 4R tau incorporation level (60%) while $p_{3\to4}$ (56%) is smaller than it, there is a small preference for homotypic 4R–4R and 3R–3R contacts over heterotypic 3R–4R contacts. We can alternatively express this preference for homotypic contact in terms of a mixing quotient, $Q = p_{3\to4}p_{4\to3}/p_{3\to3}p_{4\to4}$, defined in analogy to the equilibrium constant of mixing. Large Q values indicate a preference for heterotypic mixing while small Q values indicate a preference for block copolymerization. When there is no homotypic or heterotypic preference, i.e., when the mixing is agnostic of the present isoform, then $p_{3\to4} = p_{4\to4} = \chi_4$, and $Q = 1$. The measured $p_{4\to3}$ and $p_{3\to4}$ values correspond to a Q value of 0.75 ± 0.11, indicating a small preference for homotypic contacts, independent of the overall isoform incorporation level. To provide a molecular view of this isoform mixing, we replicated the cryoEM structure of AD PHF tau[4] into a long filament containing 360 monomers in each protofilament, and simulated the 4R and 3R tau alternation to match the measured $p_{4\to3}$ and $p_{3\to4}$ values (Fig. 3f, g). The simulation depicts both the larger abundance of 4R tau over 3R tau in the filament and the intimate and fluent molecular mixing of the two tau isoforms.

To assess whether the AD-templated recombinant tau fibrils show isoform-dependent conformation and dynamics, we measured 2D ^{15}N–^{1}H correlation NMR spectra. Dipolar hNH spectra selectively detect rigid β-sheet residues (Fig. 4a–e) whereas J-hNH INEPT spectra selectively detect nearly isotropically mobile residues (Fig. 4f–i). In addition to the mixed isoform samples, we also prepared AD-tau seeded 4R-only tau and 3R-only tau samples. All four seeded tau fibrils show very similar dipolar hNH spectral patterns. For example, the chemical shifts of the resolved Gly peaks are the same. This AD-tau seeded spectral fingerprint is distinct from the peak pattern of heparin-fibrillized 4R tau. For example, the heparin-fibrillized sample shows resolved signals of the only two Cys residues in the protein and their sequential Gly residues (C291–G292 and C322–G323)[22]. These peaks are absent from the AD-tau seeded spectra. Therefore, the AD tau fold can recruit 3R tau monomers, 4R tau monomers, or a mixture of both, without changing the conformation of the rigid β-sheet core.

The 2D INEPT spectra (Fig. 4f–i) also show similar chemical shifts for the four AD-tau-seeded samples[22,23,29], indicating that the fuzzy coat has the same motionally averaged conformation for different tau isoforms. N-terminal residues to G207 at the beginning of the P2 domain exhibit the same chemical shifts and relative intensities. In comparison, C-terminal residues show weaker intensities and larger intensity variations, indicating that the C-terminus is less mobile compared to the N-terminus. In contrast to the dipolar spectra, the INEPT spectra of AD-tau-seeded samples are similar to the INEPT spectrum of heparin-fibrillized tau. Therefore, the mobile part of the fuzzy coat has similar dynamic conformations for all isoforms in both AD tau aggregates and in vitro fibrillized tau.

Discussion

These data provide direct structural evidence of the nature of molecular mixing of 4R and 3R tau in AD brain tau filaments. The ^{1}H-detected ^{15}N-^{13}C REDOR experiment probes the molecular packing of tau isoforms along the fibril axis of AD tau, perpendicular to the C-shaped β-sheet structure[4]. We find that

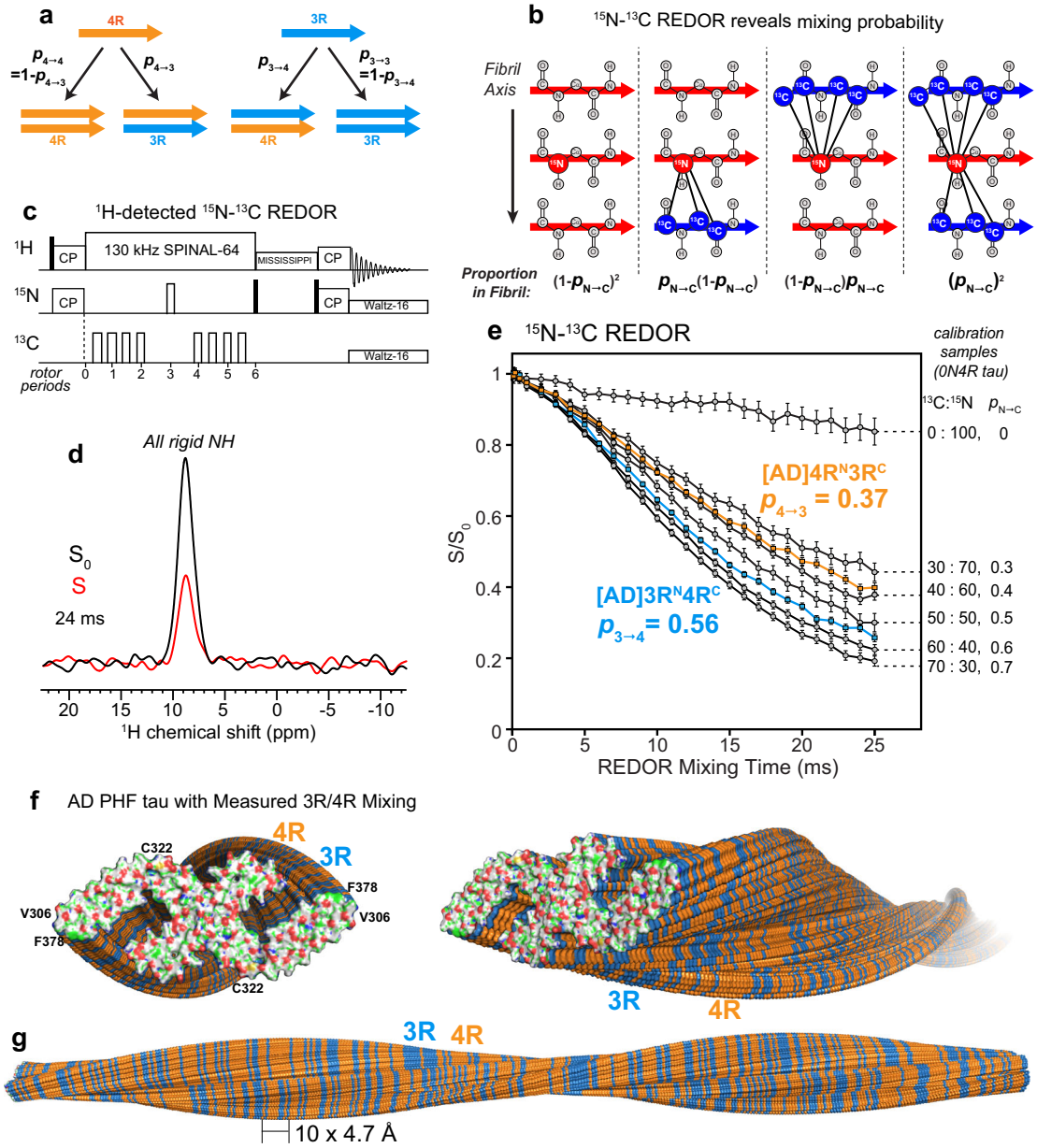


Fig. 3 ^{15}N–^{13}C REDOR NMR reveals the tau isoform mixing pattern in Alzheimer's disease brain tau filaments. a Schematic of four possible isoform combinations of two hydrogen-bonded tau molecules and their probabilities. **b** Diagram of four possible nearest neighbors of an ^{15}N-labeled tau and their probabilities in a fibril. The probability that an ^{15}N-labeled tau is followed by a ^{13}C-labeled tau is denoted by $p_{N\to C}$. Larger $p_{N\to C}$ values increase the intermolecular ^{15}N–^{13}C dipolar dephasing, which is manifested as lower REDOR S/S_0 ratios. **c** Pulse sequence for the ^{1}H-detected ^{15}N–^{13}C REDOR experiment. **d** A representative pair of the REDOR spectra, showing the total intensities of all rigid protein amides in the absence (S_0) and presence (S) of ^{13}C dephasing pulses. This spectrum was measured on [AD]4R^{N}3R^C #2 with a mixing time of 24 ms under 20 kHz MAS. **e** Measured ^{1}H-detected ^{15}N–^{13}C REDOR decay curves. The REDOR data of mixtures of ^{13}C, ^{2}H-labeled 0N4R tau and ^{15}N, ^{2}H-labeled 0N4R tau serve as calibrations (gray) for quantifying $p_{N\to C}$, which is equal to the fraction of ^{13}C-labeled protein in the mixture. The REDOR data of AD-seeded 4R^{N}3R^C indicate a probability $p_{4\to3}$ of 0.37, while the data of AD-seeded 3R^{N}4R^C give a probability $p_{3\to4}$ of 0.56. Each curve represents the average of two independently fibrillized samples. Error bars represent the propagated uncertainty from the spectral signal-to-noise ratio (one standard deviation). **f**, **g** Mixing of 4R (orange) and 3R (blue) tau in an AD PHF fibril (PDB: 5O3L). Tau isoforms along the fibril axis were simulated using the measured mixing probabilities. The simulations were done independently for each of the two protofilaments. **f** End view (left) and oblique view (right) of the PHF tau filaments. The terminal monomers' atoms are colored as H—white, C—green, N—blue, O—red, and S—yellow. **g** Sideview of the PHF tau fibril, showing a pair of protofilaments containing 360 monomers each, with a filament length of ~170 nm.

AD tau filaments recruit 4R and 3R tau monomers with a ratio of 60:40, and with a small preference for homotypic contacts. Moreover, the AD tau fibril incorporates 4R tau, 3R tau, or a mixture of the two isoforms without changing the β-sheet core structure, as seen by the similarity of the 2D dipolar hNH spectra (Fig. 4). Thus the AD tau prion incorporates either tau isoform into its pathological fold, without cross-seeding barrier between them, in contrast to in vitro assembled tau fibrils[30–32]. This fluent molecular mixing explains the biochemical observation that AD tau fibrils amplified with 4R tau, 3R tau, or both are equally capable of seeding both 4R and 3R mouse tau[18,19,25]. We hypothesize that this fluent mixing may promote fibril growth in vivo

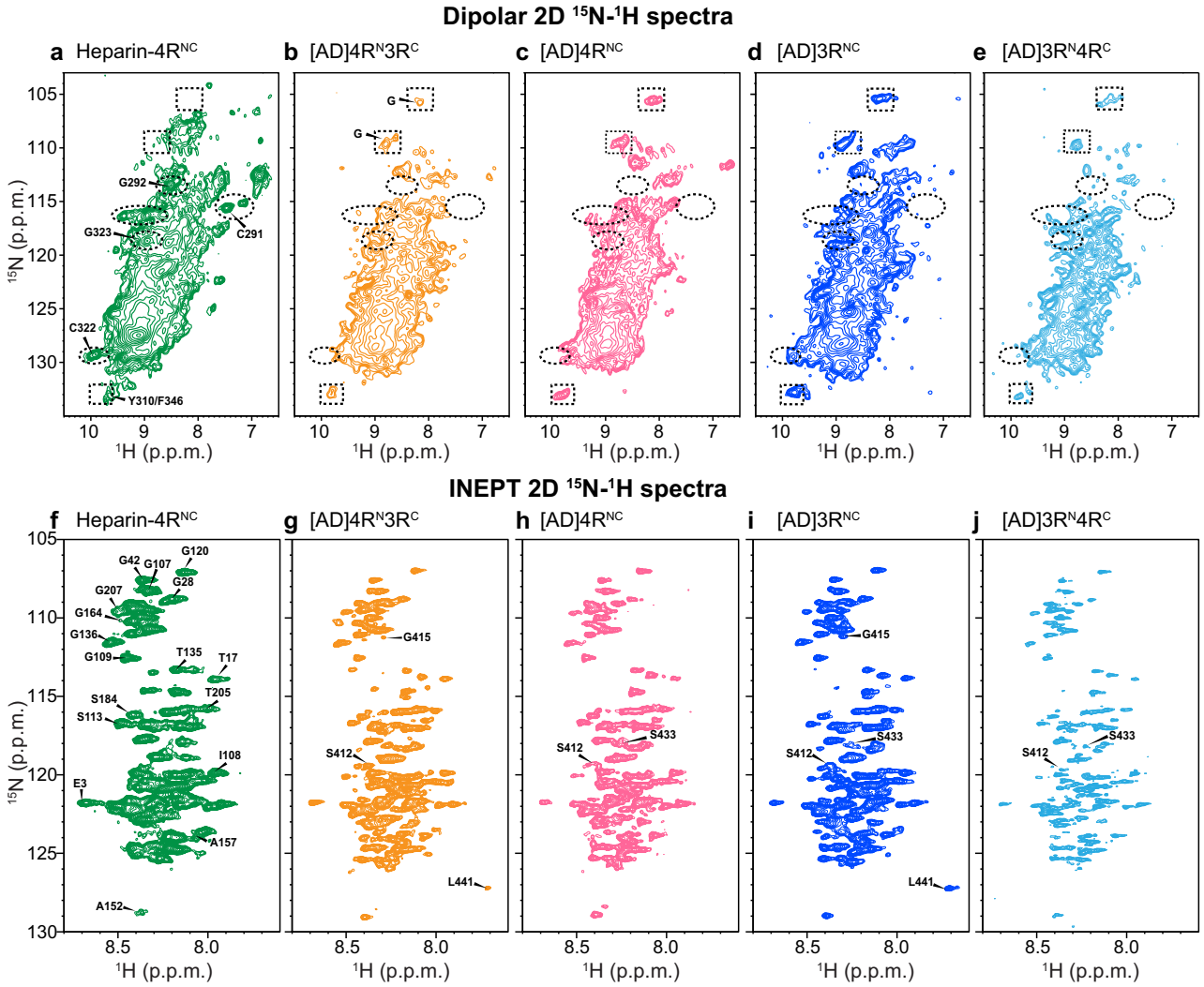


Fig. 4 2D 15N-1H correlation spectra of AD-tau-seeded filaments show isoform-independent molecular conformation and dynamics that differ from heparin-fibrilized tau. a–e 2D dipolar correlation spectra indicate a common rigid-core structure for all isoforms in AD seeded tau. **a** Heparin-fibrillized 0N4R tau. Several assignments based on 3D hCANH and hCA(co)NH spectra are indicated (Supplementary Fig. 7). **b** AD-tau seeded ^{15}N-labeled 4R tau mixed with ^{13}C-labeled 3R tau (1:1). **c** AD-tau seeded ^{15}N, ^{13}C-labeled 4R tau. **d** AD-tau seeded ^{15}N, ^{13}C-labeled 3R tau. **e** AD-tau seeded ^{15}N-labeled 3R tau mixed with ^{13}C-labeled 4R tau (1:1). Examples of peaks that are present in heparin-fibrillized tau but not in AD seeded tau are indicated by dashed ovals, while peaks that are present in AD seeded tau but not in heparin-fibrillized tau are indicated by dashed rectangles. **f–j** 2D INEPT spectra indicate similar dynamic segments for all isoforms in AD-seeded tau. Assignments for clearly resolved peaks are transferred from refs. [29,46]. **f** Heparin-fibrillized 0N4R tau. **g** AD-tau-seeded ^{15}N-labeled 4R tau mixed with ^{13}C-labeled 3R tau. **h** AD-tau-seeded ^{15}N, ^{13}C-labeled 4R tau. **i** AD-tau-seeded ^{15}N, ^{13}C-labeled 3R tau. **j** AD-tau seeded ^{15}N-labeled 3R tau mixed with ^{13}C-labeled 4R tau. Peaks from residues from the N-terminus to G207 are present in all samples, whereas C-terminal residues show varying intensities among the samples.

and play a role in making AD the most prevalent neurodegenerative disorder in humans.

Because the NMR data were obtained from AD tau seeds pooled from multiple patient brains, the 60:40 mixing and the probabilities we measured reflect the general properties of AD brain tau fibrils instead of the property of a single patient brain. The AD-seeded tau fibrils show highly reproducible 2D NMR spectra as well as reproducible REDOR dephasing, suggesting that there is no detectable molecular structural variation among the multiple AD brains. Although some biochemical studies have reported certain levels of heterogeneity in AD brain tau, these observations were all made based on total tau, including both pathogenic and non-pathogenic tau, in AD patients[33,34]. Our study instead focuses on sarkosyl-insoluble tau, which has been shown to have highly consistent fibril core structures across multiple AD patient cases[5], and which exhibits consistent seeding abilities in vitro and in vivo[19]. Therefore, these prior biochemical and biophysical data provide additional support for the isoform-independent molecular structure observed here.

The 4R and 3R tau mixing probabilities we measured here are based on the assumption that the number of residues that contributes to the β-sheet core is the same for 4R and 3R tau in the AD-tau seeded fibrils. If the number of residues in the core differs between 4R and 3R tau, then the mixing probabilities would differ from the reported value. For example, if 4R tau incorporates more β-sheet residues into the fibril core than 3R tau, then the [AD] 4R^{N}3R^{C} REDOR S/S_0 intensities would be biased high, and the true $p_{4->3}$ value would be larger than the measured value. Similarly, the true $p_{3->4}$ would be smaller than the measured value. However, three observations suggest that it is unlikely for the two isoforms to have different fibril core sizes. First, tau filaments obtained from multiple AD brains have been shown by cryoEM

to have the same β-sheet core, without signs of structural heterogeneity[4,5]. Second, the 2D hNH dipolar and INEPT spectra measured here (Fig. 4) show similar spectral peak patterns for the mixed isoform samples and the single isoform samples. Thus, the AD tau seeds are able to propagate by either isoform to develop the same structure. Third, the AD tau seeds are well amplified in our current study, which would be difficult to achieve if a monomer of the opposite isoform is added to the growing fibril with a different core size than the current isoform.

The moderate over-incorporation of 4R over 3R tau into the AD tau filament is consistent with fluorescence data of mammalian cells expressing 4R and 3R tau repeat domains. These data showed that both 4R and 3R tau are aggregated by AD brain extracts, but overexpression of 4R tau increased the total amount of the aggregates more than 3D tau did[16]. It is also consistent with the observation that 4R tau isoforms have faster fibril nucleation rates and elongation rates than 3R tau isoforms in vitro[35,36].

The isoform-independent propagation of the AD tau structure and pathological properties is noteworthy. The R2 domain contains a fibrillary hexapeptide motif like the R3 domain[37–39]. Thus, if a series of 4R tau monomers is added to the elongating filament, we would expect their R2 domains to begin to form cross-β hydrogen bonds and join the rigid core. This should gradually lead to a new fibril structure that prefers to incorporate 4R tau instead of 3R tau. The 2D NMR spectra measured here indicate that the AD tau is able to avoid this fate, even when only 4R tau monomers are added to the seed. The molecular mechanism for this avoidance is still unclear. It is possible that the posttranslational modification pattern of AD tau favors 4R tau incorporation[36] as well as predisposing the seeded structure to prevent the incorporation of the R2 domain into the β-sheet core. The AD tau fold may have evolved over its slow replication phase[40] to adapt to the environment of the AD brain[41]. The fluent molecular mixing of both 4R and 3R tau isoforms and the preservation of the β-sheet conformation regardless of the isoform could potentially be an adaptive trait to optimize prion growth and propagation. The significant amplification of the morphology and toxicity of AD tau filaments shown here opens new avenues for future characterization of the molecular structure of AD tau using a wide range of biophysical methods.

Methods

Enrichment of AD tau. AD tau was sequentially extracted from different regions of patient brains as previously described[18]. Patients were autopsied at the Center for Neurodegenerative Disease Research (CNDR) in accordance with state laws and protocols approved by the University of Pennsylvania. Informed consent was obtained from the patients' next of kin. All patients had a neuropathological diagnosis of AD at Braak stage VI. Brain tissues were obtained from the Integrated Neurodegenerative Disease Database (INDD), including six females and four males (Supplementary Table 1) whose age ranges from 55 to 92 years. Postmortem interval ranges from 4 to 23 h. As this study only uses proteins from postmortem tissues and does not involve living human subjects, it is not subject to IRB approval (45 CFR 46). All cases were screened to exclude the comorbidities of Lewy body, TDP-43, and glial tau pathologies based on immunohistochemical staining. Enzyme-linked immunosorbent assay and immunoblot measurements of tau, α-synuclein, amyloid-beta, and total protein concentrations are shown in Supplementary Table 2.

Expression and purification of perdeuterated 0N4R and 0N3R Tau. Genes for 0N4R and 0N3R tau were cloned into a pET-28a vector and transfected into *E. coli*. BL21 (DE3) competent cells (Novagen) as described previously[22,23]. Colonies from freshly transformed LB Agar plates were used to inoculate 50 ml starter cultures of H_2O LB media. All cultures were grown at 37 °C under 250 rpm shaking and contained 50 μg/ml kanamycin. After overnight growth, 4 ml of the H_2O LB starter culture was added to 25 ml of fresh D_2O LB media. This culture was grown to an OD_{600} of ~1.0 before it was added to 100 ml of fresh D_2O M9 media containing suitably labeled glucose at 2 g/l and ammonium chloride at 0.5 g/l. The M9 culture was grown to an OD_{600} of ~1.0 before it was added to 400 ml of fresh D_2O M9 media. When the 500 ml of D_2O M9 reached OD_{600} ~1.0, protein expression was induced with 0.5 mM IPTG. At this time an additional 1 g/l glucose was added. After 3–5 h of induction, cells were harvested by centrifugation and cell pellets were frozen. The D_2O M9 media contained 0.5 g/l of NH_4Cl or $^{15}NH_4Cl$ as required, and a final concentration of 3 g/l ^2H-labeled glucose or ^2H, ^{13}C-labeled glucose (Cambridge Isotope Laboratories) as required. The media contained standard M9 salts, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 50 mg/l kanamycin, 1× lyophilized MEM vitamins, and 0.2× Studier Metals.

0N4R and 0N3R tau were purified as described before[22,23]. Briefly, frozen cell pellets were thawed and homogenized by vortexing on ice in ice-cold lysis buffer. The buffer contained 20 mM Na_2HPO_4 (pH 6.8), 50 mM NaCl, 1 mM EGTA, 0.2 mM $MgCl_2$, 5 mM DTT, 0.1 mg/ml lysozyme, and 1× cOmplete™ protease inhibitor cocktail (Roche) per 50 ml lysis buffer. Lysis was performed on ice with a probe sonicator, after which cell lysate was boiled for 20 min in a water bath, then centrifuged at 15,000*g* for 40 min to remove cell debris and aggregated protein. The supernatant was purified with a cation-exchange column (self-packed with SP Sepharose Fast Flow resin; GE healthcare), then further purified using reverse phase HPLC (Agilent Zorbax 300SB-C3 column, 21.2 × 250 mm, 7 μm particle size) using an acetonitrile gradient of 5–50% in 50 min. HPLC fractions were pooled and lyophilized to yield ~10 mg/l of perdeuterated tau. The yields were similar between 0N4R and 0N3R tau.

Quantification of Tau monomers for fibrillization and seeding reactions. To precisely control the amount of tau monomers added to the fibrillization mixtures, we dissolved purified and lyophilized ^2H, ^{15}N (DN) or ^2H, ^{13}C (DC)-labeled tau in ice-cold phosphate buffer saline (PBS) buffer at 10–15 mg/ml in the absence of DTT. Monomer concentration was measured with a NanoDrop Lite (Thermo Fisher). This high concentration (10–15 mg/ml) of tau monomers was used for quantification because the tau extinction coefficients are predicted to be low: A_{280} 0.186 for 1 mg/ml 0N4R tau and 0.203 for 1 mg/ml 0N3R tau. These values do not account for increased isotopic mass. Absorbances were measured in triplicate and averaged. These gave an estimated uncertainty of ~1% in the reported ^{13}C:^{15}N monomer ratios in the reaction mixture.

Heparin-induced fibrillization of Tau. Heparin-induced tau fibrils were prepared as described before[22]. Ten micromolar DN- or DC-labeled 0N4R tau monomers (e.g., 0.4 mg/ml total 0N4R tau) and 1 mM DTT were added to pH 7.4 PBS buffer and bath sonicated for 1 min. To this solution we added 0.125 mg/ml heparin (Santa Cruz Biotech, sc203075, 8000–25,000 Da), briefly vortexed to mix, then split the solution into 1.5 ml Eppendorf tubes. The tubes were shaken at 37 °C and 1220 rpm with 2 mm orbit for 3 days. After this, the solutions were ultracentrifuged in a TLA-55 rotor at 100,000*g* for 1 h, and the fibrils were collected by discarding most of the supernatant. The fibrils were resuspended, combined, and then subjected to a second ultracentrifugation step to obtain the total pellet.

Seeded amplification of AD Tau fibrils. AD tau seeds were amplified in solutions containing final concentrations of 4 μM seeds and 36 μM total recombinant tau, along with 2 mM DTT, 5 mM PMSF, and 1% v/v of a protease inhibitor cocktail (1% w/v each of pepstatin, leupeptin, TPCK, TLCK, and trypsin inhibitor, 0.1 M EDTA), in PBS buffer. Initially, AD seeds, DTT, and protease inhibitors for 4 × 100 μl reactions were combined into a single 0.2 ml flat capped PCR tube (Fisher Scientific 14230225), with a volume of ~120 μl. The seed mixture was heat treated at 56 °C in a thermocycler for 30 min, and then sonicated by securing tubes at the antinodes of a Branson 2200 bath sonicator for 30 min. When sonicating, the tubes were secured in place within individual flotation devices which were pinned to rods taped to the sonicator. The exact positions of the tubes were adjusted for each use to maximize sonication force. The seed mixture was then split evenly into four PCR tubes, into which the appropriate amount of recombinant tau solution and PBS buffer were added to bring the final volume to 100 μl. When 4R and 3R tau monomers were mixed, each monomer was added to 18 μM final concentration; otherwise 36 μM of a single monomer was used. Reactions were wrapped in parafilm and shaken at 37 °C and 1220 rpm with 2 mm orbit for 7 days. Aliquots were taken immediately before and after 7 days of shaking as day 0 and day 7 samples, respectively. Monomer-only controls used PBS buffer instead of AD tau seeds. SDS-PAGE and neurotoxicity assays were conducted in duplicate for each AD-tau-seeded sample at day 0 and day 7. These duplicates were taken from reactions that used separately heat-treated and sonicated seed and protease inhibitor mixtures, although all protease inhibitors originally came from one stock, and all seeds came from one pool of AD brain material. After taking aliquots for biochemical assays, seeding reactions were pooled and fibrils were collected by ultracentrifugation in a TLA-55 rotor at 100,000*g* for 1 h.

Sedimentation assay of fibril samples. Sedimentation assays were conducted by mixing a 1 μl aliquot of fibrillization reaction mixture with 19 μl of PBS (0.1% sarkosyl, w/v) and ultracentrifuging for 30 min at 100,000*g* (Optima, Beckman Coulter) as previously described[19]. The supernatant was carefully removed, and the pellet was resuspended with 20 μl PBS. Lysates were mixed with loading buffer for immunoblots. Samples were separated using a 10% SDS-PAGE gel.

Tau species were revealed using antibodies 17025 (made in-house) and PHF1 (a gift from Dr. Peter Davis). The 17025 antibody was raised with recombinant human tau; thus, it has a higher affinity for recombinant tau than AD brain tau. PHF1 is more specific to the S396/S404 phosphorylated tau; thus, the bands in

PHF1 blots (Fig. 1b) mostly represent hyperphosphorylated AD tau seeds. As there are six isoforms of human tau and high molecular-weight tau species are present in the AD seeds, the banding pattern of PHF1-tau is more complex than the pattern of recombinant tau (17025-tau).

Because the AD tau seeds (PHF-1 positive bands) were sonicated before being added to the seeding reactions, ultracentrifugation on day 0 at 100,000*g* for 30 min in the sedimentation assay could not completely sediment the insoluble AD tau seeds, and a significant amount of tau seeds stay in the supernatant. After 7 days of seeding reactions, tau filaments were elongated and became easier to be pelleted, resulting in a shift in the amount of AD tau to the pellet fraction. But the total amount of AD tau in the supernatant and pellet was similar before and after the reaction. Thus, there is no evidence of kinases in the insoluble tau seeds.

Transmission electron microscopy (TEM) of Tau fibrils. After fibril growth, 1 µl aliquots of fibrillization reactions were diluted fivefold using deionized water and adsorbed onto freshly glow-discharged, 200 mesh Formvar/carbon-coated copper grids (Ted Pella) for 60 s, blotted and washed with deionized water twice, and stained with 0.5% (w/v) uranyl formate for 20 s. TEM images were measured on an FEI Tecnai G2 Spirit TWIN electron microscope. Raw image files were processed and exported using ImageJ 1.53a. The dimensions of the fibrils were sampled manually using the built-in measurement tool of ImageJ. At least five different particles from two or more different TEM images were sampled and a total of at least 10 readings of width and crossover length were averaged to give the statistics shown in Supplementary Table 3.

Pathogenicity test in primary neurons. CD1 wild-type mice were used to generate primary hippocampal neurons. Female pregnant mice were purchased from Charles River. Animals were kept in microisolator cages at a temperature between 18 and 23 °C with 40–60% humidity, with 12 h light on and 12 h light off. All animal care and experimental protocols were approved by the University of Pennsylvania's Institutional Animal Care and Use Committee. CD1 mouse cortices and hippocampi were dissected at embryo days 16–18 and dissociated with papain (Worthington Biochemical Corporation). Neurons were resuspended in neural basal medium (Gibco, 21103) with 2% B27 (Gibco), 1× Glutamax (Gibco), and 1× penicillin/streptomycin (Gibco). Plates or coverslips were coated with poly-D-lysine (0.1 mg/ml, Sigma-Aldrich) in borate buffer (0.05 M boric acid, pH 8.5) overnight at room temperature. Cells were plated at a density of 50,000 cells/cm^2 for 96- or 384-well plates. When plating, 5% fetal bovine serum (FBS) was added to the cell suspension. The plating medium was replaced with neural basal medium without FBS at day 1 in vitro (DIV1). At DIV7, the plating medium was replaced with condition medium containing 1:1 ratio of old and fresh medium. AD tau seeds or seeded reactions were first diluted in PBS to the desired tau concentrations, and then added to the cells and incubated for 14 days until DIV21. The dose of tau fibrils was converted based on cell densities (µg tau/10^6 cells). For activity measurements, if not otherwise specified, 1 µg tau/10^6 cells were used for 100% seeds (100% pathogenic tau), 0.1 µg tau/10^6 cells for 10% seeds control, and 0.01 µg tau/10^6 cells for 1% seeds control. All amplified tau variant doses were based on the total tau concentration at the start of the in vitro reaction. One-way ANOVA followed by Tukey post hoc test was conducted using 3–4 biological repeats. This number of repeats was utilized based on previous experience in similar studies obtaining statistically significant results.

Immunocytochemistry. Cells were washed four times with PBS and extracted with 1% hexadecyltrimethyl–ammonium bromide (HDTA; Sigma-Aldrich) for 10 min following 10 min of fixation with 4% paraformaldehyde and sucrose in PBS. Fixed cells were stained with primary antibodies overnight at 4 °C, and then with Alexa Fluor-conjugated secondary antibodies (Thermo Fisher Scientific) for 2 h at room temperature. Images were randomly taken using an InCell Analyzer 2200 microscope and quantified for the antibody immunoreactivity. Tau pathology was quantified as the area of occupancy × intensity/DAPI count.

Solid-state NMR experiments. All AD-tau seeded tau fibrils and heparin-fibrilized tau samples were packed into Bruker 1.3 mm magic-angle-spinning (MAS) rotors using a custom-designed ultracentrifuge packing tool. Fibril pellets were resuspended in ~120 µl of their own supernatant, loaded into the packing tool, and then spun on a Beckman Optima XL-80 with a SW60 Ti rotor at 311,000*g* for 16 h at 10 °C. For AD-tau-seeded samples, each MAS rotor contained ~4 mg hydrated material, a quarter of which was the dry mass of isotopically labeled recombinant tau. For heparin-fibrillized tau samples, ~2 mg hydrated material was packed into the rotor, in which ~1 mg was the dry mass of the protein.

Solid-state NMR experiments were conducted on a Bruker Avance III HD 600 MHz (14.1 T) spectrometer and an Avance II 800 MHz (18.8 T) spectrometers using 1.3 mm HCN MAS probes. ^{1}H chemical shifts were referenced externally to sodium trimethylsilylpropanesulfonate (DSS) at 0 ppm. ^{13}C and ^{15}N chemical shifts were referenced externally using the tripeptide *N*-formyl-Met-Leu-Phe-OMe (f-MLF). The Met Cε peak was calibrated to 14.0 ppm on the tetramethylsilane (TMS) scale while the Phe ^{15}N peak was calibrated to 110.09 ppm on the liquid ammonia scale. Most experiments were conducted under 20 kHz MAS, except for the 2D dipolar hNH and 3D hCANH and hCA(co)NH experiments, which were conducted under 55 kHz MAS. Sample temperatures were kept at 285 or 297 K, which was estimated based on the water ^{1}H chemical shift (δ_{H2O}) of 4.89 or 4.77 ppm, respectively using the equation T_{eff} (K) = 96.9 × (7.83 − δ_{H2O})[42]. Typical radiofrequency (rf) field strengths were 65–90 kHz for ^{1}H, 62.5 kHz for ^{13}C, and 45–50 kHz for ^{15}N. In all ^{1}H-detected experiments, water suppression was carried out using the MISSISPPI sequence with a ^{1}H rf field strength of 15 kHz[43]. For REDOR experiments under 20 kHz MAS, 130 kHz ^{1}H SPINAL-64 decoupling was applied during the ^{15}N–^{13}C REDOR period, while 3.6 kHz ^{15}N and ^{13}C WALTZ-16 decoupling was applied during ^{1}H detection. For 2D hNH experiments under 55 kHz MAS, 10 kHz ^{1}H, ^{15}N, and ^{13}C WALTZ-16 decoupling were applied.

One-dimensional (1D) NMR spectra for quantifying the amounts of ^{15}N-labeled and ^{13}C-labeled tau in the samples were conducted under 20 kHz MAS. Quantitative ^{13}C direct polarization (DP) spectra were measured at 285 K using a recycle delay of 5 s and 2048 scans. 1D J-hnH spectra were measured to quantify the amount of ^{15}N spins in the dynamic portion of the protein. These spectra were measured at 297 K. 1D dipolar hnH spectra were measured at 285 K to quantify the amount of ^{15}N spins in the rigid core. The experiment used ^{1}H–^{15}N and ^{15}N–^{1}H cross-polarization (CP) contact times of 1.25 and 0.4 ms, respectively. The spectra were measured with 256 scans and a recycle delay of 1.5 s.

Molecular mixing between ^{15}N-labeled and ^{13}C-labeled tau was detected using the ^{1}H-detected ^{15}N-^{13}C REDOR experiment at 285 K. A ^{15}N–^{13}C REDOR period with a variable dipolar recoupling time of 0.1 to 25 ms was inserted before the MISSISPPI period in a dipolar hnH pulse sequence (Fig. 3c). A total of 28 mixing times were measured per sample. The control spectrum (S_0) without ^{13}C pulses and the dephased spectrum (S) with ^{13}C pulses were measured back-to-back for each mixing time to avoid spectrometer drift. In total, 1536–2048 repeated scans were coadded for each mixing time. This signal averaging was separated into 6–8 blocks of the complete set of mixing times to allow re-optimization of experimental conditions between blocks. A full REDOR curve for each sample was measured in 3 days.

2D ^{15}N–^{1}H correlation spectra were measured to compare the AD-tau seeded samples and heparin-fibrillized 0N4R tau sample. The 2D dipolar hNH spectra selectively detect the rigid core. The spectra of heparin-fibrilized 0N4R tau and [AD]4R^{NC} were measured on the 800 MHz NMR, while the other samples were measured on the 600 MHz NMR. Except for the magnetic field difference, all experiments were conducted under the same conditions, including 55 kHz MAS, a sample temperature of 285 K, and 1.25 ms ^{1}H–^{15}N CP and 0.4 ms ^{15}N–^{1}H CP contact times. The 2D spectra were measured with 38 ms ^{1}H acquisition time and up to 30 ms ^{15}N evolution. Each 2D spectrum took 2–3 days of signal averaging. 2D INEPT (J-hNH) spectra were measured to observe the signals of isotropically mobile residues. The spectra were measured on the 600 MHz NMR at 297 K, with 92 ms ^{1}H acquisition and up to 70 ms ^{15}N evolution time. Each spectrum took ~2 days of signal averaging. All 2D spectra were processed using QSINE = 3 apodization.

3D hCANH and hCA(co)NH correlation spectra were used to assign part of the backbone of the rigid core of heparin-fibrillized 0N4R tau. Both experiments started from 2 ms ^{1}H–^{13}C CP and underwent up to 4.8 ms ^{13}C evolution. Intra-residue Cα-N polarization transfer was accomplished by 8 ms SPECIFIC CP in hCANH, while inter-residue CO-N transfer was carried out using consecutive 8 ms Cα-CO DREAM and 8 ms CO-N SPECIFIC CP. The ^{15}N evolution time was up to 10.5 ms, and 0.4 ms ^{15}N-^{1}H CP transferred the polarization to ^{1}H for detection. Both 3D spectra were processed using GM apodization with LB = 15 Hz and GB = 0.07 and forward linear prediction. Detailed experimental parameters can be found in Supplementary Table 5.

Two-state Markov model of 4R and 3R Tau mixing with probabilities $p_{3\to4}$ and $p_{4\to3}$. We consider a parallel-in-register cross-β fibril composed of a mixture of 4R and 3R tau. We assume that the fibril has a constant 4R:3R tau molar ratio over its entire length, and the probability $p_{4\to3}$ that a 4R tau is followed by a 3R tau is the same as the probability that a 4R tau is preceded by a 3R tau (Fig. 3a). Then the probability that the 4R tau is followed or preceded by another 4R tau is $p_{4\to4} = 1\text{-}p_{4\to3}$. The probabilities $p_{3\to3}$ and $p_{3\to4}$ are defined analogously.

We next define the mole fraction of 4R tau in one part of the fibril as χ_4. The mole fraction of 3R tau in the same part of the fibril is then $\chi_3 = 1 - \chi_4$. If we shift the portion of the fibril under consideration by one monomer, then the new mole fraction χ_4' of 4R tau is

$$\chi_4' = \chi_4 p_{4\to4} + \chi_3 p_{3\to3} \quad (1)$$

Because the 4R:3R tau concentration ratio is assumed to be constant throughout the fibril, this new mole fraction must be equal to the original mole fraction, $\chi_4' = \chi_4$. From this we obtain

$$\chi_4 = \frac{p_{3\to4}}{p_{3\to4} + p_{4\to3}} \quad (2)$$

Therefore, the measured mixing probabilities encode the mole fractions of 4R and 3R tau. These mole fractions are equivalent to the global incorporation levels of 4R and 3R tau, which reflect the preference of AD PHF tau for the two isoforms.

The preference of AD-tau seeded 4R and 3R tau to separate like block copolymers or alternate along the fibril axis can be described independently of the isoform incorporation level. To quantify this property, we define a mixing quotient,

which is analogous to the equilibrium constant of mixing:

$$Q = \frac{p_{3\to4}p_{4\to3}}{p_{4\to4}p_{3\to3}} \quad (3)$$

This mixing quotient reflects the tendency of a like-interface (4–4 or 3–3) to dissociate to form an unlike interface (4–3 or 3–4). If Q is much larger than 1, then there is a strong preference for 4R and 3R to alternate. A perfectly alternating 4R–3R–4R–3R fibril has an infinite Q. Conversely, if Q is much less than 1, then there is a strong preference for 4R–4R and 3R–3R contacts, i.e., a homotypic preference. In this block copolymerization case, $p_{4\to4}$ and $p_{3\to3}$ are much larger than $p_{3\to4}$ and $p_{4\to3}$. If there is neither preference for alternation nor neither for block copolymerization, then $p_{3\to4} = p_{4\to4} = \chi_4$, and $Q = 1$.

Therefore, assuming there are no additional correlations between non-neighboring chains, $p_{3\to4}$ and $p_{4\to3}$ completely characterize the mixing properties of a uniform fibril. The parameters Q and χ_4 are alternative and equivalent expressions of these two probabilities. $p_{3\to4}$ and $p_{4\to3}$ can be measured by ^{15}N–^{13}C REDOR, giving experimental access to the mixing quotient and the global incorporation level of 4R tau.

Extraction of $p_{4\to3}$ and $p_{3\to4}$ from REDOR curves. The probability $p_{4\to3}$ and $p_{3\to4}$ in AD seeded samples were extracted by comparing the ^{15}N-^{13}C REDOR curves of [AD]4R^{N}3R^{C} and [AD]3R^{N}4R^{C} with the REDOR curves of heparin-fibrillized calibration samples with known ratios of ^{15}N-labeled and ^{13}C-labeled 0N4R tau (Fig. 3e). The REDOR data of the AD-seeded mixed 4R/3R tau were measured in duplicate using separately seeded samples (#1 and #2 for each labeling scheme). The REDOR curves of the two replicates of [AD]4R^{N}3R^{C} were averaged, then the $p_{4\to3,\,i}$ value at each mixing time i between 8 and 25 ms was linearly interpolated from the REDOR calibration curves for $p_{N\to C} = 0.3$ and $p_{N\to C} = 0.4$ according to

$$p_{4\to3,i} = \frac{\left(\frac{S}{S_{0N4RC3R,i}} - \frac{S}{S_{040/60,i}}\right)}{\left(\frac{S}{S_{030/70,i}} - \frac{S}{S_{040/60,i}}\right)} * (30 - 40) + 40 \quad (4)$$

These individual $p_{4\to3,i}$ values were averaged to give the final $p_{4\to3}$ value. The error bar in the $p_{4\to3}$ value is propagated from the spectral signal-to-noise ratios of the calibration spectra and the AD-tau seeded spectra using the following equation :

$$\varepsilon(p_{4\to3}) = 10 * \left[\frac{1}{n}\sum_i^n \left(\left(\frac{\varepsilon_{N4RC3R,i}}{\frac{S}{S_{030/70,i}} - \frac{S}{S_{040/60,i}}}\right)^2 + \left(\frac{\left(\frac{S}{S_{0N4RC3R,i}} - \frac{S}{S_{040/60,i}}\right) * \varepsilon_{030/70,i}}{\left(\frac{S}{S_{030/70,i}} - \frac{S}{S_{040/60,i}}\right)^2}\right)^2 + \left(\frac{\left(\frac{S}{S_{0N4RC3R,i}} - \frac{S}{S_{040/60,i}}\right) * \varepsilon_{040/60,i}}{\left(\frac{S}{S_{030/70,i}} - \frac{S}{S_{040/60,i}}\right)^2}\right)^2\right)\right]^{1/2} \quad (5)$$

Here ε is the spectral noise obtained from the signal-to-noise ratio of the REDOR spectra and $n = 18$ for the number of REDOR mixing times from 8 ms to 25 ms.

The $p_{3\to4}$ value was similarly extracted by comparing the REDOR curves of [AD]3R^{N}4R^{C} samples with the REDOR data of the calibration samples for $p_{N\to C} = 0.5$ and $p_{N\to C} = 0.6$. Errors in $p_{3\to4}$ and $p_{4\to3}$ were propagated to errors in the calculated χ_4 and Q following Gauss' law of error propagation.

Among the measured REDOR calibration curves from the heparin-fibrillized 0N4R tau samples, the curve of the 100% ^{15}N-labeled 0N4R tau without ^{13}C-labeled protein already exhibits dipolar dephasing to 0.84 at 25 ms (Fig. 3e and Supplementary Fig. 3a). This indicates that natural abundance ^{13}C spins cause non-negligible dephasing to ^{15}N. This natural abundance effect is approximately corrected by dividing all measured REDOR intensities by the intensities of the 100% ^{15}N-labeled sample (Supplementary Fig. 3b).

Estimation of Tau monomer fractions in the seeded samples by 1D NMR spectra. To provide additional verification of the overall incorporation level of 4R and 3R tau in the fibril, we quantified the intensities of 1D ^{1}H-detected ^{15}N NMR spectra as well as ^{13}C spectra. Quantitative ^{13}C DP NMR spectra were measured using a 5 s recycle delay to quantify the amount of total ^{13}C-enriched tau monomers in the fibrils. 1D CP-hnH spectra were used to quantify the amount of the rigid portions of ^{15}N-enriched tau, while J-hnH spectra were used to quantify the amount of the mobile portions of ^{15}N-enriched tau (Supplementary Fig. 5). The 5 s DP ^{13}C spectra were integrated between 200 and 0 ppm, while the hnH spectra were integrated between 12 and 6 ppm. The intensity ratios of the J-hnH spectra and the ratios of the CP-hnH spectra were averaged. As the double-labeled [AD]4R^{NC} contains an equal amount of ^{13}C and ^{15}N tau, normalizing the ^{13}C and ^{15}N intensities from the mixed-labeled samples to the [AD]4R^{NC} intensities allows quantification of the ^{13}C: ^{15}N molar ratios, and hence the 4R:3R tau molar ratios. [AD]4R^{N}3R^{C} sample #1 was found to contain 60.0 ± 3.0% 4R tau, whereas sample #2 contained 62.6 ± 3.0% 4R tau (Supplementary Table 4). Following a similar procedure, we found that the two duplicates of [AD]3R^{N}4R^{C} contained 60.8% ± 3.0% and 62.1 ± 3.0% of 4R tau, respectively. The 3.0% uncertainty was estimated based on heparin-fibrillized 0N4R tau calibration samples, where the ^{15}N-labeled 4R monomer fraction derived from the 1D spectra deviated from the added fraction by 3.0%. We attribute this uncertainty to multiple polarization transfer steps in the J-hnH and CP-hnH experiments. The actual uncertainty in the monomer concentration is 1.0% based on the nanodrop reading.

REDOR simulations using SpinEvoluton and second moment analysis. To assess whether the experimentally measured ^{15}N–^{13}C REDOR curves of in vitro calibration samples are in good agreement with the known cross-β fibril geometry, and to evaluate the accuracy of the mixing probability measurement using this approach, we simulated the REDOR curves for varying ratios of ^{13}C-labeled protein and ^{15}N-labeled protein in an idealized cross-β fibril. The simulations were carried out in three ways, using (1) the SpinEvolution software, (2) the second-moment analysis for up to 9 ^{13}C spins, and (3) the second-moment analysis for up to 48 ^{13}C spins. These different approaches are compared (Supplementary Fig. 3c–e, f-h) to account for the large number of ^{13}C spins, both labeled and natural abundance ones, near each observed ^{15}N spin.

All three simulations consist of superpositions of multiple REDOR curves weighed by the probabilities of ^{13}C and ^{15}N mixing in a cross-β hydrogen-bonded fibril. Denoting the ^{13}C:^{15}N mixing ratio in the 0N4R tau calibration samples as $p_{N\to C}$, each of the four packing modes in Fig. 3b occurs with a well-defined probability: $(1\text{-}p_{N\to C})^2$ for the N–N–N packing mode, $p_{N\to C}(1- p_{N\to C})$ for the N–N–C packing mode, $(1-p_{N\to C})\,p_{N\to C}$ for the C–N–N packing mode, and $(p_{N\to C})^2$ for the C–N–C packing mode. The REDOR dephasing of the individual packing modes was weighted by these for $p_{N\to C}$ values to give the total REDOR dephasing curves for each ^{13}C:^{15}N mixing ratio. To obtain the closest ^{15}N–^{13}C distances in a standard cross-β hydrogen-bonded fibril, we used I308 in the AD PHF tau structure (PDB: 5O3L) as the reference ^{15}N spin and extracted its distances to the nearest ^{13}C spins in the two neighboring chains in PyMOL.

Simulations using SpinEvolution[27] explicitly account for all nuclear spin interactions but are restricted to a maximum of 9 ^{13}C spins for each ^{15}N spin. In these SpinEvolution simulations, we neglect the natural abundance (1.1%) ^{13}C nuclei in the central ^{15}N-labeled monomer. The coordinates of ^{13}C nuclei relative to the I308 ^{15}N were extracted in PyMOL. The internuclear distance r_{NC} ranges from 4.1 to 5.8 Å. Specifically, in the N–N–N packing mode, no ^{13}C spin was considered and hence there is no REDOR dephasing. In the N–N–C packing mode, all five ^{13}C spins within 5.8 Å were simulated (Fig. 3b). In the C–N–N packing mode, all eight ^{13}C spins within 5.8 Å were simulated. In the C–N–C packing mode, all nine ^{13}C spins within 5.5 Å were simulated (Supplementary Fig. 3i). ^{13}C–^{15}N and ^{13}C–^{13}C dipolar couplings were calculated internally by the software. The chemical shift anisotropy (CSA) of ^{15}N was neglected. All aliphatic ^{13}C nuclei were given a CSA anisotropy $\delta = \delta_{zz} - \delta_{iso}$ of 30 ppm, an asymmetry parameter $\eta = (\delta_{yy} - \delta_{xx})/\ \delta$ of 0, where δ_{xx}, δ_{yy}, and δ_{zz} are the principal values of the CSA and δ_{iso} is the average of the three principal values. All carbonyl ^{13}C nuclei were given a δ of 170 ppm and η of 0. The orientation of the CSA tensors was set to be random relative to the dipolar vector. Random coil chemical shifts were used as the isotropic chemical shift of each atom[44]. The ^{1}H Larmor frequency was set to 600 MHz, and the MAS frequency was set to 20 kHz. The rf carrier frequency for ^{13}C and ^{15}N channel were set to 100 and 120 ppm, respectively, and 8 µs rectangular 180° pulses were used for both channels. Powder averaging was carried out using the rep30 set of Euler angles based on the REPULSION power averaging scheme[45]. Options "-dw123 -re" were applied. The ^{15}N signal was simulated for REDOR mixing times of 0–50 ms in 0.1 ms steps. The SpinEvolution code used in the simulation can be found in the Source Code.

To evaluate the effects of more than nine ^{13}C spins to REDOR dephasing, we conducted a second-moment analysis for the same set of ^{13}C spins used for SpinEvolution and for up to 48 ^{13}C spins within a radius of 8.8 Å to the I308 ^{15}N in neighboring monomers (Supplementary Fig. 3i). Following Hirschinger[28], we calculated the second moment of n ^{13}C–^{15}N dipolar couplings by summing the squares of the individual couplings and dividing by 5,

$$M_2 = \frac{1}{5}\sum_i^n \left(\frac{3000\text{Hz}}{r_i^3}\right)^2. \quad (6)$$

Here 3000 Hz is the ^{13}C–^{15}N dipolar coupling constant for a 1 Å distance and the internuclear distance r is expressed in the unit of Å. The REDOR dephasing curve for these n ^{13}C-^{15}N dipolar couplings has a Gaussian shape:

$$\frac{S}{S_0} = \exp\left(-4M_2t^2/3\pi^2\right) \quad (7)$$

This second moment analysis ignores ^{13}C finite pulse effects, ^{13}C chemical shifts, and ^{13}C–^{13}C dipolar couplings. The second moment analysis was conducted using Microsoft Excel 2021, with equation embedded in the Source Data.

Supplementary Figure 3 shows that the SpinEvolution simulations and second-moment analysis produce REDOR dephasing curves that deviate in opposite ways from the measured curves after ^{13}C natural abundance correction. The 9-spin SpinEvolution simulations result in slightly higher S/S_0 values from the experimental values, whereas the 48-spin secondary moment analysis result in lower S/S_0 values than the experimental values (Supplementary Fig. 3f, h). The 9-spin second-moment analysis approaches the experimental data the closest (Supplementary Fig. 3g). The faster dephasing of the simulated REDOR curves with 48 nearest ^{13}C spins compared to dephasing by 9 ^{13}C spins indicate that

^{15}N–^{13}C distances between 5.8 and 8.8 Å cannot be neglected. Thus, SpinEvolution simulation for up to nine ^{13}C spins is an underestimate. On the other hand, the explicit SpinEvolution simulation with nine ^{13}C spins decays more slowly than the second moment analysis for nine spins, indicating that other spin interactions, including ^{13}C–^{13}C dipolar couplings and chemical shifts, slow down the ^{15}N–^{13}C dipolar dephasing. This means the second-moment analysis of 48 ^{13}C spins overestimates the true decay rate, by removing, for example, the effect of ^{13}C–^{13}C dipolar coupling truncating the ^{13}C–^{15}N difference tensors during the REDOR evolution. Therefore, neither quantum-mechanical simulations of a limited number of spins and second-moment analysis of a large number of spins without all spin interactions fully account for both long-range ^{13}C–^{15}N distances and all nuclear spin interactions. Due to these limitations, we rely on the experimentally measured calibration REDOR curves to extract the mixing probabilities of AD-tau seeded filaments accurately.

Mixed fibril construction and modeling of 4R/3R tau mixing. We constructed a cross-β filament containing 360 × 2 monomers of tau using the AD PHF tau cryo-EM structure (PDB: 5O3L) (Fig. 3f, g). The 5 × 2 monomers in the original PDB structure were duplicated, then the duplicate was translated by 4.7 Å per monomer along the fibril axis, and rotated around the fibril axis by 1° per monomer. The duplicate was then merged with the original PyMOL object. This process was repeated until the filament contained 360 × 2 monomers.

A python script was written to simulate a filament of 360 tau monomers with 4R and 3R tau mixed according to the measured $p_{3\to4}$ and $p_{4\to3}$ values. The simulation was repeated until the fraction of 4R tau followed by a 3R tau is 37 ± 1%, and the fraction of 3R tau followed by a 4R tau is 56 ± 1%. This was done independently for both protofilaments. This script can be found in Source Code.

To produce statistical models of 100-monomer filaments with varying Q and χ_4 values (Supplementary Fig. 6), we first converted Q and χ_4 to $p_{3\to4}$ and $p_{4\to3}$, and then randomly simulated the fibrils using these probabilities. These short simulations have high variance of observed Q and χ_4; thus, the 100-monomer fibrils were repeatedly simulated until both Q and χ_4 were within 10% and 5% of their desired values, respectively. This script can be found in the Source Code.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The REDOR NMR intensities and pathological activities data generated in this study are provided as a Source Data file. NMR spectra are available from M.H. upon request. The AD PHF tau cryo-EM structure used for the construction and modeling of 4R/3R tau mixing is publicly available following the link [https://doi.org/10.2210/pdb5O3L/pdb].

Code availability

The SPINEVOLUTION REDOR simulation codes, the python code for building a fibril model with desired mixing probabilities, and the python code for simulating tau fibrils with different Q and χ_4 values, are included in the Source Code.

Received: 4 January 2022; Accepted: 4 May 2022;
Published online: 27 May 2022

References

1. Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D. & Crowther, R. A. Multiple isoforms of human microtubule-associated protein-Tau—sequences and localization in neurofibrillary tangles of Alzheimers-disease. *Neuron* **3**, 519–526 (1989).
2. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* **82**, 239–259 (1991).
3. Lee, V. M., Goedert, M. & Trojanowski, J. Q. Neurodegenerative tauopathies. *Annu. Rev. Neurosci.* **24**, 1121–1159 (2001).
4. Fitzpatrick, A. W. P. et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
5. Falcon, B. et al. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathol.* **136**, 699–708 (2018).
6. Shi, Y. et al. Structure-based classification of tauopathies. *Nature* **598**, 359–363 (2021).
7. Falcon, B. et al. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
8. Braak, H. & Del Tredici, K. Potential pathways of abnormal Tau and alpha-Synuclein dissemination in sporadic Alzheimer's and Parkinson's diseases. *CSH Perspect. Biol.* **8**, a023630 (2016).
9. Goedert, M., Eisenberg, D. S. & Crowther, R. A. Propagation of Tau aggregates and neurodegeneration. *Annu. Rev. Neurosci.* **40**, 189–210 (2017).
10. Gibbons, G. S., Lee, V. M. Y. & Trojanowski, J. Q. Mechanisms of cell-to-cell transmission of pathological Tau: a review. *JAMA Neurol.* **76**, 101–108 (2019).
11. Guo, J. L. & Lee, V. M. Y. Seeding of normal Tau by pathological Tau conformers drives pathogenesis of Alzheimer-like tangles. *J. Biol. Chem.* **286**, 15317–15331 (2011).
12. Guo, J. L. & Lee, V. M. Y. Cell-to-cell transmission of pathogenic proteins in neurodegenerative diseases. *Nat. Med.* **20**, 130–138 (2014).
13. Falcon, B. et al. Conformation determines the seeding potencies of native and recombinant Tau aggregates. *J. Biol. Chem.* **290**, 1049–1065 (2015).
14. Frost, B., Jacks, R. L. & Diamond, M. I. Propagation of Tau misfolding from the outside to the inside of a cell. *J. Biol. Chem.* **284**, 12845–12852 (2009).
15. Sanders, D. W. et al. Distinct tau prion strains propagate in cells and mice and define different tauopathies. *Neuron* **82**, 1271–1288 (2014).
16. Woerman, A. L. et al. Tau prions from Alzheimer's disease and chronic traumatic encephalopathy patients propagate in cultured cells. *Proc. Natl Acad. Sci. USA* **113**, E8187–E8196 (2016).
17. DeVos, S. L. et al. Synaptic Tau seeding precedes Tau pathology in human Alzheimer's disease brain. *Front. Neurosci.* **12, 267** (2018).
18. Guo, J. L. et al. Unique pathological tau conformers from Alzheimer's brains transmit tau pathology in nontransgenic mice. *J. Exp. Med* **213**, 2635–2654 (2016).
19. Xu, H. et al. In vitro amplification of pathogenic tau conserves disease-specific bioactive characteristics. *Acta Neuropathol.* **141**, 193–215 (2021).
20. Wischik, C. M. et al. Structural characterization of the core of the paired helical filament of Alzheimer disease. *Proc. Natl Acad. Sci. USA* **85**, 4884–4888 (1988).
21. Goedert, M. et al. Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature* **383**, 550–553 (1996).
22. Dregni, A. J. et al. In vitro 0N4R tau fibrils contain a monomorphic β-sheet core enclosed by dynamically heterogeneous fuzzy coat segments. *Proc. Natl Acad. Sci. USA* **116**, 16357–16366 (2019).
23. Dregni, A. J. et al. Inclusion of the C-terminal domain in the beta-sheet core of heparin-fibrillized three-repeat Tau protein revealed by solid-state nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **143**, 7839–7851 (2021).
24. Gibbons, G. S. et al. Conformation-selective tau monoclonal antibodies inhibit tau pathology in primary neurons and a mouse model of Alzheimer's disease. *Mol. Neurodegener.* **15**, 64 (2020).
25. He, Z. H. et al. Transmission of tauopathy strains is independent of their isoform composition. *Nat. Commun.* **11**, 1–18 (2020).
26. Gullion, T. & Schaefer, J. Rotational-echo double-resonance NMR. *J. Magn. Reson.* **81**, 196–200 (1989).
27. Veshtort, M. & Griffin, R. G. SPINEVOLUTION: A powerful tool for the simulation of solid and liquid state NMR experiments. *J. Magn. Reson.* **178**, 248–282 (2006).
28. Hirschinger, J. Analytical solutions to several magic-angle spinning NMR experiments. *Solid State Nucl. Magn. Reson.* **34**, 210–223 (2008).
29. Bibow, S. et al. The dynamic structure of filamentous tau. *Angew. Chem. Int. Ed.* **50**, 11520–11524 (2011).
30. Adams, S. J., DeTure, M. A., McBride, M., Dickson, D. W. & Petrucelli, L. Three repeat isoforms of tau inhibit assembly of four repeat tau filaments. *PLoS One* **5**, e10810 (2010).
31. Dinkel, P. D., Siddiqua, A., Huynh, H., Shah, M. & Margittai, M. Variations in filament conformation dictate seeding barrier between three- and four-repeat Tau. *Biochemistry* **50**, 4330–4336 (2011).
32. Siddiqua, A. et al. Conformational basis for asymmetric seeding barrier in filaments of three- and four-repeat tau. *J. Am. Chem. Soc.* **134**, 10271–10278 (2012).
33. Dujardin, S. et al. Tau molecular diversity contributes to clinical heterogeneity in Alzheimer's disease. *Nat. Med.* **26**, 1256–1263 (2020).
34. Kim, C. et al. Distinct populations of highly potent TAU seed conformers in rapidly progressing Alzheimer's disease. *Sci. Transl. Med.* **14**, eabg0253 (2022).
35. Zhong, Q., Congdon, E. E., Nagaraja, H. N. & Kuret, J. Tau isoform composition influences rate and extent of filament formation. *J. Biol. Chem.* **287**, 20711–20719 (2012).
36. Wesseling, H. et al. Tau PTM profiles identify patient heterogeneity and stages of Alzheimer's disease. *Cell* **183**, 1699–1713 (2020).
37. von Bergen, M. et al. Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure. *Proc. Natl Acad. Sci. USA* **97**, 5129–5134 (2000).
38. 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).

39. Seidler, P. M. et al. Structure-based inhibitors of tau aggregation. *Nat. Chem.* **10**, 170–176 (2018).
40. Meisl, G. et al. In vivo rate-determining steps of tau seed accumulation in Alzheimer's disease. *Sci. Adv.* **7**, eabh1448 (2021).
41. Li, J. L., Browning, S., Mahal, S. P., Oelschlegel, A. M. & Weissmann, C. Darwinian evolution of prions in cell culture. *Science* **327**, 869–872 (2010).
42. Böckmann, A. et al. Characterization of different water pools in solid-state NMR protein samples. *J. Biomol. NMR* **45**, 319 (2009).
43. Zhou, D. H. & Rienstra, C. M. High-performance solvent suppression for proton detected solid-state NMR. *J. Magn. Reson.* **192**, 167–172 (2008).
44. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67–81 (1995).
45. Bak, M. & Nielsen, N. C. REPULSION, a novel approach to efficient powder averaging in solid-state NMR. *J. Magn. Resn* **125**, 132–139 (1997).
46. Mukrasch, M. D. et al. Structural polymorphism of 441-residue Tau at single residue resolution. *PLoS Biol.* **7**, 399–414 (2009).

Acknowledgements

This work was supported by NIH grants AG059661 to M.H. and AG17586 to V.M.-Y.L. A.J.D. is supported by an NIH Ruth L. Kirschstein Individual National Research Service Award (F31AG069418) and H.X. is supported by a grant from the BrightFocus Foundation. This study made use of NMR spectrometers at the MIT-Harvard Center for Magnetic Resonance, which is supported by NIH grant P41 GM132079. We thank Prof. Haifan Wu for fibrillizing the heparin CDN-labeled tau sample. We thank the patients and their families for donating brain tissues.

Author contributions

M.H, V.M.-Y.L, and A.J.D. designed the project. A.J.D. and P.D. expressed and purified the recombinant protein with assistance from N.E.M. H.X. and L.C. extracted and purified the human tissue samples and conducted mouse neuron assays. A.J.D. and P.D. fibrillized the brain-seeded samples and heparin-fibrillized calibration samples, with guidance from H.X. P.D. measured the TEM images. H.X. conducted sedimentation assays. A.J.D. and P.D. conducted and analyzed the solid-state NMR experiments. All authors interpreted the results of the study and wrote the paper.

Competing interests

The authors declare no competing interests.

Ethics approval

All animal protocols were approved by the University of Pennsylvania's Institutional Animal Care and Use Committee (IACUC).

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-30585-0.

Correspondence and requests for materials should be addressed to Mei Hong.

Peer review information *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

Supplementary Information

Fluent Molecular Mixing of Tau Isoforms in Alzheimer's Disease Neurofibrillary Tangles

Aurelio J. Dregni [1$], Pu Duan[1$], Hong Xu[2], Lakshmi Changolkar[2], Nadia El Mammeri[1], Virginia M.-Y. Lee[2], and Mei Hong[1]*

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[2] Department of Pathology and Laboratory Medicine, Institute On Aging and Center for Neurodegenerative Disease Research, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA

This PDF file includes:

Supplementary Tables 1 - 5
Supplementary Figures 1 to 7

Supplementary Table 1. Cases of Alzheimer's disease brains used in this solid-state NMR study.

Case no.	Disease Duration (yr)	Gender	Neuropathological diagnosis
1	11	F	Alzheimer's disease
2	6	F	Alzheimer's disease
3	7	M	Alzheimer's disease
4	9	F	Alzheimer's disease
5	11	F	Alzheimer's disease
6	9	M	Alzheimer's disease
7	8	M	Alzheimer's disease
8	16	F	Alzheimer's disease
9	3	M	Alzheimer's disease
10	8	F	Alzheimer's disease

Supplementary Table 2. Characterization of supernatant fraction from human brain extraction ("AD Seeds").

	Conc. of Tau by ELISA	Total protein Conc. by BCA	Purity (tau / total protein)	Conc. of $A\beta_{42}$ by ELISA	Conc. of $A\beta_{42}$ by ELISA	Conc. of α-synuclein by ELISA
Pooled AD-tau	0.55 µg/µl	5.28 µg/µl	10%	14.61 ng/µl	59.16 ng/µl	0.46 ng/µl

Supplementary Table 3. Morphologies of AD-seeded tau and heparin-fibrilized tau. At least 5 different particles from two or more different TEM images were sampled using ImageJ, and a total of at least 10 readings were averaged to give each of the statistics below.

Samples	Twisted fibrils			Straight fibrils
	Crossover length	Narrow width	Wide width	Width
[AD]4R^{N}3R^{C}	86.9 ± 10.6 nm	9.9 ± 1.6 nm	25.1 ± 1.8 nm	15.2 ± 0.9 nm
[AD]4R^{NC}	99.3 ± 9.4 nm	10.7 ± 1.6 nm	24.3 ± 1.7 nm	15.9 ± 1.3 nm
[AD]3R^{NC}	100.7 ± 16.5 nm	10.6 ± 1.0 nm	24.7 ± 2.7 nm	16.3 ± 1.0 nm
[AD]3R^{N}4R^{C}	93.8 ± 18.0 nm	10.0 ± 1.1 nm	25.5 ± 1.9 nm	16.0 ± 1.0 nm
Heparin-0N4R	16.1 ± 1.6 nm, long straight filaments			
Heparin-0N3R	22.9 ± 0.6 nm, straight ribbons			
Hep-0N3R/0N4R mix	13.8 ± 1.3 nm, long straight filaments			

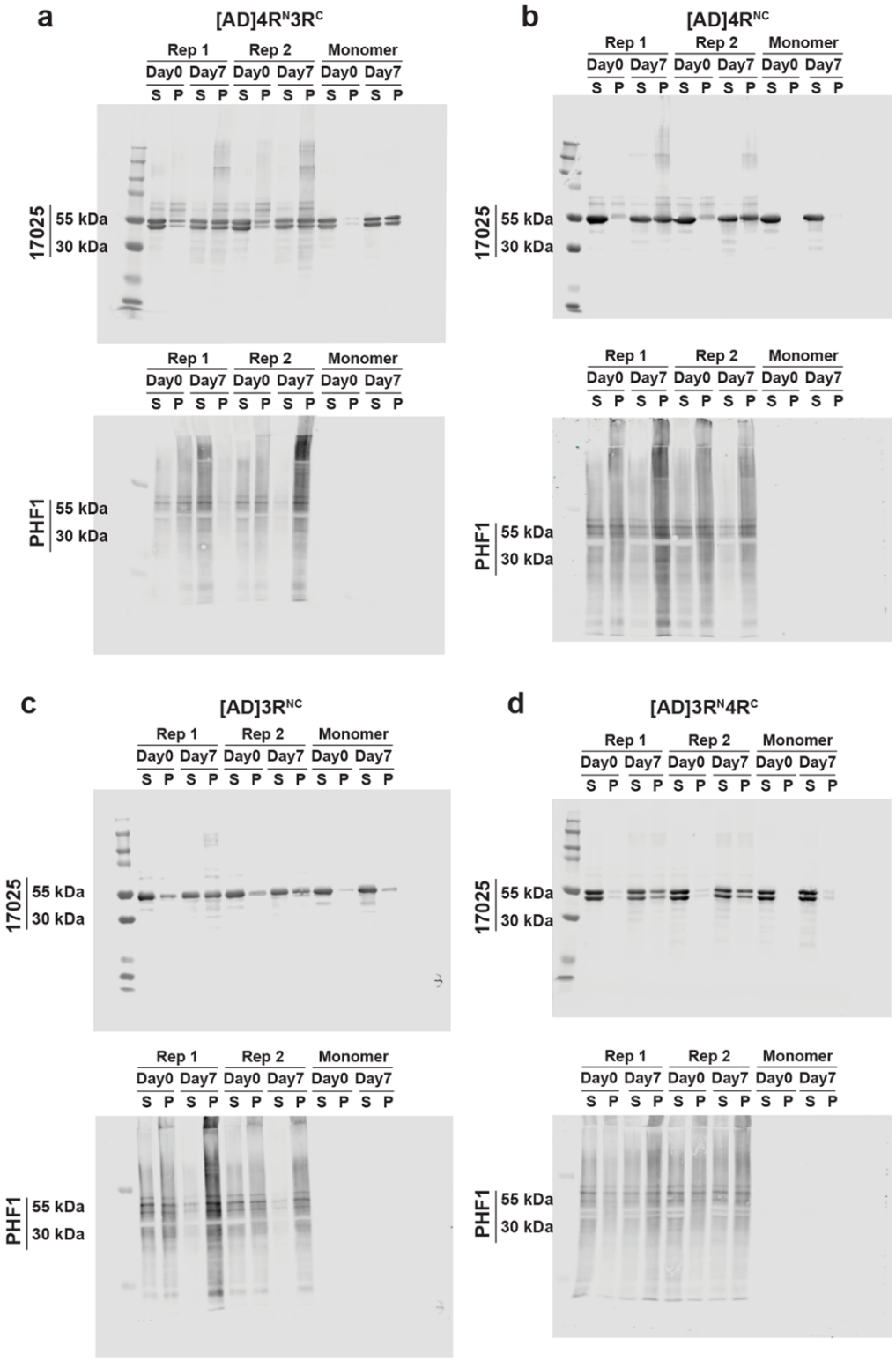
Supplementary Table 4. Concentrations of tau added to fibril growth mixtures and validation by the 1D NMR spectral quantification.

Samples	Tau monomers for fibrillization	Composition of AD seeded samples from 1D NMR spectral quantification
[AD]4R^{N}3R^{C}	18 μM ^{2}H,^{13}C-0N3R + 18 μM ^{2}H,^{15}N-0N4R	#1: 40.0% 3R, 60.0 ± 3.0% 4R, #2: 37.4% 3R, 62.6 ± 3.0% 4R
[AD]4R^{NC}	36 μM ^{2}H,^{13}C,^{15}N-0N4R	100% 4R
[AD]3R^{NC}	36 μM ^{2}H,^{13}C,^{15}N-0N3R	100% 3R
[AD]3R^{N}4R^{C}	18 μM ^{2}H,^{13}C-0N4R + 18 μM ^{2}H,^{15}N-0N3R	#1: 39.2 ± 3.0% 3R, 60.8% 4R #2: 37.9 ± 3.0% 3R, 62.1% 4R
100:0 4R	10 μM ^{2}H,^{13}C-0N4R + 0 μM ^{2}H,^{15}N-0N4R	
70:30 4R	7 μM ^{2}H,^{13}C-0N4R + 3 μM ^{2}H,^{15}N-0N4R	
60:40 4R	6 μM ^{2}H,^{13}C-0N4R + 4 μM ^{2}H,^{15}N-0N4R	
50:50 4R	5 μM ^{2}H,^{13}C-0N4R + 5 μM ^{2}H,^{15}N-0N4R	
40:60 4R	4 μM ^{2}H,^{13}C-0N4R + 6 μM ^{2}H,^{15}N-0N4R	
30:70 4R	3 μM ^{2}H,^{13}C-0N4R + 7 μM ^{2}H,^{15}N-0N4R	

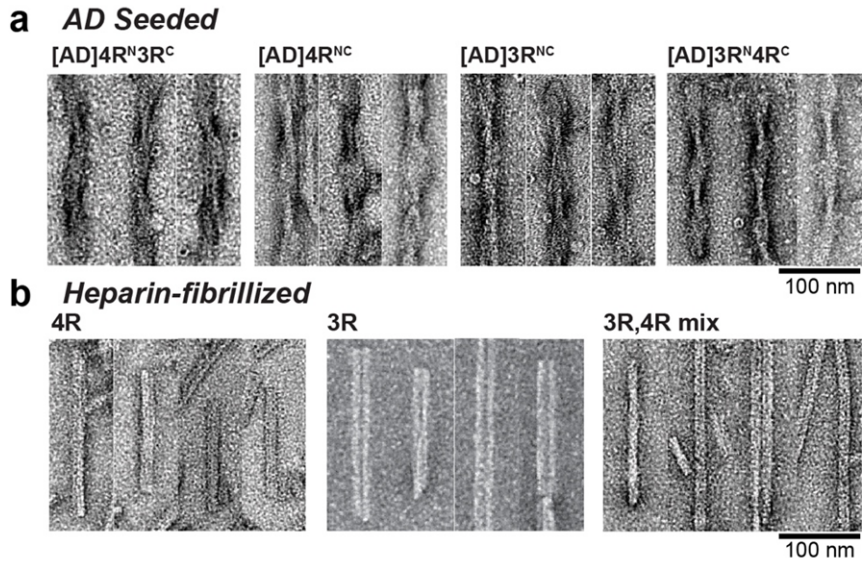
Supplementary Table 5. NMR experimental parameters in this study. All experiments were conducted using a Bruker 1.3 mm HCN probe on a Bruker Avance III HD 600 MHz spectrometer or Avance II 800 MHz spectrometer. The rf field strength for ^{1}H, ^{13}C and ^{15}N channel pulses are 83.3 kHz, 62.5 kHz and 50 kHz, respectively.

Experiment	NMR Parameters	Experimental Time per sample
1D ^{13}C DP	B_0 = 14.1 T, ns = 2048 or 6144, τ_{rd} = 5 s, τ_{dw} = 10 µs, τ_{acq} = 30.7 ms, $\nu_{H,DEC}$ = 130 kHz, ν_{MAS} = 20 kHz, $T_{bearing}$ = 278 K, δ_{H2O} = 4.89 ppm, d1 = 5 s	3 hr or 9 hr
1D J-hnH	B_0 = 14.1 T, ns = 256 or 512, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 92.2 ms, τ_{INEPT} = 10 ms, $\tau_{MISSISSIPPI}$ = 0.3 s, $\nu_{H,DEC}$ = 130 kHz, $\nu_{C,DEC}$ = 3.6 kHz, $\nu_{N,DEC}$ = 3.6 kHz, ν_{MAS} = 20 kHz, $T_{bearing}$ = 293 K, δ_{H2O} = 4.77 ppm.	10 min or 20 min
1D dipolar hnH	B_0 = 14.1 T, ns = 256 or 512, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 23.0 ms, τ_{HN} = 1250 µs, τ_{NH} = 400 µs, $\tau_{MISSISSIPPI}$ = 0.3 s, $\nu_{C,DEC}$ = 3.6 kHz, $\nu_{N,DEC}$ = 3.6 kHz, ν_{MAS} = 20 kHz, $T_{bearing}$ = 278 K, δ_{H2O} = 4.89 ppm.	10 min or 20 min
^{1}H-detected ^{15}N-^{13}C REDOR	B_0 = 14.1 T, ns = 256*6*2 to 256*8*2 for each REDOR time, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 23.0 ms, τ_{HN} = 1250 µs, τ_{NH} = 400 µs, $\tau_{MISSISSIPPI}$ = 0.3 s, τ_{REDOR} = 0.1 ms to 25 ms (28 REDOR time in total), $\nu_{H,DEC}$ = 130 kHz, $\nu_{C,DEC}$ = 3.6 kHz, $\nu_{N,DEC}$ = 3.6 kHz, ν_{MAS} = 20 kHz, ν_{C} = 62.5 kHz, $T_{bearing}$ = 278 K, δ_{H2O} = 4.89 ppm.	2-3 days per sample
2D J-hNH	B_0 = 14.1 T, ns = 64 to 160, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 92.2 ms, $t_{1,inc}$ = 274 µs, $t_{1,max}$ = 70.2 ms, τ_{INEPT} = 10 ms, $\tau_{MISSISSIPPI}$ = 0.3 s, $\nu_{H,DEC}$ = 130 kHz, $\nu_{C,DEC}$ = 3.6 kHz, $\nu_{N,DEC}$ = 3.6 kHz, ν_{MAS} = 20 kHz, $T_{bearing}$ = 293 K, δ_{H2O} = 4.77 ppm.	~2 days per sample
2D dipolar hNH	B_0 = 18.8 T (Heparin-4R^{NC}, [AD]4R^{NC}) or 14.1 T (others), ns = 400 to 600, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 38.4 ms, $t_{1,inc}$ = 225 µs (Heparin-4R^{NC}, [AD]4R^{NC}) or 300 µs (others), $t_{1,max}$ = 30 ms, τ_{HN} = 1250 µs, τ_{NH} = 400 µs, $\tau_{MISSISSIPPI}$ = 0.3 s, $\nu_{H,DEC}$ = 10 kHz, $\nu_{C,DEC}$ = 10 kHz, $\nu_{N,DEC}$ = 10 kHz, ν_{MAS} = 55 kHz, $T_{bearing}$ = 250 K, δ_{H2O} = 4.89 ppm.	2-3 days per sample
3D hCANH	B_0 = 14.1 T, ns = 24, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 38.4 ms, $t_{1,inc}$ = 160 µs, $t_{1,max}$ = 4.8 ms, $t_{2,inc}$ = 300 µs, $t_{2,max}$ = 10.5 ms, τ_{HC} = 2 ms, τ_{CN} = 8 ms, τ_{NH} = 400 µs, $\tau_{MISSISSIPPI}$ = 0.2 s, $\nu_{H,DEC}$ = 10 kHz, $\nu_{H,CN}$ = 0 kHz, $\nu_{C,DEC}$ = 10 kHz, $\nu_{N,DEC}$ = 10 kHz, ν_{MAS} = 55 kHz, $T_{bearing}$ = 247 K, δ_{H2O} = 4.89 ppm.	2 days
3D hCA(co)NH	B_0 = 14.1 T, ns = 40, τ_{rd} = 1.5 s, τ_{dw} = 15 µs, τ_{acq} = 23.0 ms, $t_{1,inc}$ = 160 µs, $t_{1,max}$ = 4.8 ms, $t_{2,inc}$ = 300 µs, $t_{2,max}$ = 10.5 ms, τ_{HC} = 2 ms, τ_{DREAM} = 8 ms, τ_{CN} = 8 ms, τ_{NH} = 400 µs, $\tau_{MISSISSIPPI}$ = 0.2 s, $\nu_{H,DEC}$ = 10 kHz, $\nu_{H,CN}$ = 0 kHz, $\nu_{C,DEC}$ = 10 kHz, $\nu_{N,DEC}$ = 10 kHz, ν_{MAS} = 55 kHz, $T_{bearing}$ = 247 K, δ_{H2O} = 4.89 ppm.	3.5 days

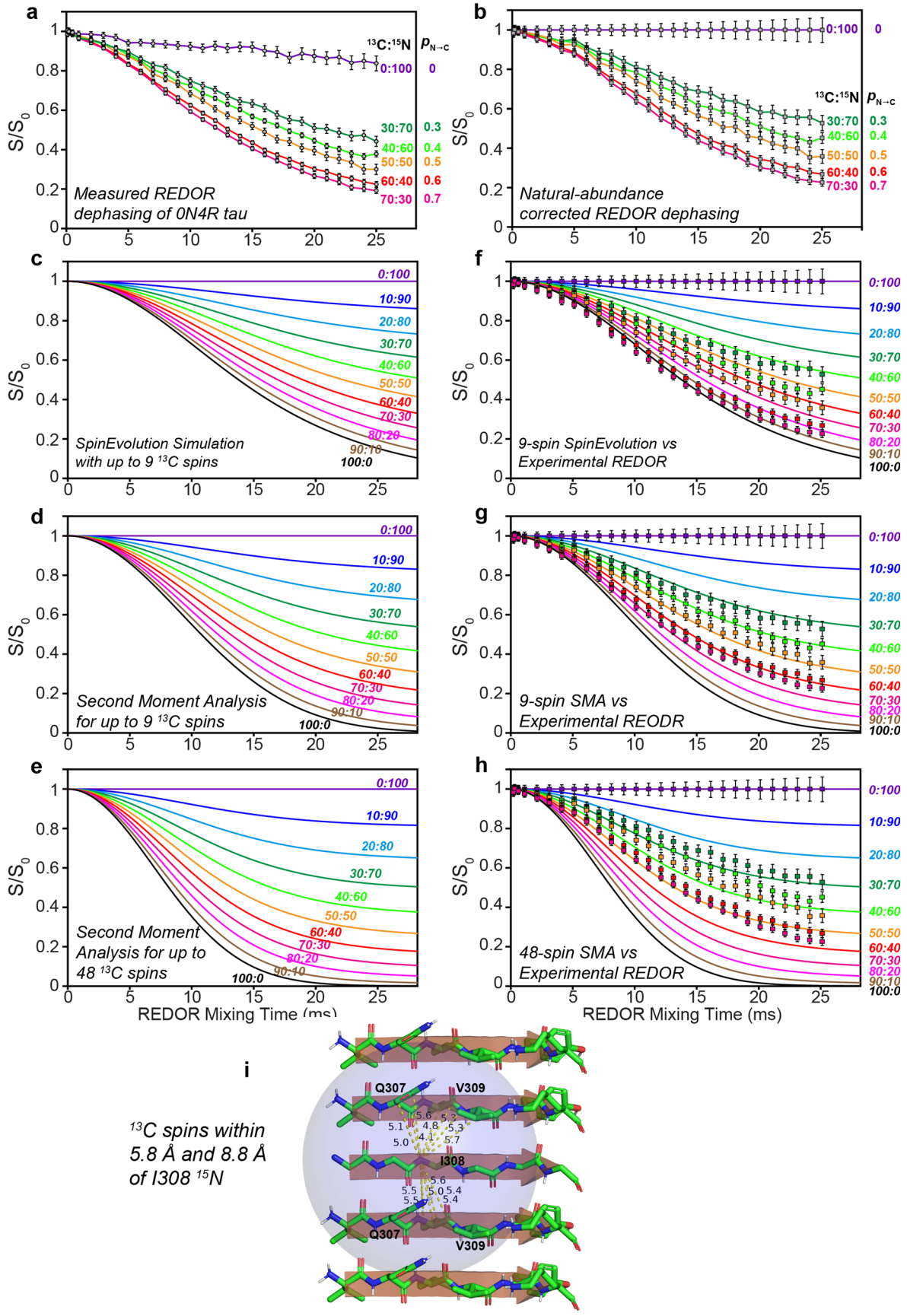
Definitions of symbols: B_0 = magnetic field; ns = number of scans per free induction decay; τ_{rd} = recycle delay; $t_{1,max}$ = maximum t_1 evolution time; $t_{1,inc}$ = t_1 increment; $t_{2,max}$ = maximum t_2 evolution time; $t_{2,inc}$ = t_2 increment; τ_{dw} = dwell-time in the direct dimension; τ_{acq} = maximum acquisition time in the direct dimension; $\nu_{X,DEC}$ = X channel (^{1}H, ^{13}C or ^{15}N) rf field strength for decoupling during acquisition; $\nu_{H,CN}$ = ^{1}H rf field strength for decoupling during ^{13}C-^{15}N SPECIFIC-CP; $T_{bearing}$ = thermocouple-reported bearing gas temperature; δ_{H2O} = ^{1}H chemical shift of water; ν_{MAS} = MAS frequency; τ_{HX} = ^{1}H-X cross polarization contact time; τ_{CN} = ^{13}C-^{15}N SPECIFIC-CP contact time; τ_{DREAM} = ^{13}C-^{13}C DREAM contact time; τ_{NH} = ^{15}N-^{1}H cross polarization contact time; $\tau_{MISSISSIPPI}$ = duration of MISSISSIPPI solvent suppression period, τ_{REDOR} = ^{15}N-^{13}C REDOR mixing time; ν_{C} = ^{13}C rf field strength for 180° pulses during REDOR.



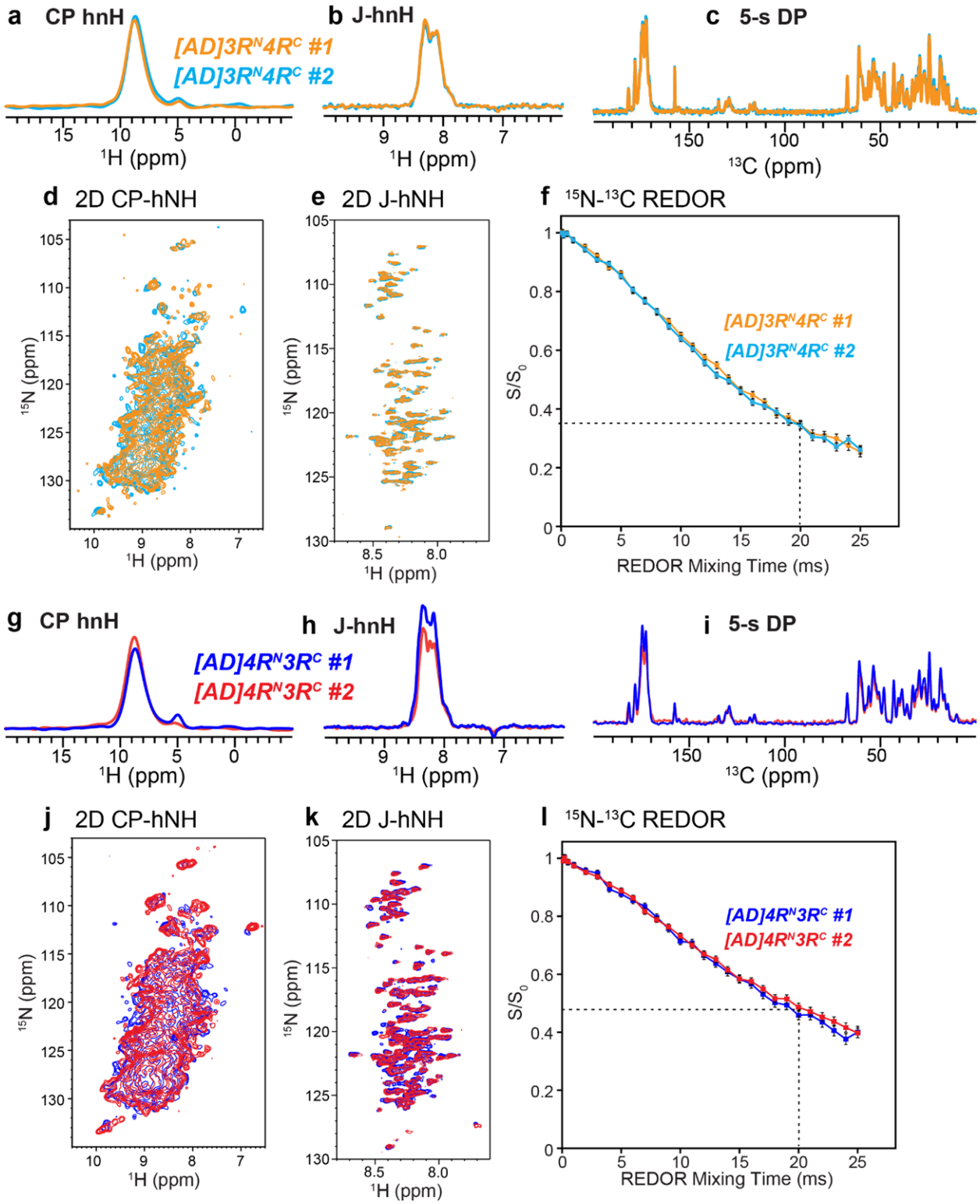
Supplementary Fig. 1. Full Western Blot images from Fig. 2a. Tau species in (**a**) [AD]4R^{N}3R^{C}, (**b**) [AD]4R^{NC}, (**c**) [AD]3R^{NC} and (**d**) [AD]3R^{N}4R^{C} were revealed by the 17025 anti-total tau antibody and PHF1 anti-phospho-tau antibody.



Supplementary Fig. 2. **TEM Images of tau filaments aligned for comparison.** (**a**) The majority of fibrils in AD-seeded samples are twisted filaments that have uniform width and crossover distance, and that match the dimensions previously reported for AD paired-helical filaments (PHF) [45]. (**b**) Heparin-fibrillized tau containing 4R tau only, 3R tau only, and both 4R and 3R tau. These fibrils are nearly exclusively straight and are distinct from AD seeded tau filaments. At least five different images from different regions of each grid were obtained for each sample.

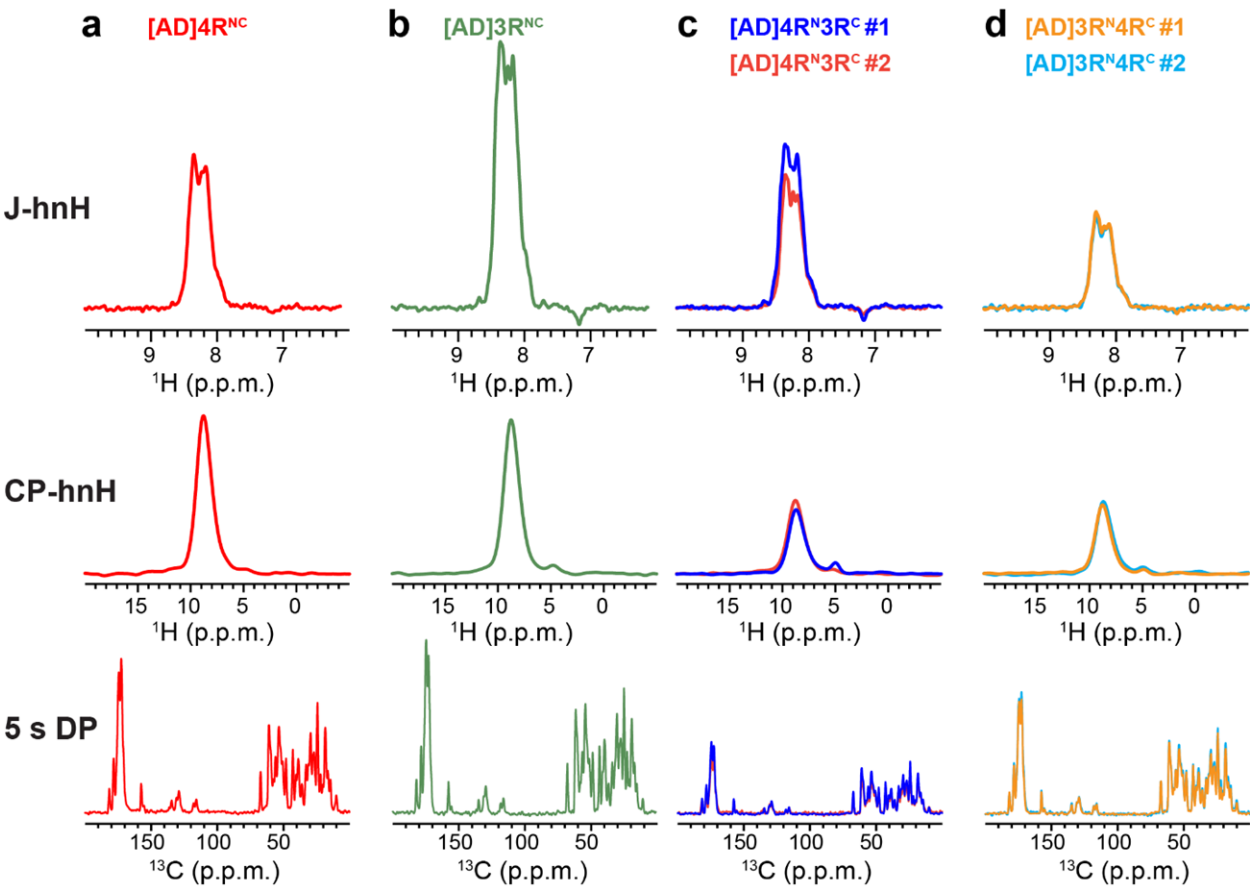

a
Measured REDOR dephasing of 0N4R tau
^{13}C:^{15}N $p_{N\to c}$
0:100 0
30:70 0.3
40:60 0.4
50:50 0.5
60:40 0.6
70:30 0.7
b
Natural-abundance corrected REDOR dephasing
^{13}C:^{15}N $p_{N\to c}$
0:100 0
30:70 0.3
40:60 0.4
50:50 0.5
60:40 0.6
70:30 0.7
c
SpinEvolution Simulation with up to 9 ^{13}C spins
0:100
10:90
20:80
30:70
40:60
50:50
60:40
70:30
80:20
90:10
100:0
d
Second Moment Analysis for up to 9 ^{13}C spins
e
Second Moment Analysis for up to 48 ^{13}C spins
f
9-spin SpinEvolution vs Experimental REDOR
g
9-spin SMA vs Experimental REODR
h
48-spin SMA vs Experimental REDOR
S/S$_0$
REDOR Mixing Time (ms)
i
^{13}C spins within 5.8 Å and 8.8 Å of I308 ^{15}N
Q307
V309
I308
Q307
V309
5.1
5.6
4.8
5.3
5.0
4.1
5.7
5.6
5.5
5.0
5.4
5.5
5.4

Supplementary Fig. 3. Measured and simulated ^{15}N-^{13}C REDOR dephasing curves of mixed labeled 0N4R tau with varying ratios of ^{13}C, ^{2}H-labeled protein and ^{15}N, ^{2}H-labeled protein. (**a**) Measured REDOR dephasing of ^{13}C and ^{15}N mixed labeled 0N4R tau with varying mixing ratios, reproduced from Fig. 3e. Natural abundance ^{13}C dephasing to ^{15}N-labeled protein is present in these data. Error bars represent the propogated uncertainty from the spectral signal-to-noise ratios. (**b**) ^{13}C natural-abundance corrected REDOR data, obtained by dividing the measured S/S_0 values of each sample by the S/S_0 values of the 100% ^{15}N-labeled tau. Error bars represent the propogated uncertainty from the spectral signal-to-noise ratios. (**c**) SpinEvolution simulation using up to 9 nearest ^{13}C spins to the I308 amide ^{15}N. The natural abundance ^{13}C nuclei in the ^{15}N-labeled monomer are ignored in these simulations. (**d**) Second moment analysis (SMA) of REDOR dephasing using the same ^{13}C spins as in (**b**). (**e**) Second moment analysis of REDOR dephasing using the nearest 48 ^{13}C spins within 8.8 Å of the ^{15}N nucleus. The second moment simulations ignore ^{13}C finite-pulse effects, ^{13}C isotropic and anisotropic chemical shifts, and ^{13}C-^{13}C dipolar couplings. (**f-h**) overlay of the measured natural-abundance corrected REDOR data in (**b**) with the SpinEvolution simulation (**c**) and second moment analysis (**d-e**). (**i**) ^{13}C spins within 5.8 Å (dashed lines) or 8.8 Å (purple sphere) of I308 amide ^{15}N in the cryoEM structure of AD PHF tau (PDB: 5O3L) [4]. The two neighboring chains that sandwich the central chain with an I308 ^{15}N label are ^{13}C-labeled either in one or both chains.

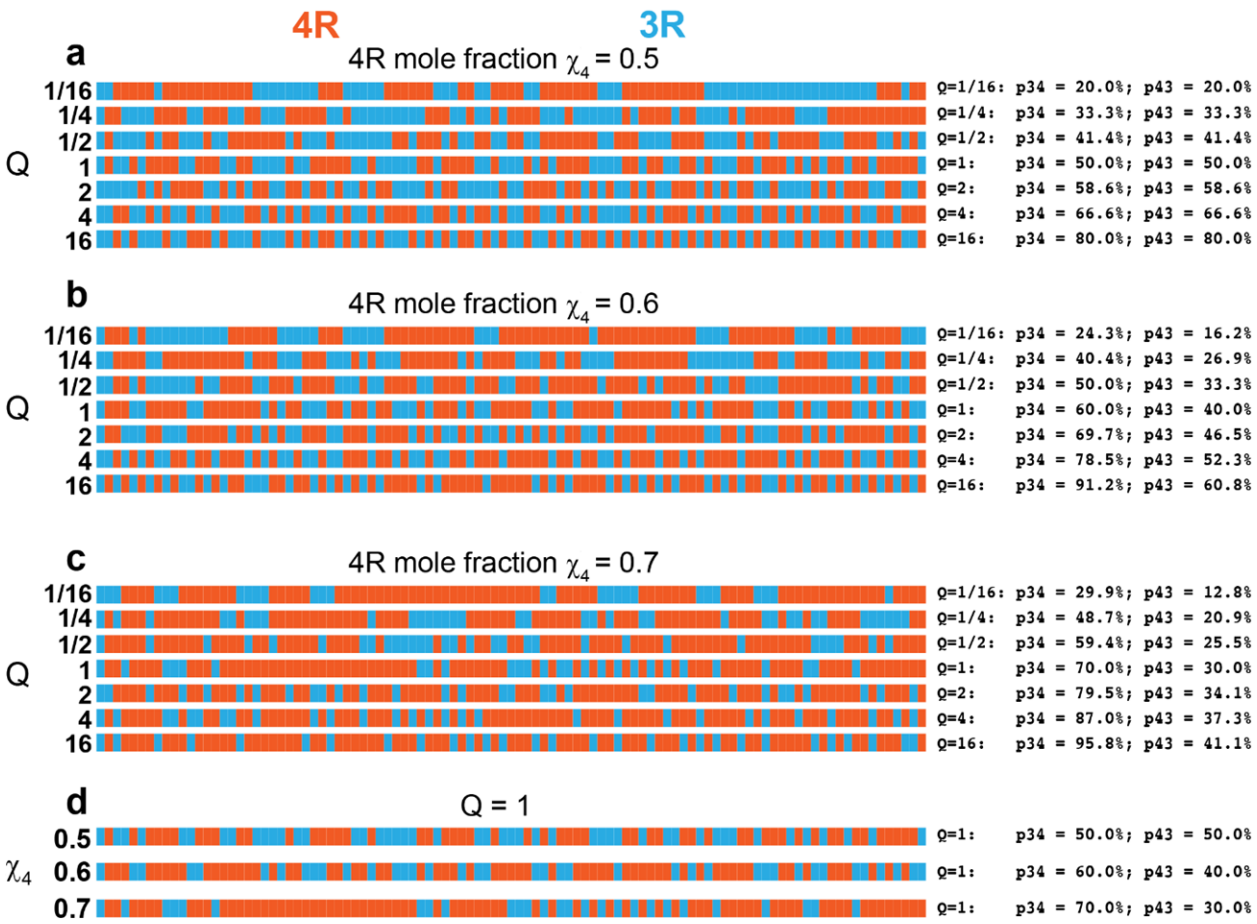


Supplementary Fig. 4. AD seeded mixed isoform tau filaments have highly reproducible structure and dynamics. (**a-f**) Comparisons of the fingerprint spectra and REDOR data of two independent samples of AD-seeded ^{15}N-labeled 3R tau mixed with ^{13}C-labeled 4R tau (1:1). (**g-l**) Comparisons of fingerprint spectra and REDOR data of two independent AD seeded ^{15}N-labeled 4R tau mixed with ^{13}C-labeled 3R tau (1:1). Each sample was seeded separately using separately

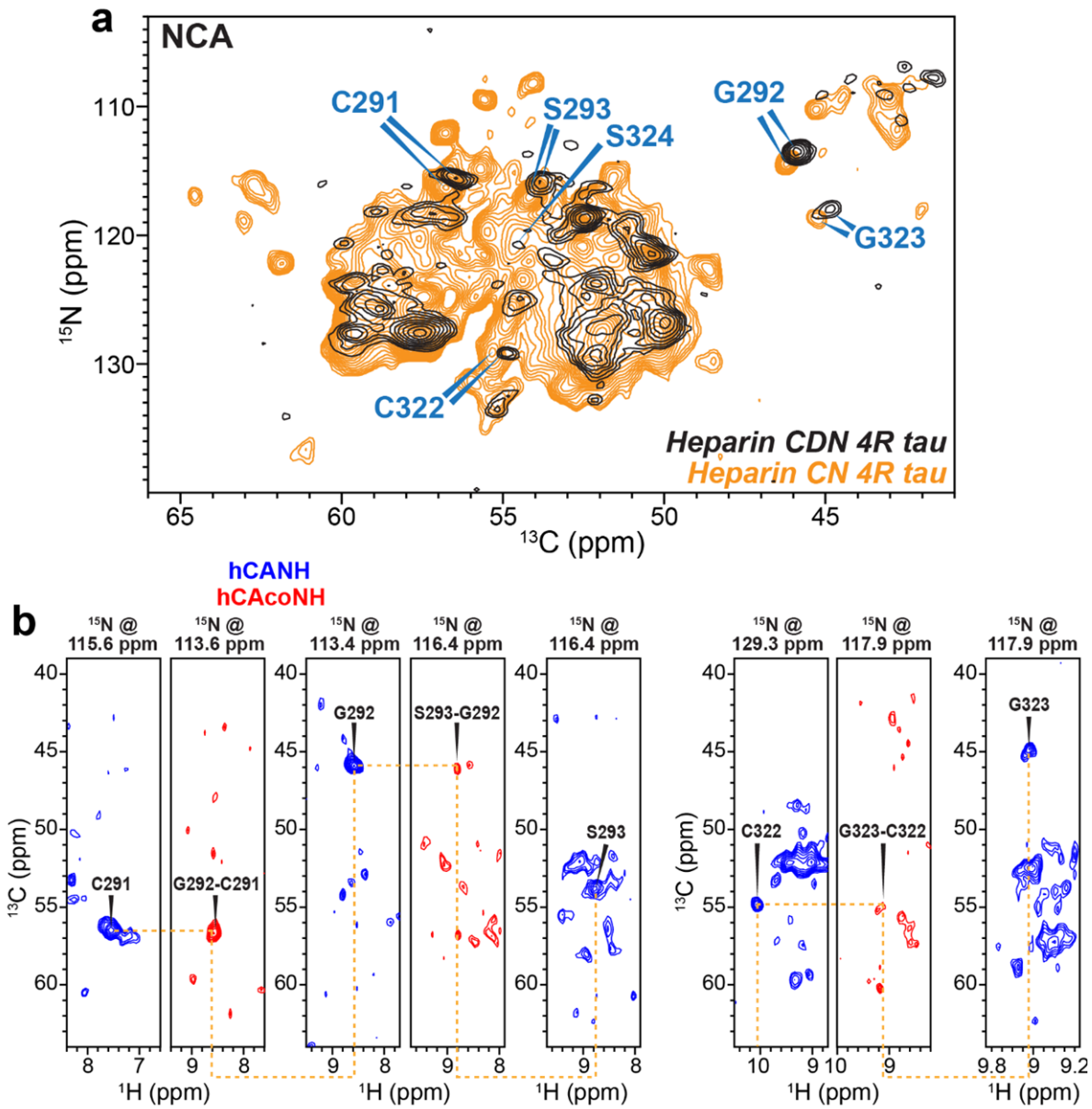
purified tau monomers. (**a-c**) 1D CP-hnH, J-hnH, and ^{13}C DP MAS spectra of duplicates #1 and #2 of [AD]3R^{N}4R^{C} tau. (**d-e**) 2D CP-hNH and J-hNH spectra of the same duplicates. (**f**) ^{1}H-detected ^{15}N-^{13}C REDOR dephasing of the duplicates. Similar comparisons between the duplicates of [AD]4R^{N}3R^{C} are shown in (**g-l**).



Supplementary Fig. 5. 1D ^{1}H-detected J-hnH spectra, ^{1}H-detected CP-hnH spectra, and ^{13}C DP-MAS spectra of AD-tau seeded tau filaments for quantification of 4R and 3R tau incorporation levels into the filaments. (**a**) [AD]4R^{NC} sample. (**b**) [AD]3R^{NC} sample. (**c**) Duplicates #1 and #2 of [AD]4R^{N}3R^{C} sample, which contains a 1:1 mixture of ^{15}N, ^{2}H-labeled 0N4R tau and ^{13}C, ^{2}H-labeled 0N3R tau. (**d**) Duplicates #1 and #2 of [AD]4R^{N}3R^{C} sample, which contains a 1:1 mixture of ^{15}N, ^{2}H-labeled 0N3R tau and ^{13}C, ^{2}H-labeled 0N4R tau.



Supplementary Fig. 6. Simulated fibril mixing schemes for varying mole fractions of 4R tau (χ_4) and mixing quotients Q. Individual 100-monomer fibrils made of 3R (blue) or 4R (orange) monomers are shown; each monomer is a vertical line. (**a**) Simulated mixing as a function of Q for $\chi_4 = 0.5$. (**b**) Simulated mixing as a function of Q for $\chi_4 = 0.6$. (**c**) Simulated mixing as a function of Q for $\chi_4 = 0.7$. (**d**) Simulated mixing as a function of χ_4 for Q = 1.



Supplementary Fig. 7. Assignment of residues ^{291}CGS^{293} and ^{322}CGS^{294} in heparin-fibrilized 4R^{NC} tau based on 3D hCANH and hCA(co)NH spectra. (**a**) 2D NC projection from the 3D hCANH spectrum of the heparin-fibrillized CDN-labeled 4R tau (black), overlaid with the previously reported 2D NCA spectrum of heparin-fibrillized CN-labeled 4R tau (orange) [21]. Residues ^{291}CGS^{293} and ^{322}CGS^{294} in both spectra are marked in cyan. The small chemical shift differences between the two samples can be attributed to deuterium isotope effects and slight difference in experimental temperatures. (**b**) CAH strips extracted at the ^{15}N chemical shifts of residues ^{291}CGS^{293} and ^{322}CG^{293} from 3D the hCANH (blue) and hCA(co)NH (red) spectra of the CDN-labeled 4R tau sample. Residues ^{291}CGS^{293} and ^{322}CG^{323} have the strongest peak intensities in these 3D ^{1}H-detected spectra, consistent with the observation of CN-labeled 4R tau sample in previously reported ^{13}C-detected 3D spectra [21].

nature portfolio

Peer Review File

Fluent Molecular Mixing of Tau Isoforms in Alzheimer's Disease Neurofibrillary Tangles

Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript by Dregni et. al. focused on tackle the problem on how 4R and 3R in AD tau tangles were mixed. It is a very difficult task. The authors first extracted the AD tau tangles from AD patients as the seeds, then the fibril were amplified and solid-state NMR REDOR was applied. This is a novel application using REDOR to obtain the information on 4R/3R mixing. The statistical explanation was written clearly. I think it is a very good paper.
I only have some minor suggestions.
1. I think if the REDOR technique applicable here, it will also have other pre-requirements. Since 13C-15N polarization transfer between the intermolecular hydrogen bonds is the key in this REDOR application, we have to assume the fibril core structure is similar between 4R and 3R in the number of residues involved in the intermolecular hydrogen bonding. The hydrogen bonding pattern would not change in the intermolecular interactions between different type of molecules once 4R and 3R are mixed. So no matter 4RN-3RC, or 3RN-4RC or the pure type, the number of these intermolecular hydrogen bonding is not changed. It may be the case here. The author did not explain much on the difference between 3R and 4R in the structure. I think since there were many reports on different types of tau fibril structures, the author cannot assume we know 3R and 4R so well. More information on the structures on 3R and 4R are needed. The brain derived tau structure AD PHF fibril (PDB: 5O3L) was used in this paper. I am not sure the structure in this paper would be the same as this published structure.
2. 10 patients' brain extract were pooled together and used. Since there was difference on these patients' onset ages and disease duration, and clinical diagnosis, I am not sure whether these extracts from different patient would represent the same type of tau fibril structure or not. However, their seeding structure seemed repeatable, suggesting the seeds only produce one type of structure. However, it is not clear for me whether the structure would be similar to the published 5O3L structure.
3. What are those circle structures in TEM images (extended data fig 2, [AD]3R and seeds after 7 days)
4. PHF1 antibody showed a distribution of bands with different sizes showing in the SDS-PAGE gel. Could you please give some explanation on these?

Reviewer #2 (Remarks to the Author):

The paper by Dregni et al represents a sound and significant contribution to the understanding of fibril formation relating to dementia. Supported by extensive biophysical and EM characterization, they present solid-state NMR data which shed new detailed light on the organization of monomers in fibrils with particular focus on how 4R and 3R tau proteins mixes in the fibril upon seeding with AD brain-derived tau – differing significantly from fibrils assembled using heparin.

Overall, the experiments and conclusions seem sound, and the outcome important for the community. As such I recommend the paper being accepted for publication in Nature Communications upon consideration of

(1) I recommend authors more carefully address the significant discrepancy between experimental and simulated REDOR curves as given in Extended Data Fig. 3. I do agree that comparison with the experimental curves obtained for samples with different labelling ratios may provide a "safe" ground for the conclusion – but still it is puzzling that the simulations do not better match the experimental data. The dephasing in the experimental curves seems to be significantly faster both in the initial and later parts of the dephasing period – which may not be fully accounted for by the explanation of non-included long-range effects. Could it be associated with non-perfect decoupling or other? On the experimental side, can the comparison with experimental data for differently labelled 0N4R

heparin-fibrillized tau samples be non-adequate for AD-brain derived mixed 4R and 3R samples? As all curves shift significantly (e.g. the experimental 40:60 curve resembles the simulated 60:40 curve), I recommend the authors explore this in more detail (including commenting on effects from fuzzy fibril regions), considering the fundamental importance of the REDOR decay profiles in building the conclusions of the paper.
(2) In light of (1), the authors should discuss the accuracy of the incorporation level percentages. For example, it is argued that a p_4->3 value is 0.37 +/- 0.03 and khi_4 is 60 +/- 2%.

Reviewer #3 (Remarks to the Author):

Recent breakthroughs in cryo-EM technology have allowed the visualization of atomic structures of tau filaments extracted from Alzheimer's disease patient brains; however, how different 3R to 4R isoforms are molecularly mixed in the fibrils remains unknown. In this manuscript, Aurelio and colleagues have utilized solid-state NMR to determine how 3R and 4R tau isoforms could potentially be recruited into the AD fibrils. In addition, the authors found that the ratio of 4R:3R incorporated into AD fibrils is in a proportion of 60:40 with a homotypic preference, and a structural difference exists between heparin-induced tau fibrils and seeded tau fibrils from AD samples.
To characterize how tau isoforms are structurally incorporated into AD seeds, the authors pooled insoluble fractions from the brains of 10 AD patients and then added 15N or 13C labeled recombinant tau monomers into the mixture. Previous studies have shown that tau seeds are heterogeneous, and different tau species could exist between patients even if their Braak staging is similar. If the pooled solution of AD seeds contained different tau fibril strains, how tau monomers are recruited into the seeds may differ depending on the fibril imaged.
The authors' findings of the fluent mixing of 3R and 4R into the AD seed is novel; however, this incorporation pattern may occur differently in vivo. Because the tau monomers the authors used were purified from bacterial cells, they were not post-translationally modified like in human brains. These post-translational modifications, such as phosphorylation, may play an essential role in electrostatic interactions and steric hindrances between the tau monomers and insoluble tau fibrils, which are missing in the author's system. While the use of human samples as seeds may give this study more human relevance, the fact that these measurements are made on in vitro derived fibrils should be more clearly stated in the abstract/conclusions.
Specific Comments:
1. While the introduction provided succinctly describes the need for the study, it is extremely brief.

2. The order of how the result is presented does not follow the order of the figures. The authors defined the mixture of AD fibrils and 4R and 3R tau for Figure 3 after they were used in figure 1 results.

3. In Figure 1.b, why is the [AD] 3RNC western blot not shown in the main figure, but moved to the supplemental data section?

4. In Figure 1.b and extended data figure 1, the authors showed that recombinant tau monomers are incorporated into insoluble mass after 7 days and stained the western blot with total tau antibody and phospho-tau antibody. It would be informative to address why more phosphorylated tau are present (more intense PHF1 bands) after incubating the tau seeds with monomers. Are kinases present in their mixture ?

5. In the primary neuron seeding experiment, that the localization of the AD tau or seeded tau fibrils is not distinguished (for example within the neurons or sitting on top of the neurons). In addition, the images of the neurons are not clear. Consider adding neuronal staining to show localization of tau fibrils and the purity of the culture.

6. Was an equal portion of each brain used to make the pooled insoluble tau sample?

Minor Comments:

1. In the bar graphs of Figure1D, what are the error bars showing?

2. In Figure 2a: TEM image quality makes it difficult to see. Extended data figure 2 of TEM images seemed to have a better quality images compared to Figure 2.

Reviewer #1 (Remarks to the Author):

The manuscript by Dregni et. al. focused on tackle the problem on how 4R and 3R in AD tau tangles were mixed. It is a very difficult task. The authors first extracted the AD tau tangles from AD patients as the seeds, then the fibril were amplified and solid-state NMR REDOR was applied. This is a novel application using REDOR to obtain the information on 4R/3R mixing. The statistical explanation was written clearly. I think it is a very good paper.

I only have some minor suggestions.

1. I think if the REDOR technique applicable here, it will also have other pre-requirements. Since 13C-15N polarization transfer between the intermolecular hydrogen bonds is the key in this REDOR application, we have to assume the fibril core structure is similar between 4R and 3R in the number of residues involved in the intermolecular hydrogen bonding. The hydrogen bonding pattern would not change in the intermolecular interactions between different type of molecules once 4R and 3R are mixed. So no matter 4RN-3RC, or 3RN-4RC or the pure type, the number of these intermolecular hydrogen bonding is not changed. It may be the case here. The author did not explain much on the difference between 3R and 4R in the structure. I think since there were many reports on different types of tau fibril structures, the author cannot assume we know 3R and 4R so well. More information on the structures on 3R and 4R are needed. The brain derived tau structure AD PHF fibril (PDB: 5O3L) was used in this paper. I am not sure the structure in this paper would be the same as this published structure.

Thanks for this interesting question. We have now added a paragraph in the Discussion to address this question:

" The 4R and 3R tau mixing probabilities that we measured from the REDOR data are based on the assumption that the number of residues that contributes to the 13-sheet core is the same for 4R and 3R tau in the AD-seeded fibril mixture. If the number of residues were different, for example, if 4R tau incorporates more 13-sheet residues into the fibril core than 3R tau, then the [AD]$4R^N3R^C$ REDOR S/S_0 intensities would be biased high, and the true $p_{4->3}$ value would be larger than the measured value. Conversely, [AD]$3R^N4R^C$ REDOR S/S_0 intensities would be biased low, and the true $p_{3->4}$ value would be smaller than the measured value. However, three lines of evidence argue against the different fibril core size between the two isoforms. First, tau filaments from multiple AD brains all show a single 13-sheet core, without any sign of structural heterogeneity {Fitzpatrick, 2017 #3;Falcon, 2018 #82}. Second, the 2D hNH dipolar and INEPT spectra (**Fig. 4**) show similar peak patterns for the two mixed isoform samples, which are also similar to the single isoform samples. Thus, the AD tau seeds are able to propagate by either isoform to develop the same structure. Third, the AD tau seeds are well amplified in our current study, which would be difficult to achieve if a monomer of the opposite isoform is added to the growing fibril with a different core size than the current isoform.

2. 10 patients' brain extract were pooled together and used. Since there was difference on these patients' onset ages and disease duration, and clinical diagnosis, I am not sure whether these extracts from different patient would represent the same type of tau fibril structure or not. However, their seeding structure seemed repeatable, suggesting the seeds only produce one type of structure. However, it is not clear for me whether the structure would be similar to the published 5O3L structure.

We have now added these sentences in the Discussion:

" Because these NMR data were obtained from AD seeds pooled from multiple patient brains, the 60 : 40 mixing and the probabilities we measured reflect the general properties of AD brain tau fibrils instead of being specific to one patient brain. The AD-seeded tau samples show highly reproducible 2D NMR spectra as well as reproducible REDOR dephasing, indicating that there is no detectable molecular structural variation among the multiple AD brains. This is consistent with the recent cryoEM observation that multiple AD brains exhibit the same C-shaped

conformation for the tau fibril core {Falcon, 2018 #82}. These results together support the notion that AD brain tau tangles have a single molecular structure. "

Although we do not attempt to determine the molecular structure of AD seeded tau studied here, the similarity of the twisted morphology of the seeded samples to AD PHF filaments and the reproducibility of AD brain tau structures in multiple brains seen by Goedert and Scheres suggest that the seeded samples likely have the same structure as 5O3L.

0. What are those circle structures in TEM images (extended data fig 2, [AD]3R and seeds after 7 days)

This is a good question. We do not know exactly what these circular objects are, but note that they are ubiquitous in the TEM images of AD brain extracted tau, for example, see an image from the Crowther's 1991 PNAS paper. The objects marked by yellow arrows have an O.D. of ~13 nm and an I.D. of ~5 nm, and are absent in *in vitro* he parin-fibrillized tau samples or tau samples made in the presence of protease inhibitors. Therefore, we suspect that these objects are impurities from the AD brain extraction procedure.

1. PHF1 antibody showed a distribution of bands with different sizes showing in the SDS-PAGE gel. Could you please give some explanation on these?

We assume the reviewer is referring to the immunoblots in Figure 1b and Extended Data Figure 1. We have now added this clarification in the experimental section:

" The 17025 antibody was raised with recombinant human tau, thus it has a higher affinity for recombinant tau than AD brain PHF tau. PHF1 is more specific to the S396/S404 phosphorylated tau, and the bands in PHF1 blots (**Fig. 1b**) mostly represent hyperphosphorylated human tau seeds. As there are six isoforms of human tau high molecular-weight tau species in the AD tau seeds, the banding pattern of PHF1-tau is much more complex than the pattern of recombinant human tau (17025-tau). "

For these reasons, the positions of PHF1 positive bands are expected to differ from the 17025 positive bands.

Reviewer #2 (Remarks to the Author):

The paper by Dregni et al represents a sound and significant contribution to the understanding of fibril formation relating to dementia. Supported by extensive biophysical and EM characterization, they present solid-state NMR data which shed new detailed light on the organization of monomers in fibrils with particular focus on how 4R and 3R tau proteins mixes in the fibril upon seeding with AD brain-derived tau – differing significantly from fibrils assembled using heparin.

Overall, the experiments and conclusions seem sound, and the outcome important for the community. As such I recommend the paper being accepted for publication in Nature Communications upon consideration of

We are glad the reviewer appreciates the importance of the work and agrees with the conclusion.

(1) I recommend authors more carefully address the significant discrepancy between experimental and simulated REDOR curves as given in Extended Data Fig. 3. I do agree that comparison with the experimental curves obtained for samples with different labelling ratios may provide a "safe" ground for the conclusion – but still it is puzzling that the simulations do not better match the experimental data. The dephasing in the experimental curves seems to be significantly faster both in the initial and later parts of the dephasing period – which may not be fully accounted for by the explanation of non-included long-range effects. Could it be associated with non-perfect decoupling or other? On the experimental side, can the comparison with experimental data for differently labelled 0N4R heparin-fibrillized tau samples be non-adequate for AD-brain derived mixed 4R and 3R samples? As all curves shift significantly (e.g. the experimental 40:60 curve resembles the simulated 60:40 curve), I recommend the authors explore this in more detail (including commenting on effects from fuzzy fibril regions), considering the fundamental importance of the REDOR decay profiles in building the conclusions of the paper.

We thank the reviewer for this question, which has prompted us to provide additional numerical simulations to explain our reasoning as to why the experimentally measured REDOR curves from the calibration samples represent the gold standard in true REDOR dephasing for given mixing ratios. In short, the numerical simulations deviate from the experiments because no program can capture all the ^{13}C spins near a ^{15}N in a cross-β fibril with realistic computational time while still taking into account multiple spin interactions quantum mechanically. We have now added this additional analysis in Extended Data Figure 3. The key take-home message is that SpinEvolution under-estimates dipolar dephasing whereas second moment analysis over-estimates dipolar dephasing. Below is one paragraph in the Methods section that summarizes these simulations:

"The 9-spin SpinEvolution simulations result in slightly higher S/S_0 values whereas the 48-spin secondary moment analysis result in lower S/S_0 values than the experimental data. The 9-spin second-moment analysis approaches the experimental data the closest. The faster dephasing of the simulated REDOR curves with 48 nearest ^{13}C spins (**Supplementary Fig. 3f**) compared to dephasing by 9 ^{13}C spins indicate that ^{15}N-^{13}C distances between 5.8 Å and 8.8 Å cannot be neglected. Thus, SpinEvolution simulation for up to 9 ^{13}C spins is an underestimate. On the other hand, the explicit SpinEvolution simulation with 9 ^{13}C spins decays more slowly than the second moment analysis for 9 spins, indicating that other spin interactions, including ^{13}C-^{13}C dipolar couplings and chemical shifts, moderately slow down the ^{15}N-^{13}C dipolar dephasing. This means the second-moment analysis of 48 ^{13}C spins overestimates the true decay rate. Therefore, neither simulation approach fully accounts for both long-range ^{13}C-^{15}N distances and all nuclear spin interactions. Due to these limitations, we rely on the experimentally measured calibration REDOR curves to extract the mixing probabilities of AD-tau seeded filaments accurately."

We now show the ^{13}C natural abundance corrected experimental data in Figure S3:

"Among the measured REDOR calibration curves from the heparin-fibrillized 0N4R tau samples, the curve of the 100% ^{15}N-labeled 0N4R tau without ^{13}C-labeled protein already exhibits dipolar dephasing to 0.84 at 25 ms (**Fig. 3e** and **Supplementary Fig. 3a**). This indicates that natural abundance ^{13}C spins cause non-negligible dephasing

to ^{15}N. This natural abundance effect can be approximately corrected by dividing all measured REDOR intensities by the intensities of the 100% ^{15}N-labeled sample (**Supplementary Fig. 3b**)."

As to the validity of using calibration samples of heparin-fibrillized 0N4R to calibrate the AD seeded samples' REDOR data, we first note that our hNH-REDOR experiment is conducted with a short 400 μs ^{15}N-^{1}H CP step, which should be sufficient to make the contributions of the fuzzy domains to the REDOR negligible. Under this condition, we are primarily observing the dephasing within the parallel-in-register 4.8 Å cross-β amyloid core. Whatever the size and shape of the cores in heparin fibrillized 0N4R tau and [AD]3R/4R tau, the REDOR dephasing is primarily reporting the rigid core contacts from amide nitrogens to the carbons of the same residue or flanking residues in the monomers above or below the ^{15}N-labeled chain in the fibril. Thus, heparin-fibrillized 0N4R tau, where all the monomers are chemically identical, represents an excellent calibration sample for the REDOR experiment. Even if one could make heparin-fibrillized tau using 0N4R and 0N3R monomers with different ratios, the uncertainty of the incorporation ratio would make it impossible to obtain a better calibration curve than what we currently have.

(2) In light of (1), the authors should discuss the accuracy of the incorporation level percentages. For example, it is argued that a p_4->3 value is 0.37 +/- 0.03 and khi_4 is 60 +/- 2%.

As explained in (1), we believe the experimental REDOR calibration by 0N4R tau is the best approach.

Reviewer #3 (Remarks to the Author):

Recent breakthroughs in cryo-EM technology have allowed the visualization of atomic structures of tau filaments extracted from Alzheimer's disease patient brains; however, how different 3R to 4R isoforms are molecularly mixed in the fibrils remains unknown. In this manuscript, Aurelio and colleagues have utilized solid-state NMR to determine how 3R and 4R tau isoforms could potentially be recruited into the AD fibrils. In addition, the authors found that the ratio of 4R:3R incorporated into AD fibrils is in a proportion of 60:40 with a homotypic preference, and a structural difference exists between heparin-induced tau fibrils and seeded tau fibrils from AD samples.

To characterize how tau isoforms are structurally incorporated into AD seeds, the authors pooled insoluble fractions from the brains of 10 AD patients and then added 15N or 13C labeled recombinant tau monomers into the mixture. Previous studies have shown that tau seeds are heterogeneous, and different tau species could exist between patients even if their Braak staging is similar. If the pooled solution of AD seeds contained different tau fibril strains, how tau monomers are recruited into the seeds may differ depending on the fibril imaged.

This is an interesting question. We have now added a paragraph in the first paragraph of the Discussion section to address this question.

" Because the NMR data were obtained from AD tau seeds pooled from multiple patient brains, the 60 : 40 mixing and the probabilities we measured reflect the general properties of AD brain tau fibrils instead of the property of a single patient brain. The AD-seeded tau fibrils show highly reproducible 2D NMR spectra as well as reproducible REDOR dephasing, suggesting that there is no detectable molecular structural variation among the multiple AD brains. Although some biochemical studies have reported certain levels of heterogeneity in AD brain tau, these observations were all made based on total tau, including both pathogenic and non-pathogenic tau, in AD patients [32,33]. Our study instead focuses on sarkosyl-insoluble tau, which has been shown to have highly consistent fibril core structures across multiple AD patient cases [5], and which exhibits consistent seeding abilities *in vitro* and *in vivo* [18]. Therefore, these biochemical and biophysical data provide additional evidence for the isoform-independent molecular structure observed here. "

The authors' findings of the fluent mixing of 3R and 4R into the AD seed is novel; however, this incorporation pattern may occur differently in vivo. Because the tau monomers the authors used were purified from bacterial cells, they were not post-translationally modified like in human brains. These post-translational modifications, such as phosphorylation, may play an essential role in electrostatic interactions and steric hindrances between the tau monomers and insoluble tau fibrils, which are missing in the author's system. While the use of human samples as seeds may give this study more human relevance, the fact that these measurements are made on in vitro derived fibrils should be more clearly stated in the abstract/conclusions.

We agree that our data obtained on bacterially produced recombinant tau do not probe whether and how posttranslational modifications affect the 4R/3R tau mixing in human brains. On a molecular level, it is difficult to imagine how PTMs would influence the mixing of 3R and 4R tau in AD brains. This question would require future investigations.

Specific Comments:
1. While the introduction provided succinctly describes the need for the study, it is extremely brief.

The original text was written for the journal Nature, which prizes brevity and puts background information into the abstract. We have now shortened the abstract and expanded the introduction according to the Nature Communications format.

2. The order of how the result is presented does not follow the order of the figures. The authors defined the mixture of AD fibrils and 4R and 3R tau for Figure 3 after they were used in figure 1 results.

We have updated the results and the captions for figure 1.

3. In Figure 1.b, why is the [AD] 3RNC western blot not shown in the main figure, but moved to the supplemental data section?

This is due to space limitation. Since Extended Data Figure 1 shows all four samples' western blot, we do not feel that the main text figure should enumerate all details. All four main text figures are already very large, so we try to delegate the less important details to extended data figures.

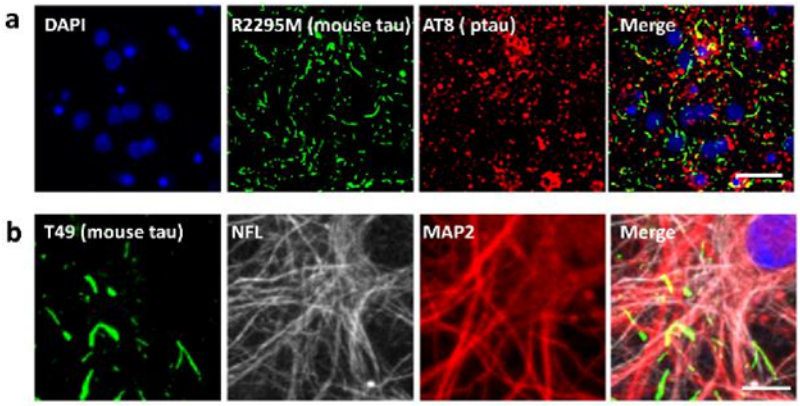
4. In Figure 1.b and extended data figure 1, the authors showed that recombinant tau monomers are incorporated into insoluble mass after 7 days and stained the western blot with total tau antibody and phospho-tau antibody. It would be informative to address why more phosphorylated tau are present (more intense PHF1 bands) after incubating the tau seeds with monomers. Are kinases present in their mixture ?

To demonstrate homogeneity in tau filament sizes, the AD-tau seeds (PHF-1 positive bands) were sonicated intensely before being added to the cell-free seeding reactions. Because of this intense sonification step, ultracentrifugation at 100,000 g for 30 min in the sedimentation assay could not completely sediment the insoluble AD-tau seeds at day 0 and a significant amount of tau seeds stay in the supernatant. After 7 days of seeding reaction, tau filaments were elongated and became easier to be pelleted, resulting in a shift in the amount of AD-tau in the pellet fraction. However, the total amount of AD-tau (supernatant + pellet) was not significantly altered before and after the reaction. Thus, there is no evidence of kinases existing in the insoluble tau seeds.

We have now added this clarification in the Methods section on sedimentation assays.

5. In the primary neuron seeding experiment, that the localization of the AD tau or seeded tau fibrils is not distinguished (for example within the neurons or sitting on top of the neurons). In addition, the images of the neurons are not clear. Consider adding neuronal staining to show localization of tau fibrils and the purity of the culture.

In the neuron-based activity assay, we used R2295M mouse tau specific antibody to reveal the seeded mouse tau aggregates in primary mouse neurons. The neurons were treated with 1% HDTA, a detergent, to remove soluble proteins in the cell culture. This paradigm allows us to specifically show the amount of aggregated mouse tau induced by pathogenic tau seeds without the confounding factor of exogenous tau seeds. To demonstrate the presence of the exogenous human tau and endogenous mouse tau, below we attach immunofluorescence images showing neurons treated with AD tau for 14 days and stained with AT8 (S202/T205 pTau) for AD-tau and R2295M for mouse tau. pTau immunoreactivity shows homogenous distribution of AD-tau seeds, while mouse tau antibody only reveals aggregated mouse tau (a). Further evidence could be found in our previous publication {Xu, 2021 #4}. To elucidate the presence of aggregated mouse tau in cultured neurons and their localization, we show below images of neurons treated with AD-tau seeds for 14 days and extracted with methanol to remove immunoreactivity from soluble mouse tau while preserving cytoskeleton proteins, followed by immunofluorescent staining with monoclonal antibody T49 (for mouse tau), MAP2 (in dendrites) and NFL (in axons) (b). These images show that insoluble mouse tau is primarily surrounding MAP2 and NFL in neurons, indicating their neuronal localization.



Characterization of mouse tau pathology in neuron-based activity assay. CD1 mouse primary hippocampal neurons (PHN) were treated with AD-tau for 14 days at DIV7, fixed 1%HDTA + 4% PFA and stained with R2295M anti-mouse tau antibody for mouse tau aggregates and AT8 anti-ptau antibody for AD-tau seeds (**a**). Scale bar = 25 µm. PHN treated the same way were fixed with methanol and stained with T49 antibody for mouse tau, NFL antibody for neurofilament light chain, and MAP2 (**b**). Scale bar = 10 µm.

6. Was an equal portion of each brain used to make the pooled insoluble tau sample?

Brain lysates were prepared according to the availability of the fresh frozen tissues from each case. Please find this information added in extended data table 1.

Minor Comments:
1. In the bar graphs of Figure1 D, what are the error bars showing?

We have now clarified the figure caption as:

" Statistical analysis was performed using One-way ANOVA. Multiple comparison between groups was conducted using Tukey post hoc test. On the graph, * stands for $P < 0.05$, ** stands for $P < 0.01$, *** stands for $P < 0.001$ in the multiple comparison tests. Each data point stands for one biological repeat. "

2. In Figure 2a: TEM image quality makes it difficult to see. Extended data figure 2 of TEM images seemed to have a better quality images compared to Figure 2.

We thank the reviewer for this comment. We have now swapped the two images, to place the full TEM views in figure 2 of the main text, and the aligned fibril images in Extended data Figure 2.

REVIEWERS' COMMENTS

Reviewer #2 (Remarks to the Author):

The work of the authors revising the manuscript in view of reviewers comments is appreciated. Overall the paper stands clearer now.

Still, however, remains the dilemma between using heparin-fibrillized 0N4R samples or numerical simulations of REDOR spectra to compare with the corresponding spectra from AD-brain derived mixed 4R and 3R samples:

1) Indeed the authors included many more simulation/moment analysis data in an attempt to get at better match to the experimental curves. This is highly appreciated, however, the representation of data in Supplementary Figure 3 and the long discussion on what resembles best the experimental spectra are very difficult to evaluate - not least because the earlier existing direct overlay of experimental and simulated data now is taken out, leaving the reader to compare decaying signals in different panels of supplementary figure 3 when following the discussion. I would recommend the authors helping the reader in evaluating the quality of the simulations by adding a few figure panels overlaying experimental and the chosen simulated data, and on the these clearly point out which match is chosen for the conclusions relating to the different experimental data - i.e., the chosen x:y experimental and x:y simulated cures. This may reduce the length of the discussion and certainly give the reader the opportunity to follow the connect between data and simulations.

2) Also when saying that comparison with experimental spectra is chosen as the solution, it seems important to comment on potential challenges comparing heparin-fibrillated samples with AD-seed samples.

Reviewer #3 (Remarks to the Author):

The authors have addressed our comments

REVIEWERS' COMMENTS

Reviewer #2 (Remarks to the Author):
The work of the authors revising the manuscript in view of reviewers comments is appreciated. Overall the paper stands clearer now.

Still, however, remains the dilemma between using heparin-fibrillized 0N4R samples or numerical simulations of REDOR spectra to compare with the corresponding spectra from AD-brain derived mixed 4R and 3R samples:

1) Indeed the authors included many more simulation/moment analysis data in an attempt to get at better match to the experimental curves. This is highly appreciated, however, the representation of data in Supplementary Figure 3 and the long discussion on what resembles best the experimental spectra are very difficult to evaluate - not least because the earlier existing direct overlay of experimental and simulated data now is taken out, leaving the reader to compare decaying signals in different panels of supplementary figure 3 when following the discussion. I would recommend the authors helping the reader in evaluating the quality of the simulations by adding a few figure panels overlaying experimental and the chosen simulated data, and on the these clearly point out which match is chosen for the conclusions relating to the different experimental data - i.e., the chosen x:y experimental and x:y simulated curves. This may reduce the length of the discussion and certainly give the reader the opportunity to follow the connect between data and simulations.

We have now added panels (f-h) in Supplementary Fig. 3 to show overlays of the natural abundance corrected experimental REDOR curves for different calibration samples with three different simulations. These overlays more explicitly show that none of the simulation methods – whether SpinEvolution for up to 9 ^{13}C spins or second-moment analysis for 48 ^{13}C spins - accurately capture the experimentally measured REDOR dephasing curves for the different calibration samples. The second-moment analysis for up to 9 ^{13}C spins approach the experimental data the closest. These inaccuracies in the simulations can all be understood. Therefore, the most accurate way to determine the 4R/3R mixing probability in the AD tau fibril is to use the experimentally measured calibration curves of the ^{15}N/^{13}C mixed labeled samples, rather than using simulations. This is clarified in the Methods section.

2) Also when saying that comparison with experimental spectra is chosen as the solution, it seems important to comment on potential challenges comparing heparin-fibrillated samples with AD-seed samples.

We have now added a sentence in the Results section to clarify the reason for using heparin-fibrillized tau to calibrate the REDOR dephasing of AD-seeded tau:

" The fact that in vitro fibrilized tau and AD-seeded tau do not have the same rigid core structure does not affect the use of intermolecular REDOR for determining the statistics of mixing, because the REDOR experiment only depends on the existence of parallel-in-register cross-β packing, which is true for both heparin-fibrillized tau and AD tau. "

Reviewer #3 (Remarks to the Author):

The authors have addressed our comments.

Source Codes for

Fluent Molecular Mixing of Tau Isoforms in Alzheimer's Disease Neurofibrillary Tangles

Aurelio J. Dregni [1$], Pu Duan[1$], Hong Xu[2], Lakshmi Changolkar[2], Nadia El Mammeri[1], Virginia M.-Y. Lee[2], and Mei Hong[1*]

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[2] Department of Pathology and Laboratory Medicine, Institute On Aging and Center for Neurodegenerative Disease Research, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA

* Corresponding author: Mei Hong: meihong@mit.edu
$ These authors contributed equally to this work.

SpinEvolution Source Code for REDOR simulation

SpinEvolution Code for the case of ^{13}C-chain → ^{15}N-chain → ^{13}C-chain:

```
****** The System **********************************
spectrometer(MHz)  600
spinning_freq(kHz) 20.0
channels           N15 C13
nuclei             N15 C13 C13 C13 C13 C13 C13 C13 C13 C13
atomic_coords      I308_5p55_343.txt
cs_isotropic       0 -47 70 -65 -40 70 -82 -40 -62 -75 ppm
csa_parameters     1 0 0 0 0 0 ppm
csa_parameters     2 30 0 10 45 0 ppm
csa_parameters     3 170 0 90 30 0 ppm
csa_parameters     4 30 0 50 10 0 ppm
csa_parameters     5 30 0 80 60 0 ppm
csa_parameters     6 170 0 5 20 0 ppm
csa_parameters     7 30 0 49 10 0 ppm
csa_parameters     8 30 0 25 60 0 ppm
csa_parameters     9 30 0 50 20 0 ppm
csa_parameters     10 30 0 30 20 0 ppm
j_coupling         *
quadrupole         *
dip_switchboard    *
csa_switchboard    *
exchange_nuclei    *
bond_len_nuclei    *
bond_ang_nuclei    *
tors_ang_nuclei    *
groups_nuclei      *
******* REDOR Pulse Sequence ***********************
CHN 1
timing(usec)      (50)512     42     8    50       (50)512
power(kHz)         0           0    62.5   0         0
phase(deg)         0           0     0     0         0
freq_offs(kHz)     0           0     0     0         0
CHN 2
timing(usec)    (redor.pp)   17  8     25  redor.pp (redor.pp)
power(kHz)         *            0   62.5  0     *           *
phase(deg)         *            0   0       0     *           *
freq_offs(kHz)     *            0   0       0     *           *

******* Variables ***********************************
scan_par1d rf/62.5:-1.74:52.06/
power_2_1_2 = rf
power_2_1_4 = rf
power_2_2_2 = rf
power_2_2_5 = rf
power_2_2_7 = rf
power_2_3_2 = rf
power_2_3_4 = rf
******* Options *************************************
rho0               I1x
observables        I1p
EulerAngles        rep30
```

n_gamma *
line_broaden(Hz) *
zerofill *
FFT_dimensions *
options -dw123 -re

"atomic_coords" file for the case of ^{13}C-chain → ^{15}N-chain → ^{13}C-chain:

130.013	103.502	142.851		N	ILE	F	0
129.061	101.533	138.265		CA	GLN	below	5.080813025
130.095	102.462	138.877		C	GLN	below	4.108649413
127.663	102.083	138.528		CB	GLN	below	5.120975493
131.652	104.3		138.441	CA	ILE	below	4.771920473
130.98		105.609	138.057	C	ILE	below	5.325126665
128.329	107.077	139.319		CG2	VAL	below	5.300141979
130.205	104.843	148.098		CA	ILE	Above	5.419054715
131.602	104.671	147.465		CB	ILE	above	5.018015345
132.221	103.35		147.919	CG1	ILE	above	5.530189147

SpinEvolution Code for the case of ^{13}C-chain → ^{15}N-chain → ^{15}N-chain:

```
****** The System **********************************
spectrometer(MHz)  600
spinning_freq(kHz) 20.0
channels           N15 C13
nuclei             N15 C13 C13 C13 C13 C13
atomic_coords      I308_5p8_344.txt
cs_isotropic       0 -40 70 -60 -85 -90 ppm
csa_parameters     1 0 0 0 0 0 ppm
csa_parameters     2 30 0 10 45 0 ppm
csa_parameters     3 170 0 90 30 0 ppm
csa_parameters     4 30 0 50 10 0 ppm
csa_parameters     5 30 0 80 60 0 ppm
csa_parameters     6 30 0 5 20 0 ppm
j_coupling         *
quadrupole         *
dip_switchboard    *
csa_switchboard    *
exchange_nuclei    *
bond_len_nuclei    *
bond_ang_nuclei    *
tors_ang_nuclei    *
groups_nuclei      *
******* REDOR Pulse Sequence **********************
CHN 1
timing(usec)      (50)512     42     8    50       (50)512
power(kHz)         0           0    62.5   0         0
phase(deg)         0           0     0     0         0
freq_offs(kHz)     0           0     0     0         0
CHN 2
timing(usec)    (redor.pp)   17  8     25  redor.pp (redor.pp)
power(kHz)         *            0   62.5  0     *           *
phase(deg)         *            0   0       0     *           *
freq_offs(kHz)     *            0   0       0     *           *
******* Variables ***********************************
scan_par1d rf/62.5:-1.74:52.06/
power_2_1_2 = rf
power_2_1_4 = rf
power_2_2_2 = rf
power_2_2_5 = rf
power_2_2_7 = rf
power_2_3_2 = rf
power_2_3_4 = rf
******* Options *************************************
rho0               I1x
observables        I1p
EulerAngles        rep30
n_gamma            *
line_broaden(Hz)   *
zerofill           *
FFT_dimensions     *
options            -dw123 -re
```

"atomic_coords" file for the case of ^{13}C-chain → ^{15}N-chain → ^{15}N-chain:

130.013	103.502	142.851	N	ILE	F	0
130.205	104.843	148.098	CA	ILE	Above	5.419054715
129.585	106.179	147.718	C	ILE	above	5.571104199
131.602	104.671	147.465	CB	ILE	above	5.018015345
132.221	103.35	147.919	CG1	ILE	above	5.530189147
133.458	102.93	147.167	CD1	ILE	above	5.551852394

SpinEvolution Code for the case of ^{15}N-chain → ^{15}N-chain → ^{13}C-chain:

```
****** The System **********************************
spectrometer(MHz)  600
spinning_freq(kHz) 20.0
channels           N15 C13
nuclei             N15 C13 C13 C13 C13 C13 C13 C13 C13
atomic_coords      I308_5p8_443.txt
cs_isotropic       0 -47 70 -65 -67 -40 75 -70 -82 ppm
csa_parameters     1 0 0 0 0 0 ppm
csa_parameters     2 30 0 10 45 0 ppm
csa_parameters     3 170 0 90 30 0 ppm
csa_parameters     4 30 0 50 10 0 ppm
csa_parameters     5 30 0 80 60 0 ppm
csa_parameters     6 30 0 5 20 0 ppm
csa_parameters     7 170 0 49 10 0 ppm
csa_parameters     8 30 0 25 60 0 ppm
csa_parameters     9 30 0 50 20 0 ppm
j_coupling         *
quadrupole         *
dip_switchboard    *
csa_switchboard    *
exchange_nuclei    *
bond_len_nuclei    *
bond_ang_nuclei    *
tors_ang_nuclei    *
groups_nuclei      *
******* REDOR Pulse Sequence **********************
CHN 1
timing(usec)      (50)512     42     8    50       (50)512
power(kHz)         0          0    62.5   0         0
phase(deg)         0          0     0     0         0
freq_offs(kHz)     0          0     0     0         0
CHN 2
timing(usec)    (redor.pp)   17  8     25  redor.pp (redor.pp)
power(kHz)         *           0   62.5  0     *           *
phase(deg)         *           0   0       0     *           *
freq_offs(kHz)     *           0   0       0     *           *
******* Variables ***********************************
scan_par1d rf/62.5:-1.74:52.06/
power_2_1_2 = rf
power_2_1_4 = rf
power_2_2_2 = rf
power_2_2_5 = rf
power_2_2_7 = rf
power_2_3_2 = rf
power_2_3_4 = rf
******* Options ***************************************
rho0               I1x
observables        I1p
EulerAngles        rep30
n_gamma            *
line_broaden(Hz)   *
zerofill           *
FFT_dimensions     *
options            -dw123 -re
```

"atomic_coords" file for the case of ^{13}C-chain → ^{15}N-chain → ^{15}N-chain:

130.013	103.502	142.851		N	ILE	F	0
129.061	101.533	138.265		CA	GLN	below	5.080813025
130.095	102.462	138.877		C	GLN	below	4.108649413
127.663	102.083	138.528		CB	GLN	below	5.120975493
127.402	103.439	137.901		CG	GLN	below	5.59676603
131.652	104.3		138.441		CA	ILE	below	4.771920473
130.98		105.609	138.057		C	ILE	below	5.325126665
129.384	108.127	139.63			CB	VAL	below	5.671076353
128.329	107.077	139.319		CG2	VAL	below	5.300141979

SpinEvolution pulse sequence file “redor.pp”:

```
17   0      0    0
8    62.5   0    0
17   0      0    0
8    62.5   90   0
```

Bruker Topspin 3.5 pulse program code for ^{1}H-detected ^{15}N-^{13}C REDOR

```
;1H detected 15N{13C} REDOR
;Avance III version
;f1 : H
;f2 : N
;f3 : C
;o1 : H offset, on resonance with water peak (~5 ppm, water suppression!)
;o2 : N offset, center of 15N signal (~119 ppm)
;o3 : C offset, center of 13C signal (~100 ppm, using '-DTC')
;p3 : H 90 hard pulse at pl2
;p21 : N 90 hard pulse at pl21
;p22 : N 180 hard pulse at pl21
;p33 : C 180 hard pulse at pl3 for redor
;p25 : HN CP at sp42 (H) & sp43 (N), (~1 to 3 ms)
;p45 : NH CP at sp46 (H) & sp47 (N), (~400 to 800 us)
;pl2 : H hard pulse power
;pl3 : C hard pL for REDOR
;pl12 : H dec power ('tppm15' at >70 kHz for <14 kHz MAS; 'waltz16_12nofq' at ~10 kHz for fast MAS)
;pl13 : H dec power during H2O suppression (~15 kHz, 'cwX_13nofq', 'cwY_13nofq')
;pl16 : N dec power ('waltz16_16nofq' at ~7 kHz)
;pl18 : C dec power ('waltz16_18nofq' at ~7 kHz, using '-DTC')
;pl21 : N hard pulse power (can be optimized with '-DN90')
;sp42 : H HN CP power
;sp43 : N HN CP power
;sp46 : H NH CP power
;sp47 : N NH CP power
;d1 : recycle delay; 1 to 5 times T1 (~0.8 to 1 s)
;d19 : delay for water suppression (~100 to 300 ms)
;d5: Total REDOR time, tune by l10
;cpdprg1 : H dec ('tppm15' at >70 kHz for <14 kHz MAS; 'waltz16_12nofq' at ~10 kHz for fast MAS)
;cpdprg2 : N dec ('waltz16_16nofq' at pl16 (10 kHz))
;cpdprg3 : C dec ('waltz16_18nofq' at pl18 (10 kHz), using '-DTC')
;cpdprg4 : H Water suppression along X ('cwX_13nofq' at pl13 (15 kHz))
;cpdprg5 : H Water suppression along Y ('cwY_13nofq' at pl13 (15 kHz))
;pcpd1 : H dec pulse
;pcpd2 : N dec pulse: 35.35 us ('waltz16_16nofq' ~7 kHz)
;pcpd3 : C dec pulse: 35.35 us ('waltz16_18nofq' ~7 kHz, using '-DTC')
;spnam42 : H shape (ramp up for NH CP, e.g. 'ramp.70100.1000')
;spnam43 : N shape (e.g. 'square.1000' for HN CP (=no shape))
;spnam46 : H shape (ramp down for NH CP, e.g. 'ramp.10070.1000')
;spnam47 : N shape (e.g. 'square.1000' for NH CP (=no shape))
;inf1 :  1/SW(N) = 2 * DW(N)
;in0 : = inf1
;l31: spinning speed in Hz
;l10 : REDOR mix = L10*tr, must be even
;ZGOPTNS : -DTC : switch on C decoupling
;  -DN90: N 90 degree pulse check
;  or blank
;FnMODE : States-TPPI
;ns : MIN. 4 (full: 8)
;ds : 2 or 4

;#########################################################
;# hNHREDOR                                    #
;# Dregni, Duan… Hong, Nature Communications 2022          #
```

```
;#########################################################

;$COMMENT=1H detected 15N{13C} REDOR
;$CLASS=BioSolids
;$DIM=1D
;$TYPE=H detection
;$SUBTYPE=CP, Heteronuclear

prosol relations=<biosolHCN>

#include <HNC_defs.incl>
       ; defines H:f1, N:f2, C:f3

#include <trigg.incl>
       ; definition of external trigger output

"acqt0=-(p1*2/3.1416)-0.5u"     ; baseopt correction

"spoff42=0.0"              ;############################
"spoff43=0.0"              ;#     ensure correct     #
"spoff46=0.0"              ;#     shape offsets      #
"spoff47=0.0"              ;############################

"p2=p1*2"
"p22=p21*2"

  "d11=0.5s/l31-p22/2"                    ;to center square refocusing pulse p22
  "d25=0.25s/l31-1u"                      ;REDOR. Used at beginning and end.
  "d26=d25-p21/2"                         ;Account for N flip up
  "d27=0.25s/l31-p33/2"                   ;REDOR, altered for square timing
  "d28=0.25s/l31"                         ;Tr/4, REDOR
  "d5=l10/l31"

"in0=inf1"               ;############################
"d0=1u"                  ;#    t1_init => 0, 0     #
"l0=0"                  ;############################

                       ;############################
;$EXTERN                      ;# python insertion point #
                       ;############################

Prepare, ze

;#########################################################
;#          Start of Active Pulse Program          #
;#########################################################

Start, 30m  do:N

#ifdef TC
  0.5u do:C
#endif         /* end of TC */

  d1
  d5
```

```
  trigg

if "l0>0"
{
  "d51=d0-1u"
}

;#########################################################
;#            Initial excitation & HN CP            #
;#########################################################

  (p3 pl2 ph1):H

  (p25:sp42 ph0):H  (p25:sp43 ph2):N

;------------NC REDOR------------------------------

  1u cpds1:H                                ;composite pulse decoupling scheme for REDOR

  d25 pl3:C

6 d27
  p33:C ph15^
  d27
lo to 6 times l10                   ; L0 still must be even


  d28
  d11
  (p22 pl21 ph10):N                 ;square pulse - for nonselective appplications
  d11
  d28

8 d27
  p33:C ph15^
  d27
lo to 8 times l10

  d26
  1u do:H

;----------End of REDOR-----------------------------

#ifdef TC
if "l0>0"
{
  1u cpds1:H
  (center (d51) (p33 pl3 ph20):C)
  0.5u do:H  pl13:H
}
#else
if "l0>0"
{
  0.5u cpds1:H
  d51
```

```
  0.5u do:H  pl13:H
}
#endif          /* end of TC */


;#########################################################
;#            Water suppression                          #
;#########################################################

  (p21 pl21 ph3):N          ; brings magn. to z

  0.5u cpds4:H
  d19*0.25
  0.5u do:H

  0.5u cpds5:H
  d19*0.25
  0.5u do:H

  0.5u cpds4:H
  d19*0.25
  0.5u do:H

  0.5u cpds5:H
  d19*0.25
  0.5u do:H

  (p21 pl21 ph4):N          ; brings magn. to y

;#########################################################
;#          15N hard pulse optimization                  #
;#########################################################

#ifdef N90                  ; brings magn. to z.
  (p21 pl21 ph22):N
#endif          /* end of N90 */

;#########################################################
;#                  N-H CP                               #
;#########################################################

  (p45:sp47 ph5):N  (p45:sp46 ph6):H

;#########################################################
;#                  Acquisition                          #
;#########################################################

#ifdef TC
  1u cpds2:N  cpds3:C
#else
  0.5u pl16:N
  0.5u cpds2:N
#endif          /* end of TC */

  gosc ph31     ;start ADC with ph31 signal routing
```

```
#ifdef TC
  1m do:C  do:N
#else
  1m do:N
#endif          /* end of TC */

lo to Start times ns

30m mc #0 to Start
F1PH(calph(ph2, +90), caldel(d0, +in0) & calclc(l0, 1))

HaltAcqu, 1m
exit

;#######################################
;#           Phase Cycle           #
;#######################################

ph1 = 1 3                    ; 1H 90 hard pulse
ph0 = 0                      ; 1H HN CP Spin lock
ph2 = 1 1 1 1 3 3 3 3        ; 15N HN CP Spin lock
ph3 = 0                      ; 15N 1st 90 hard pulse (flip to z)
ph4 = 0 0 2 2                ; 15N 2nd 90 hard pulse (flip back)
ph5 = 1                      ; 15N NH CP Spin lock
ph6 = 1                      ; 1H NH CP Spin lock
ph10= 1 1 1 1 1 1 1 1
      3 3 3 3 3 3 3 3        ;F1 180, normal shaped pulse

ph15=  0 1 0 1               ; xy-4. F3 phase
ph31= 1 3 3 1 3 1 1 3        ; receiver

#ifdef TC
ph20= 0                       ; C 180 hard pulse
#endif                  /* end of TC */

#ifdef N90
ph22= 0 0 2 2                ;N hard pulse
#endif                  /* end of N90 */

;#########################################################
```

Python Source Code for Building Fibril Model (Fig 3f, g)

```
#Simulates a N residue fibril with given p34 and p43.
#Then computes observed probabilities in the simulated fibril.
#Returns "results", a string of '3', and '4' indicating the simulated fibril.
#Also returns obsProbs: (p43 observed in sim fibril - input p43, p34 observed in sim fibril - p34)
#Repeat simulation until observed probabilities are within one percentage point of input probabilities

from random import random

p34 = 0.56
p33 = 1-p34
p43 = 0.37
p44 = 1-p43

results = '3'
N = 360
for idx in range(N):
    r = random()
    if results[-1] == '3':
        if r < p34:
            results = results + '4'
        else:
            results = results + '3'
    else:
        if r < p43:
            results = results + '3'
        else:
            results = results + '4'

print(results)
obs4 = results.count('4')
obs3 = results.count('3')
obs43 = results.count('43')
obs44 = obs4 - obs43
obs34 = results.count('34')
obs33 = obs3 - obs34
#obsProbs = (obs4/N,obs44/obs4,obs43/obs4,obs34/obs3,obs33/obs3) #p4, p44, p43, p34, p33
obsProbs = (obs43/obs4-p43,obs34/obs3-p34)
obsProbs
```

Python Source Code for Fibril Models with Varying Q and χ_4 (Supplementary Fig. 7)

```
#Simulates an N residue fibril with given chi4 and Q.
#First computes p34 and p43 from chi4 and Q (A quadratic equation)
#Then simulates fibril based on p34 and p43.
#Variable "Results" is printed and contains the simulated fibril.
#Also returns obsParams: (Chi4 observed in sim fibril, input Chi4, Q observed in sim fibril, input Q)
#Repeat simulation until observed and input parameters are within appropriate tolerance.

from random import random
import math as mat
chi4 = 0.7 #macroscopic preference = p34 / (p43 + p34)
Q = 1/16 #microscopic mixing parameter = p34 p43 / (p33 p44).
          #Choose slightly off 1.0, as equations fail for Q=1 exactly
W = (1-chi4)/chi4
p34 = 1/(2*W*(Q-1)) * (Q*(1+W) - math.sqrt((Q**2) *((1+W)**2)-4*Q*W*(Q-1)))
p43 = W*p34

p33 = 1-p34
p44 = 1-p43
probs = (p34, p43, p33, p44)
probs
results = '3'
N = 100
for idx in range(N):
   r = random()
   if results[-1] == '3':
      if r < p34:
         results = results + '4'
      else:
         results = results + '3'
   else:
      if r < p43:
         results = results + '3'
      else:
         results = results + '4'

print(results)
obs4 = results.count('4')
obs3 = results.count('3')
obs43 = results.count('43')
obs44 = obs4 - obs43
obs34 = results.count('34')
obs33 = obs3 - obs34
obsParams = (obs4/N, chi4, obs43*obs34/(obs44*obs33),Q)
obsParams
```