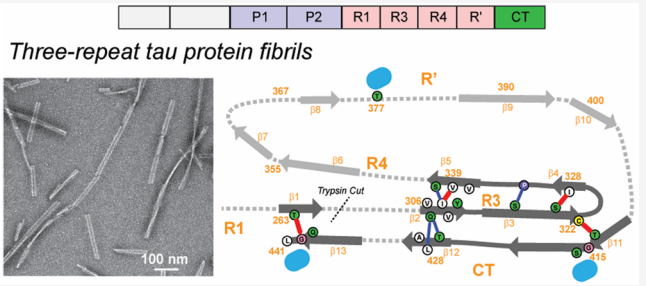
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Inclusion of the C-Terminal Domain in the β-Sheet Core of Heparin-Fibrillized Three-Repeat Tau Protein Revealed by Solid-State Nuclear Magnetic Resonance Spectroscopy

Aurelio J. Dregni,[$] Harrison K. Wang,[$] Haifan Wu, Pu Duan, Jia Jin, William F. DeGrado, and Mei Hong*

Cite This: J. Am. Chem. Soc. 2021, 143, 7839−7851
Read Online

ABSTRACT: Many neurodegenerative diseases such as Alzheimer's disease are characterized by pathological β-sheet filaments of the tau protein, which spread in a prion-like manner in patient brains. To date, high-resolution structures of tau filaments obtained from patient brains show that the β-sheet core only includes portions of the microtubule-binding repeat domains and excludes the C-terminal residues, indicating that the C-terminus is dynamically disordered. Here, we use solid-state NMR spectroscopy to identify the β-sheet core of full-length 0N3R tau fibrillized using heparin. Assignment of ^{13}C and ^{15}N chemical shifts of the rigid core of the protein revealed a single predominant β-sheet conformation, which spans not only the R3, R4, R′ repeats but also the entire C-terminal domain (CT) of the protein. This massive β-sheet core qualitatively differs from all other tau fibril structures known to date. Using long-range correlation NMR experiments, we found that the R3 and R4 repeats form a β-arch, similar to that seen in some of the brain-derived tau fibrils, but the R1 and R3 domains additionally stack against the CT, reminiscent of previously reported transient interactions of the CT with the microtubule-binding repeats. This expanded β-sheet core structure suggests that the CT may have a protective effect against the formation of pathological tau fibrils by shielding the amyloidogenic R3 and R4 domains, preventing side-on nucleation. Truncation and post-translational modification of the CT *in vivo* may thus play an important role in the progression of tauopathies.

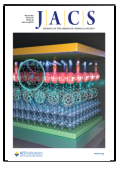


INTRODUCTION

Tau is an intrinsically disordered microtubule-binding protein that aggregates into intracellular inclusions in a number of neurodegenerative diseases.[1] These aggregates are cross-β amyloid fibrils, whose structured cores are comprised of parts of the microtubule-binding repeat domains (R) R1−R4 and R′ (Figure 1a). The remainder of the protein is dynamically disordered and appears as a "fuzzy coat" around the fibril core in electron micrographs.[2−4] Adult human brains express six tau isoforms, distinguished by whether they contain zero, one, or two 29-residue acidic inserts (0N, 1N, or 2N tau) in the N-terminal region and the presence or absence of the 31-residue R2 repeat (3R or 4R tau). Different tauopathies are characterized by tau proteins with different isoform compositions. For example, tau fibrils in Alzheimer's disease (AD) and chronic traumatic encephalopathy (CTE) are comprised of both three-repeat (3R) and four-repeat (4R) tau, in Pick's disease (PiD) are solely 3R tau, while in corticobasal degeneration (CBD) are solely 4R tau.[5] The β-sheet structures of tau fibrils in AD,[6] CTE,[7] PiD,[8] and CBD[9] brains have been determined using cryo-electron microscopy (cryo-EM). These results show that each tauopathy is characterized by a distinct β-sheet core conformation,[10] which is the same for patients with the same disease. Moreover, these conformationally specific tau aggregates can spread between neurons in connected brain regions by recruiting soluble tau to adopt the same pathological conformation.[11−14] Therefore, mapping the structures of tau aggregates is important for understanding the structure-disease relationship and for designing imaging and therapeutic agents to diagnose and treat tauopathies.

To date, the known rigid cores of tau filaments from patient brains comprise 73−107 residues (Table 1), all spanning R3, R4, and part of R′. Additional β-sheet residues are found in the R2 region in CBD tau and the R1 region in PiD tau. No residues before K254 or after E380 have been observed in any of the brain tau fibril cores. The β-strands in the core hydrogen-bond in a parallel-in-register fashion across hundreds of molecules along the fibril axis, giving rise to the spine of the

Received: March 29, 2021
Published: May 13, 2021



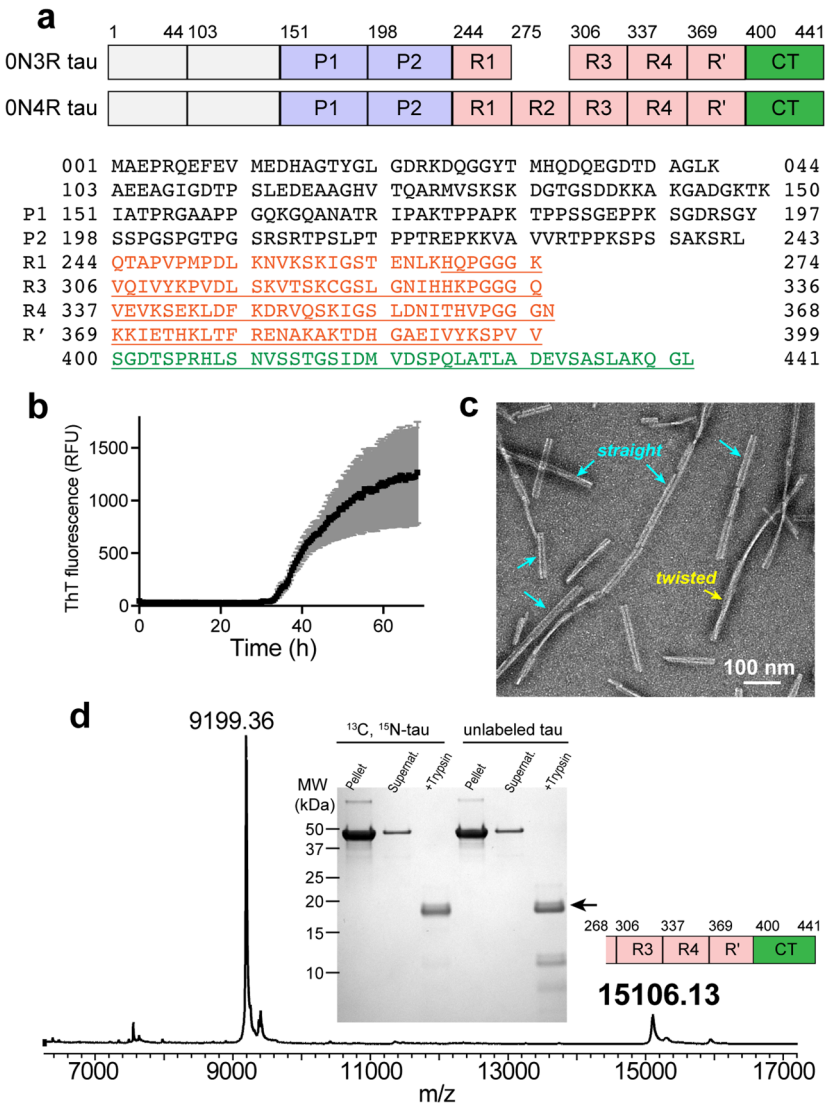


Figure 1. Heparin-fibrillized 0N3R tau forms well-ordered amyloid fibrils with a single ultrastructural morphology. (a) Comparison of the domains of 0N3R and 0N4R tau and the sequence of 0N3R tau. Due to the absence of N-terminal inserts, residue 44 precedes 103 in both 0N4R and 0N3R tau. 0N3R tau does not contain R2, thus residue 274 is followed by residue 306. Underlined residues correspond to the trypsin-resistant core of 0N3R tau. (b) ThT fluorescence of heparin-induced fibrillization of 0N3R tau, showing a lag time of ~33 h. (c) Representative negative-stain TEM image of 0N3R tau, showing that the fibrils are predominantly straight with a width of ~23 nm, together with a small fraction of twisted fibrils. (d) SDS-PAGE gel and mass spectrum of 0N3R tau fibrils after trypsin digestion. In the gel, both ^{13}C, ^{15}N-labeled tau and unlabeled tau show a dominant band whose molecular weight is 15.1 kDa, corresponding to residues 268−441. A much weaker band with a molecular weight of 9199 (residues 354−441) is sometimes observed. Pellet and supernatant lanes before trypsin digestion show a single band at the intact protein molecular weight, confirming the purity of the protein used for solid-state NMR measurements. The mass spectrum corresponds to unlabeled tau.

Table 1. Locations of the β-Sheet Core of Full-Length Tau Fibrils Reported so Far

Tau fibril source	PDB code	Structurally assigned β-sheet core	No. residues	Microtubule-binding repeats	Longest trypsin resistant domain
AD tau (3R and 4R)	6HRE[6]	V306−F378	73	R3, R4, R′	H268-R406[15]
PiD tau (3R)	6GXS[21]	K254−F378	94	R1, R3, R4, R′	L243-K385[15]
CTE tau (3R and 4R)	6NWP[7]	K274/S305−R379	75	R3, R4, R′	
CBD tau (4R)	6VHA[9]	K274−E380	107	R2, R3, R4, R′	H268−K395[15]
In vitro 0N4R tau[24]		K274−K340	67	R1, R2, R3, R4	H268−K340[24]
In vitro 2N4R tau, snake	6QJH[23]	G272−H330	28	R1, R2, R3	
In vitro 0N3R tau		G262−L441	149	R1, R3, R4, R′, CT	H268−L441

long filaments seen in transmission electron microscopy (TEM) images. In addition to high-resolution cryo-EM structures, trypsin digestion followed by mass spectrometry has provided useful information about the size of the tau fibril cores in diseased brains.[15] These trypsin-resistant regions are overall consistent with the cryo-EM results but can be larger than the structurally resolved β-sheet core: the most C-terminal residue reported so far is R406 at the beginning of the C-terminal domain (CT).[15]

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

In addition to aggregating in patient brains, tau can also form fibrils *in vitro* with the help of anionic cofactors such as polyglutamate, RNA, and the sulfonated glycosaminoglycan heparin.[16,17] Extensive biochemical and biophysical studies of these *in vitro* fibrillized tau revealed structural similarities to brain tau fibrils, but their full three-dimensional fold usually differs from the brain tau fibrils, and the structure and morphology of these *in vitro* tau fibrils are also sensitive to the protein construct length and fibrillization conditions. Solid-state NMR studies of K19, a three-repeat tau construct that spans residues Q244-E372, found a β-arch formed by R3 and R4,[18] with a rigid β-strand at V306−S320 and a disordered segment at K321−S324.[19,20] This motif was later observed in PiD tau[21] and CBD tau.[9] Solution NMR studies of heparin-fibrillized 2N4R tau found residues 212−399 to be undetectable, indicating that this domain is either rigid or undergoes intermediate-time scale motion.[22] Cryo-EM studies of heparin-fibrillized 2N4R tau detected a β-sheet core consisting of residues G272−H330 in the dominant polymorph,[23] but several other polymorphs with moderately different structures were also found. In comparison, a solid-state NMR study of heparin-fibrillized 0N4R tau found a single conformation for the β-sheet core,[24] manifested by a single set of peaks in 2D and 3D correlation NMR spectra. Spectral assignment revealed that this β-sheet core spans residues H268−K340 and adopts a β-arch that places the R3 hexapeptide motif in close proximity to the ^{291}CGS^{293} triplet in R2, similar to the CBD tau fold.[23] Dynamics NMR experiments showed that the protein becomes increasingly mobile from this rigid core toward the two termini: the R1 to R4 domains are semirigid; the proline-rich region and the R′ domain are semimobile with order parameters of ∼0.4, whereas the N-and C-termini are nearly isotropically mobile.[24]

Although 4R tau fibrils of both *in vivo* and *in vitro* origins have now been investigated in detail, *in vitro* 3R tau is much less understood. A single study of heparin-fibrillized 2N3R tau[23] showed a small β-sheet core spanning residues G272−H330, which packs in parallel with another protofibril to form an asymmetric dimer. The unique cysteine of the protein, C322, is packed closely between the two monomers, suggesting the possible formation of an intermolecular disulfide bond.[23] This cryo-EM structure differs dramatically from the *in vivo* PiD 3R tau structure, which exhibits a larger, 93-residue, β-sheet core folded into a strand-turn-strand topology as a single protofibril. Given the unusual features of the *in vitro* 2N3R tau structure and the fact that proteolysis during fibril formation has been shown to cause polymorphism to *in vitro* tau fibrils,[24] we set out to investigate the heparin-fibrillized 3R tau structure in more detail. Such *in vitro* studies of both 4R and 3R tau filaments are also necessary for understanding how 3R and 4R tau proteins mix in dual-isoform tauopathies such as AD.

Here, we present a study of the molecular conformation and hydration of heparin-fibrillized 0N3R tau using magic-angle-spinning (MAS) solid-state NMR and complementary biophysical measurements. Trypsin fingerprinting and J-coupling-based 2D correlation experiments indicate that the entire CT of the protein is immobilized, in stark contrast to 0N4R tau and all *in vivo* tau fibrils known to date. 2D and 3D dipolar correlation NMR spectra confirm that the CT is part of the β-sheet core, together with R3, R4, and R′. Long-mixing time spectra indicate that the CT folds over the R3 domain. The incorporation of the entire CT into the fibril core suggests important biological roles of the CT in regulating the formation of pathological tau filaments.

MATERIALS AND METHODS

Expression of Unlabeled 0N3R Tau. The 0N3R gene was cloned into a pET-28a vector, and the plasmid was transfected into *E. coli.* BL21 (DE3) competent cells (Novagen). A starter culture was grown in 20 mL of LB medium containing 50 μg/mL kanamycin. After overnight growth at 37 °C under shaking at 215 rpm, a 10 mL culture was inoculated into 1 L of LB medium containing 50 μg/mL kanamycin. Cells were grown at 37°C and 215 rpm until the OD_{600} reached 0.8−0.9 and were induced with 1 mM IPTG for 3 h.

Expression of Isotopically Labeled 0N3R Tau. Uniformly ^{13}C, ^{15}N-labeled 0N3R tau with reverse-labeled lysine was expressed in *E. coli* and purified by heat denaturation, cation exchange chromatography, and reversed-phase HPLC. Specifically, a starter culture was grown in 20 mL of LB medium containing 50 μg/mL kanamycin. After overnight growth at 37 °C under shaking at 215 rpm, a 10 mL overnight culture was inoculated into 1 L of LB medium containing 50 μg/mL kanamycin. Cells were grown at 37°C and 215 rpm until the OD_{600} reached 0.5. Cells were spun down at 3000 g and 4°C for 5 min, and the cell pellet was resuspended in 1 L of minimal media containing M9 salts (1x), 0.5 g/L $^{15}NH_4Cl$, 3 g/L of $^{13}C_6$ labeled D-glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 mg/L kanamycin, and vitamin/mineral supplements. We also added 100 mg/L unlabeled lysine to reduce spectral congestion. This concentration suppressed the lysine intensities to ∼20% of those of fully labeled residues, thus weak intensities are detected for some of the Lys residues. Cells were then regrown in minimal medium at 37 °C and 215 rpm for 1.5 h before induction with 0.5 mM IPTG for 4 h.

Purification of 0N3R Tau. Cells from 1 L expression were pelleted and resuspended in 50 mL of lysis buffer containing 20 mM MES (pH 6.8), 1 mM EGTA, 0.2 mM $MgCl_2$, 5 mM DTT, and 1x cOmplete protease inhibitor cocktail (Roche). Cells were lysed using a probe sonicator. The cell lysate was then boiled in a hot water bath for 20 min and spun at 40000 *g* for 40 min to remove cell debris and protein aggregates. The supernatant containing 0N3R tau was first purified with a cation exchange column (self-packed with SP Sepharose Fast Flow resin, GE Healthcare). Fractions containing 0N3R tau were collected and further purified by reversed-phase HPLC (Agilent Zorbax 300SB-C3 column, 21.2 × 250 mm, 7 μm particle size) using an acetonitrile gradient of 5−50% in 45 min. HPLC fractions (>95% pure, determined by SDS-PAGE with Coomassie blue staining) were pooled and lyophilized to yield 0N3R tau as a powder. The yields of purified proteins are ∼10 mg and ∼3.5 mg per liter culture for unlabeled and labeled 0N3R tau, respectively.

0N3R Tau Fibrillization and Kinetics. 0N3R Tau was fibrillized in sealed 1.5 mL eppendorf tubes containing 0N3R tau (1 mg/mL), TCEP (1 mM, pH adjusted to pH 7), and heparin (0.16 mg/mL, Santa Cruz Biotech, sc203075, 8000−25000 Da) in 1 x PBS solution (pH 7.4). The solution was shaken at 37°C and 1200 rpm for 3 days. To monitor the aggregation kinetics, 10 μM ThT was added to a small aliquot (300 μL) of the mixture, and the fluorescence signal (λ_{ex} = 444 nm, λ_{em} = 485 nm) was monitored on a Spectramax M5 plate reader (Molecular Devices).

Electron Microscopy. Fibril samples were adsorbed onto freshly glow-discharged, 200-mesh Formvar/carbon-coated copper grids (Ted Pella), washed with 0.1 M ammonium acetate followed by 0.01 M ammonium acetate, and stained with 2% (wt/vol) uranyl acetate. TEM images were taken on an FEI Tecnai T12 electron microscope (UCSF EM core).

Trypsin Fingerprinting. A 100 μL portion of unlabeled 0N3R tau fibrils (1 mg/mL) were treated with freshly prepared trypsin (Sigma T8003, 1 mg/mL) in 1x PBS by incubation at 37°C and 1400 rpm for 30 min. The trypsin-resistant material was pelleted at 100000 *g* and 4°C for 30 min, washed once with 1 x PBS, and resuspended in 24 μL 1 x PBS for SDS-PAGE and mass spectrometry analysis. For mass analysis, 12 μL of resuspended solution was mixed with 100 μL

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

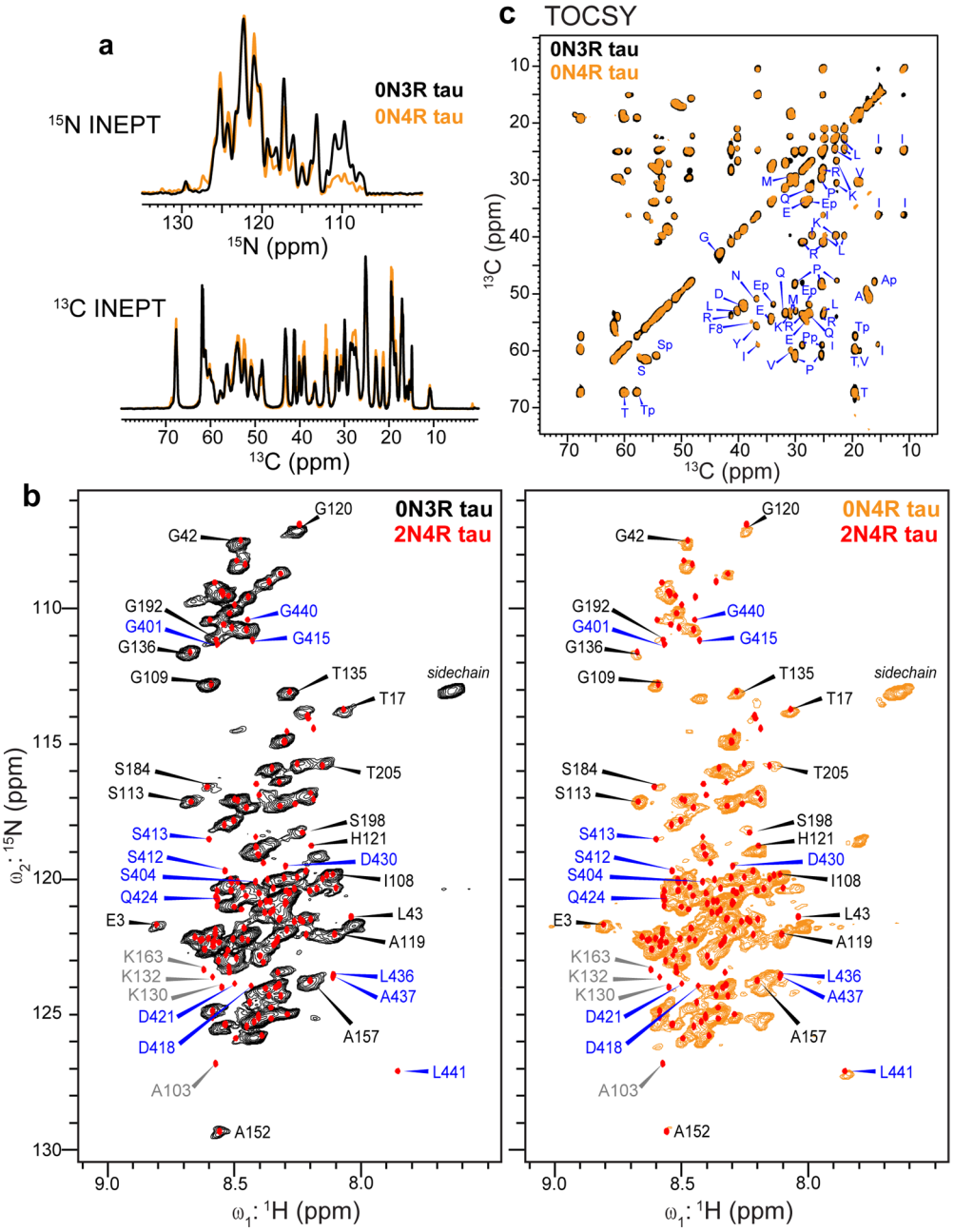


Figure 2. C-terminal domain of 0N3R tau fibrils is immobilized. (a) 1D refocused INEPT ^{15}N and ^{13}C spectra of 0N3R tau (black), overlaid with the 0N4R tau spectra (orange). These INEPT spectra selectively detect highly mobile residues. (b) 2D ^{1}H−^{15}N INEPT correlation spectra of 0N3R tau fibrils (left) and of 0N4R tau fibrils (right).[24] Superimposed onto each MAS spectrum are the solution NMR chemical shifts of 2N4R tau protofibrils (red points).[22] Assignments for the N1 and N2 residues are removed since these domains are absent in 0N3R and 0N4R tau. Resolved N-terminal residues are assigned in black. Resolved 2N4R CT signals (assigned in blue) are missing in the 0N3R tau spectrum, but many of them are present in the 0N4R tau spectrum, indicating that the rigid core of 0N3R tau extends to the C-terminus. Several Lys peaks (assigned in gray) are too weak to be detected in the spectra due to reduced labeling of lysines. The Gly peaks are weaker in the 0N4R tau spectrum than the 0N3R tau spectrum due to reverse labeling of Gly in the former. (c) 2D ^{13}C−^{13}C TOCSY spectrum of 0N3R tau (black), showing residue-type specific random coil ^{13}C chemical shifts. The spectrum agrees well with the 0N4R tau spectrum (orange). Residues that precede Pro, for example, the serine residues marked as Sp, manifest different chemical shifts from those that do not precede Pro.

of denaturing buffer (7.2 M GdnHCl, 100 mM Tris pH 7.5, 10 mM TCEP), and the mixture was incubated at 95°C for 1 h. The sample was desalted using the StageTip (Pierce C18 Tips, Thermo Scientific) and dried by speedVac. The resuspended solution was mixed with sinapinic acid and analyzed on a Shimazu AXIMA Performance MALDI TOF/TOF mass spectrometer. To identify the sequence of the trypsin-resistant material, the average masses were searched against a sequence database generated using ProteinProspector (http://prospector.ucsf.edu/prospector/mshome.htm) with the following settings: protein 0N3R tau, enzyme trypsin, max. missed cleavages 99, mass type average, mass tolerance 200 ppm.

Solid-State NMR Samples and Experiments. The fibrillized tau aggregates were collected by ultracentrifugation at 100000 *g* for 30 min at 4 °C. Pellets were combined and washed twice with water to reduce salt concentration to an estimated ∼10 mM. The final pellet contained ∼20 mg of dry tau mass with 3-fold excess water. This was reduced to 1-fold excess through slow desiccation, then the hydrated fibrils were transferred into a 3.2 mm thin-wall Revolution NMR

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

pencil rotor using tabletop centrifugation with a series of funnels made of plastic pipet tips. The final sample contained ~18 mg 0N3R tau and ~17 mg solution.

All solid-state NMR spectra were measured on a Bruker Avance 800 MHz (18.8 T) spectrometer using a BlackFox 3.2 mm HCN MAS probe. ^{13}C chemical shifts were referenced externally to the adamantane CH_2 chemical shift at 38.48 ppm on the tetramethylsilane (TMS) scale, and ^{15}N chemical shifts were referenced to the ^{15}N peak of ^{15}N-acetylvaline at 122.0 ppm on the liquid ammonia scale. Typical rf field strengths were 54−83 kHz for ^{1}H, 16−50 kHz for ^{13}C, and 25−36 kHz for ^{15}N. ^{1}H chemical shifts of a hydrated DSS-containing POPC membrane sample were used as an external standard to calibrate the ^{1}H chemical shifts and measure the sample temperature. Reported temperatures are estimated sample temperatures based on the probe thermocouple and the measured water ^{1}H chemical shift. Spectra that preferentially detect mobile residues were measured using scalar couplings. These include $^{1}H-^{15}N$ and $^{1}H-^{13}C$ INEPT spectra and 2D $^{13}C-^{13}C$ TOCSY spectra. Spectra that preferentially detect rigid residues were measured using cross-polarization from ^{1}H to ^{13}C, ^{1}H to ^{15}N, and ^{15}N to ^{13}C. In addition, the mixing sequences for these rigid-residue selective 2D and 3D experiments are dipolar in nature, including a $^{13}C-^{13}C$ CORD[25] and a modified CO-Cα BSH−CP sequence[26] (Figure S1). Most experiments were conducted under 10.5 kHz or 14 kHz MAS (Table S1). For many experiments, a short $^{1}H-^{13}C$ CP contact time of 70 μs was used to select only the most rigid β-sheet core signals while excluding the cumulative dynamic signals from mobile regions of the protein. Long NMR experiments were run in blocks of 1−3 days with field drift corrected between each block. Multiple blocks of these 2D and 3D spectra were added in the time domain before Fourier transformation.

All MAS NMR spectra were acquired and processed using TopSpin, and chemical shift assignment was conducted in SPARKY. Typical processing used either QSINE apodization with SSB = 3 or GM apodization with LB = −20 Hz and GB = 0.05, all in TopSpin. Multidimensional spectra are plotted with a 1.2 x scale between successive contour lines. Backbone (ϕ, ψ) dihedral angles were calculated from the assigned chemical shifts using the TALOS-N software,[27] after adding 2.0 ppm to all TMS-referenced experimental ^{13}C chemical shifts to convert to the DSS scale. Long-range contacts were obtained from four long-mixing 2D and 3D spectra.

Water-edited 2D CC and 2D NCA spectra were measured using a water-selective ^{1}H echo that consists of a 180° Gaussian pulse lasting 10 rotor periods in duration, surrounded by one rotor period before and after the pulse. This water-selective T_2 period was followed by a ^{1}H mixing period of 4 ms to transfer the water ^{1}H polarization to protein protons. With vanishing mixing time, less than 0.1% of the protein signal remained in the spectrum, while a 100 ms ^{1}H mixing period produced an equilibrated protein spectrum with 50% of the full intensity of the unedited protein spectra. This is roughly consistent with the similar mass of protein and water in the rotor. Therefore, for calculating the water transfer ratio S/S_0, we used the unedited 2D CC and NCA spectra multiplied by a factor of 0.50 as the S_0 intensity. The S/S_0 intensity ratios were extracted from the individual peaks in the 2D spectra, after correcting for the number of scans, and after averaging over both sides of the diagonal for the 2D CC spectra. Error bars were propagated from the signal-to-noise ratios of 1D cross-sections of resolved peaks in the 2D spectra, using empty regions of the 2D spectra as the noise.

RESULTS

Heparin-Fibrillized 0N3R Tau Exhibits a Single Predominant Ultrastructural Morphology. We expressed recombinant 0N3R tau in *E. coli* and purified it by heat denaturation, cation exchange chromatography, and reverse-phase HPLC. These purification conditions minimize trace contaminants, which can affect tau nucleation and hence fibril morphology. Fibrils were formed by adding 0.16 mg/mL heparin to 1.0 mg/mL tau monomers in phosphate buffer that contains 1 mM of the reducing agent TCEP. The solution was shaken at 1200 rpm at 37 °C for 3 days. The tau and heparin concentrations were optimized to maximize the pelletable mass, as assessed by SDS-PAGE. These conditions gave a fibrillization yield of >90%. Thioflavin-T fluorescence data (Figure 1b) showed that fibrils formed with a lag time of ~33 h, much longer than the ~1 h lag time of 0N4R tau fibrils prepared under similar conditions. TEM images show ~23 nm wide filaments that are predominantly straight, together with a small fraction (<15%) of twisted fibrils (Figure 1c). Trypsin digestion followed by mass spectrometry revealed that the dominant trypsin-resistant fragment has a molecular weight of 15.1 kDa, which corresponds to residues K268 to L441 (Figure 1d). This segment encompasses the end of R1, the full R3, R4 and R′ repeats, and the entire CT. This expanded core is unexpected, as no tau fibrils known to date have a β-sheet core that extends beyond E380.

2D NMR Spectra Confirm That the Rigid Domain of 0N3R Tau Extends to the C-Terminus. To verify the size of the fibril core and obtain site-specific structural information, we measured one- and two-dimensional MAS NMR spectra of ^{13}C, ^{15}N-labeled protein. To simplify the spectra, we suppressed Lys labeling by adding 100 mg/L unlabeled Lys to ^{13}C, ^{15}N-labeled growth media. 1D ^{15}N spectra (not shown) indicate a residual Lys labeling level of ~20%. We preferentially detect dynamic residues using J-coupling based experiments and immobilized residues using cross-polarization (CP) based experiments. 1D ^{13}C and ^{15}N INEPT spectra (Figure 2a) show narrow line widths of 0.4 ppm for ^{13}C (resolution limited by $^{13}C-^{13}C$ J-coupling) and 0.4 ppm for ^{15}N, and the intensity envelopes are similar to that of the 0N4R tau spectra. A 2D $^{1}H-^{15}N$ INEPT correlation spectrum (Figure 2b) of 0N3R tau further resolved the amide ^{1}H signals, and its overlay with the previously reported solution-state ^{1}H and ^{15}N chemical shifts of 2N4R protofibrils[22] shows excellent agreement after removing the N1 and N2 residues, which are absent in 0N3R tau. This agreement indicates that the conformation of the dynamic portion of 0N3R tau is similar to that of 2N4R tau. The $^{1}H-^{15}N$ 2D INEPT spectrum of 0N4R tau fibrils[24] also agrees well with the 2N4R tau chemical shifts. These consistencies allow us to transfer the solution-state chemical shifts of 2N4R tau to the solid-state spectra of 0N3R and 0N4R tau fibrils' mobile domains.

It can be seen that many N-terminal residues' peaks in 2N4R tau are present in both 0N3R and 0N4R tau spectra, and their chemical shifts agree well among the three tau samples. These residues include, for example, T17, G42, G109, S113, and S184 (Figure 2b). In contrast, the resolved C-terminal residues' peaks, assigned in blue, are missing in the 0N3R tau spectrum, while many are present in the 0N4R tau spectrum, indicating that the C-terminal domain is not isotropically mobile in 0N3R tau. These missing signals include, for example, those of L441, A437, and L436. Several other peaks such as D418, Q424, and D430 are weak in the 0N4R tau spectrum but are completely missing in the 0N3R tau spectrum. As expected, residue A103, which is preceded by T102 in 2N4R tau but by K44 in 0N3R tau, resonates at different chemical shifts between 2N4R tau and 0N3R tau. Finally, several Lys peaks such as K130 and K163 are absent in the 0N3R spectrum due to the reduced Lys labeling. A 2D $^{13}C-^{13}C$ TOCSY spectrum (Figure 2c) mostly exhibits residue-type random coil chemical shifts, confirming the presence of isotropically dynamic residues in 0N3R tau. The only exceptions are residues preceding Pro, whose

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

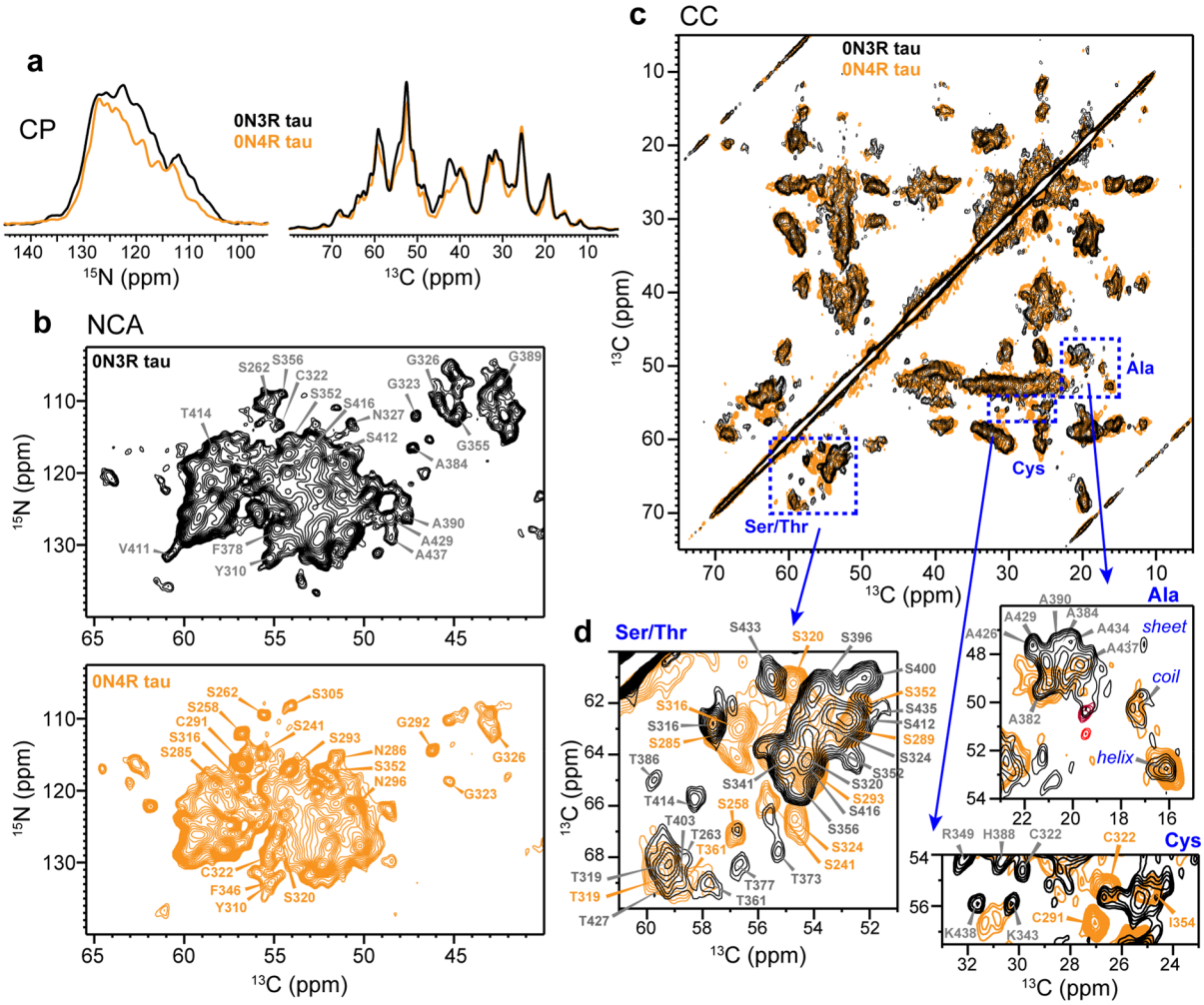


Figure 3. Fingerprint CP-MAS NMR spectra of 0N3R tau, showing the signals of immobilized residues. (a) 1D ^{13}C and ^{15}N CP spectra of 0N3R (black) and 0N4R (orange) tau fibrils. The ^{13}C spectra were measured using a 70 μs ^{1}H−^{13}C CP contact time to suppress the signals of mobile residues. (b) 2D NCA spectra of 0N3R (black) and 0N4R (orange) tau. The 0N3R tau spectrum displays many more Ala peaks than the 0N4R tau spectrum, consistent with the inclusion of the R′ and CT in the rigid core. (c) 2D ^{13}C−^{13}C correlation spectrum of 0N3R tau (black), overlaid with the 0N4R tau spectrum (orange). (d) Insets of the Ser/Thr, Ala, and Cys regions of the 2D CC spectrum. Assignments obtained from 3D correlation data are indicated for some of the signals.

chemical shifts differ from non-Pro-preceding residues.[28] The TOCSY spectrum of 0N3R tau is superimposable with the TOCSY spectrum of 0N4R tau within the spectral line width, confirming that the lack of R2 does not change the dynamically averaged conformation of the fuzzy coat.

To directly observe rigid β-sheet residues, we turned to ^{13}C and ^{15}N CP-MAS experiments (Figure 3a). The fingerprint 2D ^{15}N−^{13}Cα (NCA) correlation spectrum shows line widths of 0.5−0.7 ppm for ^{13}C and 1.2−1.5 ppm for ^{15}N for resolved peaks (Figure 3b and Figure S2), indicating that the β-sheet core has a relatively homogeneous molecular conformation. Compared to 0N4R tau, 0N3R tau exhibits more Ala peaks, which resonate at 47−50 ppm for Cα and 117−130 ppm for ^{15}N. Tau has no Ala between residues 250−380 (spanning R1, R3, R4, and R′), whereas residues 380−441 (spanning R′ and the CT) include seven Ala (Figure 1a, Table S2). Thus, the large number of Ala peaks in the NCA spectrum is consistent with the inclusion of the C-terminus into the β-sheet core. Analysis of 3D correlation spectra (*vide infra*) allowed us to assign all seven Ala, confirming this conclusion.

Complementing the 2D NCA spectrum, we measured a 2D ^{13}C−^{13}C (CC) correlation spectrum using 23 ms ^{13}C spin diffusion and a short ^{1}H−^{13}C CP contact time of 70 μs to suppress the signals of dynamic residues (Figure 3c and Figure S3). We focus on the Cys, Ser/Thr, and Ala regions of the spectrum to assess the size and conformational homogeneity of the β-sheet core (Figure 3d). Assignment of some of the well-resolved peaks based on 3D correlation spectra (*vide infra*) is indicated. The longest isoform of tau contains only two cysteines, C291 and C322, at equivalent positions in the R2 and R3 repeats. C322 is present in both 3R and 4R tau whereas C291 only exists in 4R tau. On the basis of the characteristic chemical shifts of reduced Cys and connectivities in the 3D correlation spectra, we assigned the C322 Cα-Cβ peak at (54.8, 30.0) ppm. This peak is weaker and has different chemical shifts from the C322 peak in 0N4R tau,[24] indicating that the C322 structure and dynamics differ between the two proteins. In the Ser/Thr region of the spectrum, 0N3R tau exhibits a similar number of Ser peaks as in 0N4R tau but more Thr peaks than 0N4R tau.[24] Full-length 3R and 4R tau isoforms both contain 4 Thr's in R1−R4 repeats and 6 Thr's in

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

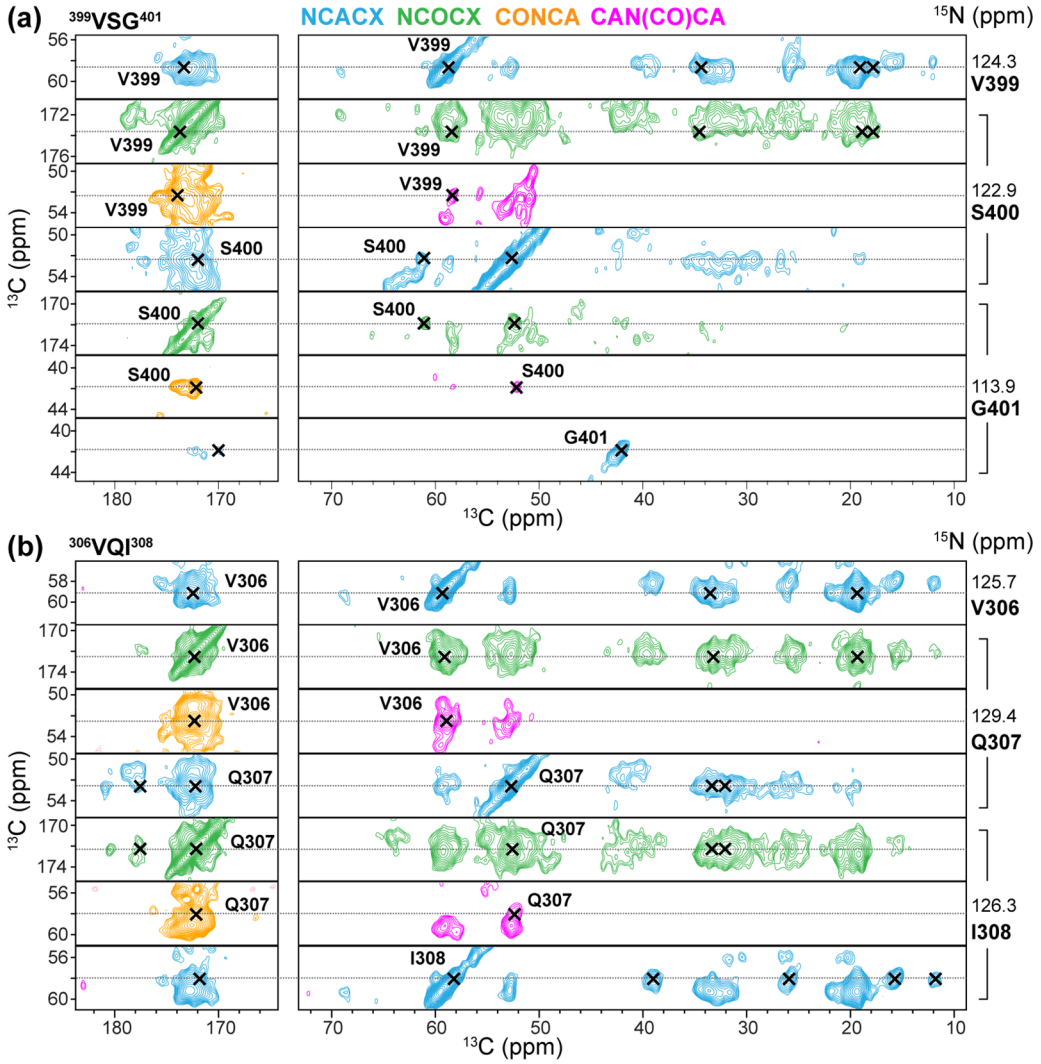


Figure 4. Representative strips of four 3D correlation spectra of 0N3R tau fibrils for chemical shift assignment. (a) Assignment of the ^{399}VSG^{401} triplet. (b) Assignment of the ^{306}VQI^{308} segment in the hexapeptide motif at the beginning of R3. The NCACX (blue) spectrum shows the intraresidue chemical shifts of residue *i*, while the NCOCX (green), CAN(CO)CA (lavender), and CONCA (orange) spectra correlate the chemical shifts of residue *i*-1 with the chemical shifts of residue *i*. Additional strips can be found in Figure S4.

the R′ and CT (Table S2). Thus, the detection of a larger number of Thr peaks in the CC spectrum supports an expanded β-sheet core in 0N3R tau. In the well-resolved Ala region of the spectrum, both 0N3R and 0N4R tau exhibit random coil, α-helical, and β-sheet Cα-Cβ cross-peaks. However, the coil and helical Ala signals are absent in the 3D spectra of 0N3R tau, indicating that the non-β-sheet Ala residues are semimobile. In the β-sheet Ala chemical shift region, more cross-peaks are observed for 0N3R tau than for 0N4R tau, further supporting the larger β-sheet core in 0N3R tau.

Chemical Shift Assignment of 0N3R Tau from 3D Correlation Spectra. To determine the site-specific backbone conformation of the 0N3R tau fibril core from ^{13}C and ^{15}N chemical shifts, we measured and assigned four 3D ^{15}N−^{13}C-^{13}C correlation spectra. The NCACX spectrum correlates intraresidue ^{15}N, ^{13}Cα, and sidechain ^{13}C chemical shifts, whereas the NCOCX, CONCA, and CAN(CO)CA spectra correlate the chemical shifts of two sequential residues.[26,29−32] An 80 ms ^{13}C CORD mixing period was used in the NCACX and NCOCX experiments. Using standard reverse backbone walks from residue *i* to *i*-1 and forward walks from residue *i* to *i*+1, we readily assigned 20 residues (SI Methods). This includes, for example, ^{400}SG^{401}, which exhibits a well-resolved CONCA peak at (172, 114.3, 41.9) ppm for the S400 ^{13}CO, G401 ^{15}N, and Cα resonances (Figure 4, Figure S4), which are confirmed by peaks in the other three 3D spectra. Additional assignments were obtained by augmenting the backbone walk with several additional strategies, including exploiting the sequence uniqueness of amino acid triplets for residues of the same type and using nonsequential short and medium-range inter-residue correlations to validate the assignment (SI Methods). Using this integrated approach, we assigned 104 out of 149 residues between S262 and L441 (Table S3, Table S4). The percentage of assigned residues is similar to that of 0N4R tau,[24] even though the core of 0N3R tau is twice as large. Importantly, all intense and resolved peaks in the 3D spectra are assigned. As a result, any minor conformation would represent at most 10−15% of the fibrils based on the spectral sensitivity, in agreement with the TEM data (Figure 1c). The 45 unassigned residues out of 149 either give weak signals due to low labeling

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

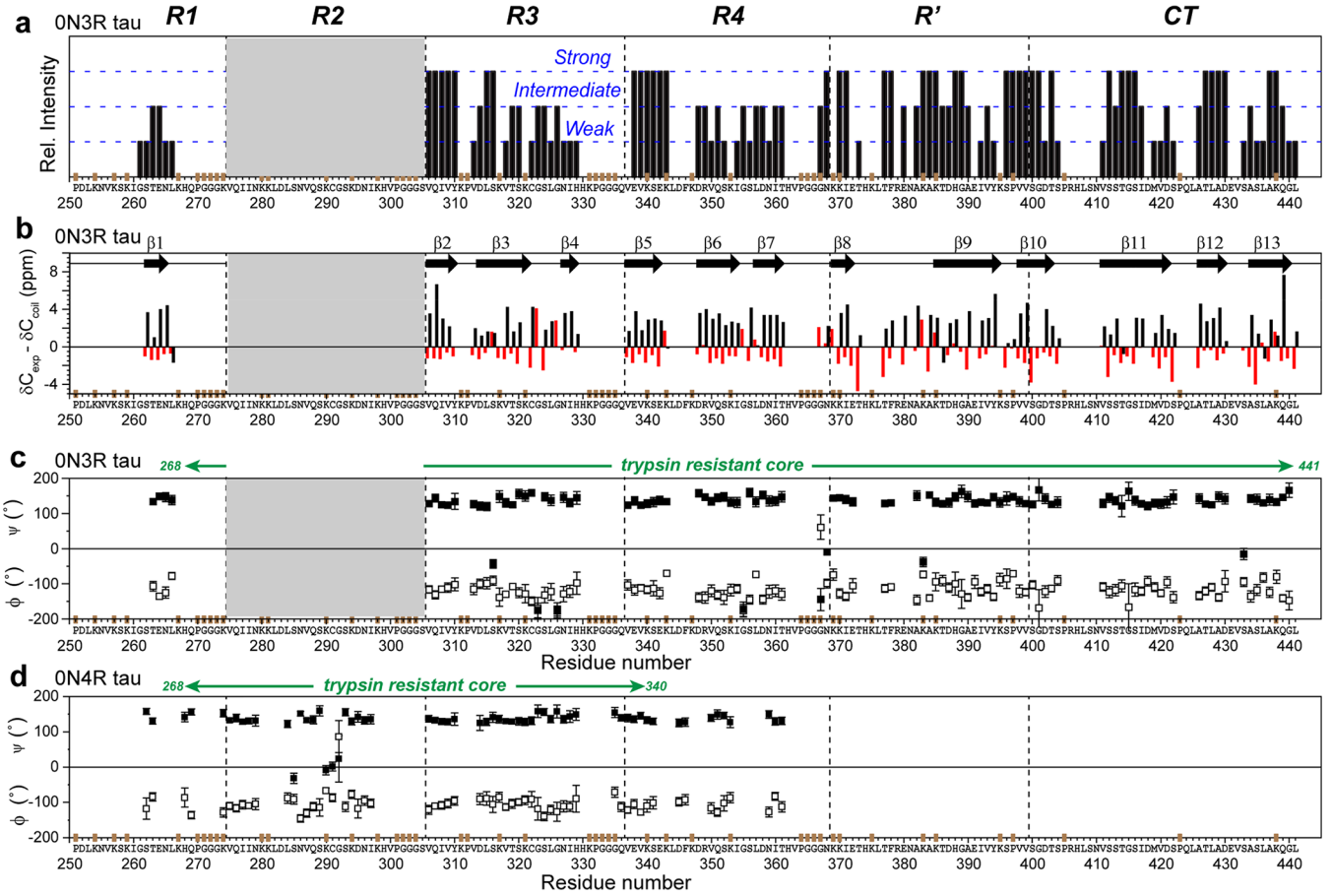


Figure 5. Backbone conformation of the 0N3R tau fibril core obtained from ^{13}C and ^{15}N chemical shifts. (a) Relative peak intensities of assigned residues, categorized qualitatively as strong, intermediate and weak based on comparing resolved peaks of the same type. Strong peaks indicate rigid and ordered residues. R1 residues have weaker intensities than residues in other domains. The positions of Pro (P), GGG triplets, and unlabeled Lys (K) residues are indicated in brown on the *x*-axis. (b) Cα (red) and Cβ (black) secondary chemical shifts. Negative Cα and positive Cβ secondary shifts are indicative of a β-strand conformation. Thirteen β-strands can be identified (arrows). Since not all residues are assigned, the exact number of distinct β-strands in the protein may differ. (c) Backbone torsion angles predicted from the measured chemical shifts.[27] (d) (ϕ, ψ) torsion angles of heparin-fibrillized 0N4R tau for comparison.[24]

levels and dynamics or are unresolved due to sequence repetition such as PGGG (Figure S5).

The resolved and assigned peaks in the aliphatic region of the 2D CC spectrum show a single set of chemical shifts per residue, indicating a single conformation for the β-sheet core (Figure S3). The sole cysteine, C322, has a Cβ chemical shift (29.9 ppm) indicating a reduced thiol. We observed 30 out of 42 residues in the CT, including the terminal ^{437}AKQGL^{441}, confirming the rigidity of the C-terminal residues. The intensities of R3, R4, R′, and CT residues are similarly high in these dipolar correlation spectra, while the R1 signals are less intense (Figure 5a), consistent with the start of the core from the middle of R1. Relative peak intensities of assigned residues are categorized qualitatively as strong, intermediate, and weak based on comparison between resolved peaks of the same residue type. Most Cα chemical shifts are smaller than the random coil chemical shifts whereas most Cβ chemical shifts are larger, indicating that the core adopts a predominantly β-sheet conformation (Figure 5b). The (ϕ, ψ) torsion angles obtained from these chemical shifts[27] indicate 13 β-strands in the trypsin-resistant core (Figure 5c). 0N3R tau shares common β-strands with 0N4R tau at V306−Y310 and V313−H330 in R3, V337−S341, V350−K353, and N359-T361 in R4, and G261-S262 in R1 (Figure 5d). However, we did not detect β-sheet residues past T361 in 0N4R tau, while 0N3R tau exhibits additional β-strands at K385−S400, V411−S422, A426-A429, and A434-G440 (Figure 5c).

Water Accessibility of 0N3R Tau Fibrils. We investigated site-specific water accessibility of 0N3R tau to gain further insights into the fibril core structure. Water polarization transfer to amyloid proteins reveals which side chains are located in dry steric zippers and which ones are well hydrated and likely surface-exposed, thus constraining the three-dimensional fold of the fibril core. We measured water-edited 2D CC and NCA spectra of the protein (Figure S6a,c)[33−35] by selecting the water ^1H magnetization using a ^1H T_2 filter, then transferring it to the protein during a ^1H mixing period. For 2D CC-detected water polarization transfer, hydrogen exchange with labile hydroxy protons is expected to favor the intensities of hydroxy-containing residues such as Ser and Thr. To minimize this effect, we conducted the water-edited 2D experiments at a sample temperature of −5 °C to slow down the chemical exchange rates to less than 10 s^{-1} for Ser and Thr.[35−37] For the water-edited 2D NCA spectrum, water polarization transfer to amide protons is probed, which is relatively insensitive to the presence or absence of side chain protons. We evaluate the water-transferred intensity ratios (S/

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

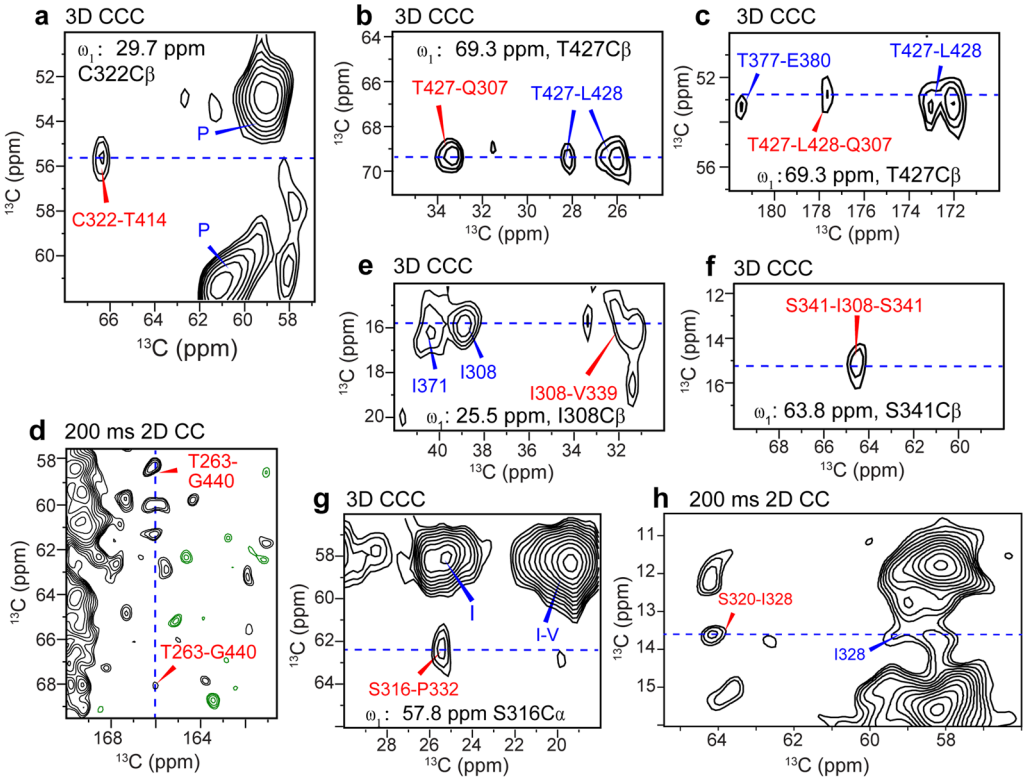


Figure 6. Long-range correlation peaks (assigned in red) that define the tertiary fold of the 0N3R tau fibril. Some of the short-range cross-peaks and amino-acid type assignment of strong peaks are also indicated (blue). (a) C322-T414 correlation. (b) T427-Q307 correlation. (c) L428-Q307 correlations. (d) T263-G440 correlations. (e) I308−V339 correlation. (f) S341−I308 correlation. (g) S316−P332 correlation. (h) S320−I328 correlation. Cross-sections extracted from the dashed blue lines are given in Figure S7.

S_0) between a 4 ms water-edited spectrum (S) and a 100 ms control spectrum (S_0), shown as hydration maps (Figure S6b,d). Using the well-resolved 2D CC cross-peaks of Ser and Thr and the well-resolved 2D NCA peaks of Gly and Ala, we compared the relative hydration of these residues. No systematic intensity differences are detected between Ser/Thr and Ala/Gly, confirming that chemical exchange is slow under our experimental conditions and water-edited CC and NC spectra similarly probe the exposure of the protein backbone to water. Interestingly, the water accessibilities show significant variations among residues of the same type. Among Ser residues, S416 has the highest water accessibility, with an S/S_0 value of 0.38 ± 0.013, while other Ser residues have significantly lower S/S_0 values of 0.23−0.27 (Figure S6e). The water accessibility of S416 can be readily seen in the water-edited 2D CC spectrum as an intense Cα-Cβ cross-peak (Figure S6b). Among Thr residues, T377 exhibits the highest water accessibility. The G415 and G440 N−Cα cross-peaks also show much higher water-transferred intensities than the other Gly. Taken together, these results indicate that the ^{415}GS^{416} segment and the C-terminus are two well-hydrated regions of the protein (Figure S6f).

Long-Range Contacts Reveal the Three-Dimensional Fold of the Rigid Core of 0N3R Tau. To constrain the tertiary structure of the fibril core, we measured four long-mixing spectra: two 2D CC spectra with 200 and 450 ms ^{13}C spin diffusion, a 3D NCACX spectrum with 450 ms spin diffusion, and a 3D ^{13}C−^{13}C-^{13}C (CCC) correlation spectrum with 400 ms spin diffusion (Figure 6 and Figure S7). Inter-residue correlations in these spectra indicate that the respective carbons are within ∼8 Å in space. We count multiple correlations between the same two residues as one contact. Analysis of these four spectra yielded 90 medium-range contacts, four unambiguous long-range contacts, and four ambiguous long-range contacts, which derive from chemical shifts with partial overlap but cannot be explained by any short-range contacts (Table S5). Originally, we found 93 putative long-range peaks in the 3D correlation spectra; however, most of them could be potentially explained by a short-range contact between residues with resonances within 0.5 ppm of the peak. We thus present only the small number of long-range peaks that cannot be explained by any short-range contacts.

These eight long-range contacts provided clear constraints on the tertiary fold of the 0N3R tau core. One set of long-range contacts indicates that R3 is packed against the CT. The 3D CCC spectrum shows a C322-T414 contact at (ω_1, ω_2, ω_3) chemical shifts of (29.5, 55.4, 66.4) ppm (Figure 6a), which can be assigned to a C322Cβ-C322Cα-T414Cβ correlation. This C322-T414 contact is supported by cross-peaks for Q307 polarization transfer to T427 and L428 in the 3D CCC spectrum (Figure 6b,c). Since Q307 is 15 residues before C322 while T427 and L428 are 13−14 residues after T414 in the sequence, these two long-range contacts indicate antiparallel packing of R3 with the CT. In addition, we observed multiple correlations between G440 and T263 (Figure 6d), which constrain the penultimate residue of the protein to be in close proximity with the N-terminus of the trypsin-resistant core. Together, these long-range contacts establish that the CT is stacked against the R1−R3 segment in an antiparallel fashion.

The second set of long-range contacts, between I328 and S320 in R3 and between I308 of R3 and V339 of R4, defines

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

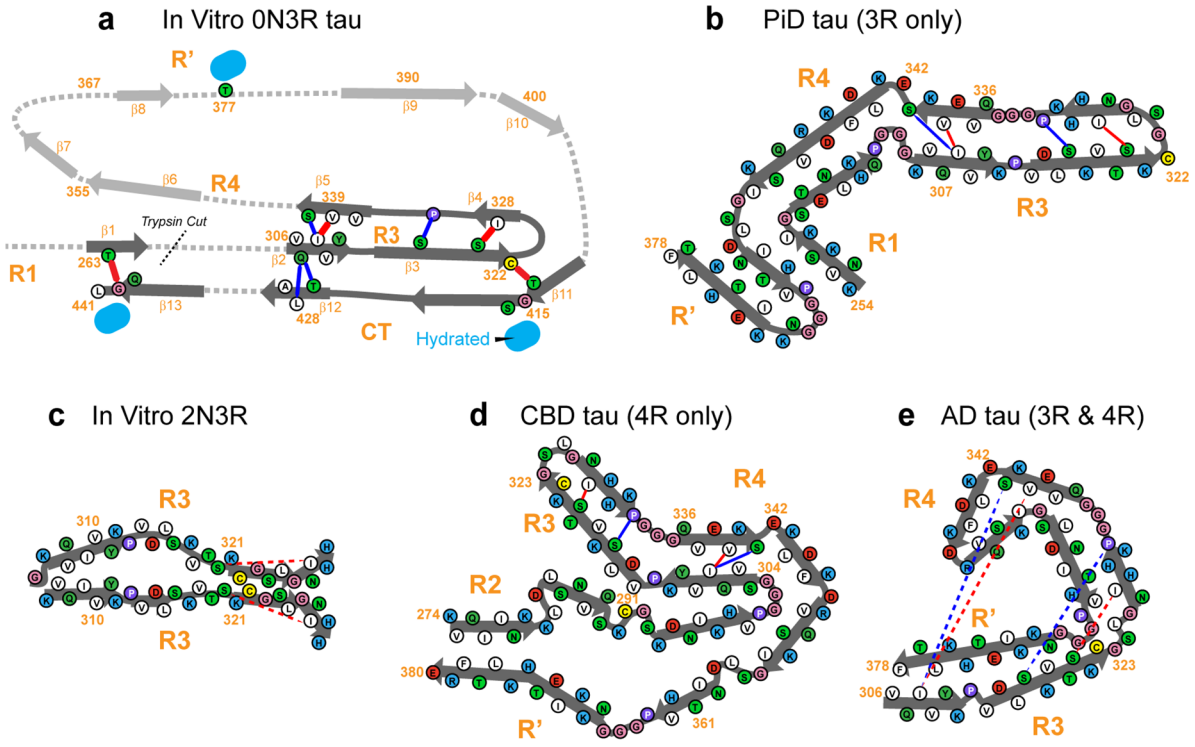


Figure 7. Provisional structural model of 0N3R tau, compared to other tau fibril structures known to date. (a) Tertiary fold of heparin-fibrillized 0N3R tau. Unambiguous long-range contacts are shown as red lines while ambiguous contacts are shown as blue lines. The tertiary fold of residues 342−411 is not known and is indicated by gray arrows and dotted lines, but most of these residues are rigid, ordered, and β-sheet. The R3-CT interactions could also be intermolecular, leading to a domain-swapped dimer, an example of which is shown in Figure S8. (b) Tertiary fold of the Pick's disease 3R tau fibril core (PDB code: 6GX5).[21] (c) Tertiary fold of 2N3R tau fibril core (PDB code: 6QJQ).[23] (d) Tertiary fold of the CBD tau fibril core (PDB code: 6TJO).[7] (e) Tertiary fold of the AD tau fibril core (PDB code: 5O3L).[6] For the structures b−e, the long-range contacts seen in 0N3R tau are annotated where possible, to indicate the different folds of the proteins.

the packing of R3 and R4 repeats. Two cross-peaks define these contacts: a (64.1, 13.6) ppm peak in the 200 ms 2D CC spectrum (Figure 6h) and a (25.9, 16.1, 31.7) ppm peak in the 3D CCC spectrum (Figure 6e). For each peak, the first residue's assignment is unique, while the second residue's assignment can be disambiguated by eliminating assignments that violate the β-strand geometry and contradict the unambiguously assigned peaks. In this way, we can assign the first peak to an S320Cβ-I328Cδ1 contact and the second peak to an I308Cδ1-V339Cβ contact (Table S5).

These long-range contacts, although sparse, define the major qualitative features of the three-dimensional fold of the 0N3R tau fibril core (Figure 7a). In this provisional model, the fibril core has an elongated C-shape resembling an alligator head: the R3 hexapeptide motif ^{306}VQIVYK^{311} forms a steric zipper with the R4 pseudohexapeptide motif ^{337}VEVKSE^{342}, whereas the other side of R3 stacks against CT. The R3-R4 β-arch is accomplished by a turn at G323-G326 and resembles the R3-R4 β-arch in PiD tau (Figure 7b). However, the steric zipper registry in the heparin-fibrillized 0N3R tau may be shifted by two residues from the PiD tau, in which V306 is in close contact with V339 instead of S341. In 0N3R tau, the CT packs against the other side of R3, stretching from T414, which contacts C322, past T427, which contacts the R3 hexapeptide motif, to the C-terminal G440 and L441, which contact T263 in R1. The hydration data are consistent with this provisional model, showing that the last 30 residues of the CT domain have higher hydration than the rest of the fibril core (Figure S6f). This tertiary fold does not account for the stretch of residues from E342 to N411, which form the top of the C-shape, even though the majority of these residues are assigned and are relatively rigid based on their intensities. The lack of tertiary structure information for these residues is due to a lack of information about long-range contacts involving these residues. The R′ region is expected to be packed against the second half of R3 to establish the R3-CT contacts. Peaks defining these long-range contacts involving R′ most likely exist in our current spectra, but are overlapped with resonances representing potential short-range contacts, and thus cannot be used as evidence of the tertiary fold. Further experiments are needed to test this hypothesis and define the structure of this region of the protein.

DISCUSSION

The NMR and biochemical data presented here unambiguously indicate that the β-sheet core of *in vitro* 0N3R tau fibrils adopts a single predominant molecular conformation that contains a reduced C322, and this β-sheet core encompasses the entire R3, R4, R′, and CT domains. The latter differs qualitatively from all *in vivo* and *in vitro* tau fibril core structures known to date. Compared to *in vivo* tau structures, the present 0N3R tau fold shows the highest resemblance to CBD 4R tau[9] (Figure 7d) and partial resemblance to PiD 3R tau[21] (Figure 7b). All three proteins share a common R3-R4 side chain-stacked β-arch. CBD tau also possesses the same topology of an elongated C-shape as 0N3R tau, despite having four microtubule-binding repeats. Due to the additional R2 repeat, the outermost β-strand in CBD tau is in R′ instead of

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

the CT. In both CBD and PiD tau structures, the R3 residues ^{322}CGSLG^{326} participate in the turn of the strand-turn-strand fold, and the R3 and R4 hexapeptide motifs stack to form a steric zipper, with I308 contacting V337 and V339. The current 0N3R tau model has a similar steric zipper, but I308 shifts in registry slightly to contact V339 and S341 (Figure 7a). In CBD tau, the R3 hexapeptide residues Q307 and V309 form another steric zipper with R2 residues around I297. In comparison, the current heparin-fibrillized 0N3R tau displays a second steric zipper between Q307/V309 and T427 in the CT. AD tau[6] (Figure 7e) also exhibits a C-shaped fold, but the core is much smaller. Future characterization of the most stable steric zippers in these tau fibrils will be of interest to obtain predictive insights into the folding landscape of the tau protein.

A recently reported heparin-fibrillized 2N3R tau structure[23] (Figure 7c) found a very small fibril core containing only the R3 domain; moreover, the fibril core forms an asymmetric dimer, potentially stabilized by an intermolecular C322−C322 disulfide bond. However, this same study reported polymorphic 0N4R tau fibrils, which were subsequently reproduced under conditions of proteolysis during fibril formation.[24] Addition of protease inhibitors removed the polymorphism and yielded a single morphology for the 0N4R tau fibrils. On the basis of these observations, we suggest that the unusually small 2N3R tau β-sheet core from this study may have arisen from proteolysis during fibril formation or may have resulted from differences in fibrillization conditions such as protein concentration, buffer composition, or shaking.

The most significant and striking finding of the current study is the incorporation of the entire C-terminal domain of tau into the fibril core, which has not been observed in any other tau fibril structures to date. However, *transient* interactions between CT and the rest of the protein have been reported in many studies. For example, fluorescence resonance energy transfer (FRET) data and electron paramagnetic resonance data of site-specifically labeled 2N4R tau monomers indicate interactions between the CT and R2-R3 residues as well as between the CT and the N terminus of the protein.[38] For example, a FRET distance of ∼20 Å was measured between V432 in the CT and Y310 in R3 in monomeric tau. This is qualitatively consistent with the Q307-T427/L428 contacts seen in the 0N3R tau fibrils. Solution NMR studies of 2N4R tau protofibrils showed that residues 409−426 in the CT transiently contact C322 in R3,[22,39] also consistent with the current results. These data suggest that the CT-R3 interaction observed in the 0N3R tau fibril may represent a solidification of the transient contacts of the protein sampled in solution.

Since a larger amyloid is harder to break up than smaller β-sheet cores, why has this large β-sheet core not been observed in tau filaments from patient brains? Would this expanded core of 3R tau be toxic or protective, if it is present *in vivo*? We hypothesize that the CT-incorporated β-sheet core is protective against pathogenic fibril formation for the following reasons. The CT is chemically heterogeneous *in vivo* due to posttranslational modification and proteolytic truncation. Heterogeneous phosphorylation of 13 residues in the CT,[40] along with caspase cleavage at D421 and other residues,[41] are well documented in diseased brains. Caspase cleavage that removes the last ∼20 residues of the protein has been found to be an early event in aging and AD pathology.[41] Given this chemical heterogeneity, it is difficult to imagine that a C-terminal amyloid could be as cooperatively formed into infinite sheets *in vivo* as the less heterogeneous inner core that contains the more amyloidogenic R3-R4 regions. Instead, it would appear that the CT would be protective by folding back over the inner core, preventing side-on nucleation, a key step in amyloid formation.[42] Indeed, a large body of biochemical data has shown that truncations of the CT as well as the N-terminal disordered region accelerate fibril formation *in vitro*,[43,44] accelerate fibril propagation between cells,[45] enhance seeding by AD brain tau,[43] and increase tau phosphorylation.[1,43,46] These biochemical data thus support the notion that the CT is protective against pathological aggregation, by shielding the aggregation-prone microtubule-binding repeats from interacting with each other and from hyperphosphorylation. This protective effect implies that any CT-containing fibril core that might nucleate *in vivo* would be out-competed in growth by smaller and faster-aggregating species, hence making a CT-including expanded β-sheet core minimally present in diseased brains. This hypothesis is supported by the fact that 0N3R tau fibrillized with a 30-fold longer lag time than 0N4R tau under similar *in vitro* experimental conditions,[24] consistent with a high kinetic barrier for including the CT into the β-sheet core. We speculate that the lack of the C291-containing R2 domain in 3R tau might play a significant role in raising the energy barrier and hence slowing down the rate of fibril formation. C291 participates in crucial steric zipper interactions with the R3 hexapeptide motif in both heparin-fibrillized 0N4R tau[24] and CBD 4R tau.[9] Despite the high kinetic barrier, the fact that we captured this CT-R3 interaction in our 0N3R tau sample suggests that the energetic stability of this structure, when the CT is free of posttranslational modifications, is not substantially weaker than the stability of the smaller cores seen in diseased brains.

Since the long-range R3-CT correlations we observed were obtained from uniformly ^{13}C, ^{15}N labeled protein without mixing, we cannot rule out that some of the cross-peaks might be intermolecular instead of intramolecular in origin. However, even if this were true, for example, caused by protofibril assembly into dimeric species, the same R3-CT interactions would still exist, but in a domain-swapped manner (Figure S8). This would have no impact on the conclusion that the R3 domain is protected and structurally stabilized by the CT.

In conclusion, we have employed solid-state NMR spectroscopy and complementary biophysical measurements to characterize the structure of 0N3R tau fibrils formed using heparin. We find that the rigid core of 0N3R tau fibrils extends from the end of R1 all the way to the C-terminus of the protein in a single molecular conformation. An R3-R4 β-arch and R3-CT antiparallel stacking give rise to an elongated C-shaped fold, which bears similarities to the CBD 4R tau fold. The incorporation of the entire CT into the β-sheet core suggests that the previously reported transient interactions between the CT and other regions of tau can be stabilized under suitable conditions. The lack of such C-terminus-including fibril cores in the brains of tauopathies suggests that cellular mechanisms exist to prevent the CT from aggregating in diseased brain. Elucidation of these mechanisms will be important to provide a fuller understanding of the misfolding pathways of tau in diseases, cellular determinants of distinct tau strains, as well as the folding of tau in healthy brains.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c03314.

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

An NMR pulse sequence, additional 2D and 3D correlation spectra of 0N3R tau, assignment summary, water-edited spectra, and tables of NMR experimental conditions, amino acid sequence distribution, assigned chemical shifts, and inter-residue correlations (PDF)

AUTHOR INFORMATION

Corresponding Author

Mei Hong − *Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0001-5255-5858; Email: meihong@mit.edu

Authors

Aurelio J. Dregni − *Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States*

Harrison K. Wang − *Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States*

Haifan Wu − *Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158-9001, United States*

Pu Duan − *Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0002-7395-4353

Jia Jin − *Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158-9001, United States*

William F. DeGrado − *Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158-9001, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c03314

Author Contributions

[$]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

NMR chemical shifts have been deposited in the Biological Magnetic Resonance Bank (BMRB) with ID number 50785.

ACKNOWLEDGMENTS

This work is supported by NIH grant AG059661 to M.H. and AG002132 to W.F.D. A.J.D. is supported by an NIH Ruth L. Kirschstein Individual National Research Service Award (F31AG069418). This study made use of NMR spectrometers at the MIT-Harvard Center for Magnetic Resonance, which is supported by NIH grant P41 GM132079.

REFERENCES

(1) Wang, Y.; Mandelkow, E. Tau in physiology and pathology. *Nat. Rev. Neurosci.* **2016**, *17*, 5−21.

(2) Wischik, C. M.; Novak, M.; Edwards, P. C.; Klug, A.; Tichelaar, W.; Crowther, R. A. Structural characterization of the core of the paired helical filament of Alzheimer disease. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 4884−8.

(3) Crowther, R. A. Straight and Paired Helical Filaments in Alzheimer-Disease Have a Common Structural Unit. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 2288−2292.

(4) Wegmann, S.; Medalsy, I. D.; Mandelkow, E.; Muller, D. J. The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E313−21.

(5) Goedert, M.; Eisenberg, D. S.; Crowther, R. A. Propagation of Tau Aggregates and Neurodegeneration. *Annu. Rev. Neurosci.* **2017**, *40*, 189−210.

(6) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017**, *547*, 185−190.

(7) Falcon, B.; Zivanov, J.; Zhang, W.; Murzin, A. G.; Garringer, H. J.; Vidal, R.; Scheres, S. H. W. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **2019**, *568*, 420−423.

(8) Falcon, B.; Zhang, W.; Schweighauser, M.; Murzin, A. G.; Vidal, R.; Garringer, H. J.; Goedert, M. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathol.* **2018**, *136*, 699−708.

(9) Zhang, W.; Tarutani, A.; Newell, K. L.; Murzin, A. G.; Matsubara, T.; Falcon, B.; Scheres, S. H. W. Novel tau filament fold in corticobasal degeneration. *Nature* **2020**, *580*, 283−287.

(10) Arakhamia, T.; Lee, C. E.; Carlomagno, Y.; Duong, D. M.; Kundinger, S. R.; Wang, K.; Fitzpatrick, A. W. P. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell* **2020**, *180*, 633−644.

(11) Goedert, M.; Spillantini, M. G. Propagation of Tau aggregates. *Mol. Brain* **2017**, *10*, 18.

(12) Gibbons, G. S.; Lee, V. M. Y.; Trojanowski, J. Q. Mechanisms of Cell-to-Cell Transmission of Pathological Tau A Review. *Jama Neurology* **2019**, *76*, 101−108.

(13) Morozova, O. A.; March, Z. M.; Robinson, A. S.; Colby, D. W. Conformational features of tau fibrils from Alzheimer's disease brain are faithfully propagated by unmodified recombinant protein. *Biochemistry* **2013**, *52*, 6960−7.

(14) Sanders, D. W.; Kaufman, S. K.; DeVos, S. L.; Sharma, A. M.; Mirbaha, H.; Li, A.; Diamond, M. I. Distinct tau prion strains propagate in cells and mice and define different tauopathies. *Neuron* **2014**, *82*, 1271−88.

(15) Taniguchi-Watanabe, S.; Arai, T.; Kametani, F.; Nonaka, T.; Masuda-Suzukake, M.; Tarutani, A.; Hasegawa, M. Biochemical classification of tauopathies by immunoblot, protein sequence and mass spectrometric analyses of sarkosyl-insoluble and trypsin-resistant tau. *Acta Neuropathol.* **2016**, *131*, 267−280.

(16) Goedert, M.; Jakes, R.; Spillantini, M. G.; Hasegawa, M.; Smith, M. J.; Crowther, R. A. Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature* **1996**, *383*, 550−553.

(17) Kampers, T.; Friedhoff, P.; Biernat, J.; Mandelkow, E. M.; Mandelkow, E. RNA stimulates aggregation of microtubule-associated protein tau into Alzheimer-like paired helical filaments. *FEBS Lett.* **1996**, *399*, 344−9.

(18) Andronesi, O. C.; von Bergen, M.; Biernat, J.; Seidel, K.; Griesinger, C.; Mandelkow, E.; Baldus, M. Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 5922−5928.

(19) Daebel, V.; Chinnathambi, S.; Biernat, J.; Schwalbe, M.; Habenstein, B.; Loquet, A.; Lange, A. beta-Sheet core of tau paired helical filaments revealed by solid-state NMR. *J. Am. Chem. Soc.* **2012**, *134*, 13982−9.

(20) Xiang, S.; Kulminskaya, N.; Habenstein, B.; Biernat, J.; Tepper, K.; Paulat, M.; Linser, R. A Two-Component Adhesive: Tau Fibrils Arise from a Combination of a Well-Defined Motif and Conformationally Flexible Interactions. *J. Am. Chem. Soc.* **2017**, *139*, 2639−2646.

(21) Falcon, B.; Zhang, W.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Vidal, R.; Goedert, M. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **2018**, *561*, 137−140.

(22) Bibow, S.; Mukrasch, M. D.; Chinnathambi, S.; Biernat, J.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. The dynamic structure of filamentous tau. *Angew. Chem., Int. Ed.* **2011**, *50*, 11520−4.
(23) Zhang, W.; Falcon, B.; Murzin, A. G.; Fan, J.; Crowther, R. A.; Goedert, M.; Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *eLife* **2019**, *8*, e43584.
(24) Dregni, A. J.; Mandala, V. S.; Wu, H.; Elkins, M. R.; Wang, H. K.; Hung, I.; Hong, M. In vitro 0N4R tau fibrils contain a monomorphic β-sheet core enclosed by dynamically heterogeneous fuzzy coat segments. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 16357−16366.
(25) Hou, G. J.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T. Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J. Magn. Reson.* **2013**, *232*, 18−30.
(26) Shi, C. W.; Fasshuber, H. K.; Chevelkov, V.; Xiang, S. Q.; Habenstein, B.; Vasa, S. K.; Lange, A. BSH-CP based 3D solid-state NMR experiments for protein resonance assignment. *J. Biomol. NMR* **2014**, *59*, 15−22.
(27) Shen, Y.; Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **2013**, *56*, 227−241.
(28) 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* **1995**, *5*, 67−81.
(29) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **1998**, *95*, 1197−1207.
(30) Hong, M. Resonance Assignment of 13C/15N Labeled Proteins by Two- and Three-Dimensional Magic-Angle-Spinning NMR. *J. Biomol. NMR* **1999**, *15*, 1−14.
(31) Rienstra, C. M.; Hohwy, M.; Hong, M.; Griffin, R. G. 2D and 3D 15N-13C-13C NMR chemical shift correlation spectroscopy of solids: assignment of MAS spectra of peptides. *J. Am. Chem. Soc.* **2000**, *122*, 10979−10990.
(32) Hung, I.; Gan, Z. H. Spin-locking and cross-polarization under magic-angle spinning of uniformly labeled solids. *J. Magn. Reson.* **2015**, *256*, 23−29.
(33) Wang, T.; Jo, H.; DeGrado, W. F.; Hong, M. Water Distribution, Dynamics, and Interactions with Alzheimer's β-Amyloid Fibrils Investigated by Solid-State NMR. *J. Am. Chem. Soc.* **2017**, *139*, 6242−6252.
(34) Gelenter, M. D.; Smith, K. J.; Liao, S.-Y.; Mandala, V. S.; Dregni, A. J.; Lamm, M. S.; Hong, M. The peptide hormone glucagon forms amyloid fibrils with two coexisting β-strand conformations. *Nat. Struct. Mol. Biol.* **2019**, *26*, 592−598.
(35) Dregni, A. J.; Duan, P.; Hong, M. Hydration and Dynamics of Full-Length Tau Amyloid Fibrils Investigated by Solid-State Nuclear Magnetic Resonance. *Biochemistry* **2020**, *59*, 2237−2248.
(36) Liepinsh, E.; Otting, G. Proton exchange rates from amino acid side chains - Implications for image contrast. *Magn. Reson. Med.* **1996**, *35*, 30−42.
(37) Williams, J. K.; Hong, M. Probing membrane protein structure using water polarization transfer solid-state NMR. *J. Magn. Reson.* **2014**, *247*, 118−127.
(38) Jeganathan, S.; von Bergen, M.; Brutlach, H.; Steinhoff, H. J.; Mandelkow, E. Global hairpin folding of tau in solution. *Biochemistry* **2006**, *45*, 2283−2293.
(39) Mukrasch, M. D.; Bibow, S.; Korukottu, J.; Jeganathan, S.; Biernat, J.; Griesinger, C.; Zweckstetter, M. Structural Polymorphism of 441-Residue Tau at Single Residue Resolution. *PLoS Biol.* **2009**, *7*, 399−414.
(40) Noble, W.; Hanger, D. P.; Miller, C. C.; Lovestone, S. The importance of tau phosphorylation for neurodegenerative diseases. *Front. Neurol.* **2013**, *4*, 83.
(41) Wesseling, H.; Mair, W.; Kumar, M.; Schlaffner, C. N.; Tang, S.; Beerepoot, P.; Steen, J. A. Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's Disease. *Cell* **2020**, *183*, 1699−1713.
(42) Törnquist, M.; Michaels, T. C. T.; Sanagavarapu, K.; Yang, X.; Meisl, G.; Cohen, S. I. A.; Linse, S. Secondary nucleation in amyloid formation. *Chem. Commun.* **2018**, *54*, 8667−8684.
(43) Gu, J. L.; Xu, W.; Jin, N. N.; Li, L. F.; Zhou, Y.; Chu, D. D.; Liu, F. Truncation of Tau selectively facilitates its pathological activities. *J. Biol. Chem.* **2020**, *295*, 13812−13828.
(44) Gamblin, T. C.; Chen, F.; Zambrano, A.; Abraha, A.; Lagalwar, S.; Guillozet, A. L.; Cryns, V. L. Caspase cleavage of tau: Linking amyloid and neurofibrillary tangles in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 10032−10037.
(45) Stöhr, J.; Wu, H.; Nick, M.; Wu, Y.; Bhate, M. P.; Condello, C.; DeGrado, W. F. A 31-residue peptide induces aggregation of tau's microtubule-binding region in cells. *Nat. Chem.* **2017**, *9*, 874−881.
(46) Berry, R. W.; Abraha, A.; Lagalwar, S.; LaPointe, N.; Gamblin, T. C.; Cryns, V. L.; Binder, L. I. Inhibition of tau polymerization by its carboxy-terminal caspase cleavage fragment. *Biochemistry* **2003**, *42*, 8325−8331.

https://doi.org/10.1021/jacs.1c03314
J. Am. Chem. Soc. 2021, 143, 7839−7851

Supporting Information

Inclusion of the C-Terminal Domain in the β-Sheet Core of Heparin-Fibrillized Three-Repeat Tau Protein Revealed by Solid-State Nuclear Magnetic Resonance Spectroscopy

Aurelio J. Dregni [1$], Harrison K. Wang [1$], Haifan Wu [2], Pu Duan [1], Jia Jin [2], William F. DeGrado [2], and Mei Hong [1*]

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

[2] Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158

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

Additional resonance assignment details for 0N3R tau

The large size of the β-sheet core of 0N3R tau makes its NMR spectra congested, even though the ^{13}C and ^{15}N line widths of individual resolved peaks are actually relatively narrow: 0.5-0.7 ppm and 1.2-1.5 ppm, respectively. These linewidths are similar to or better than the linewidths of some well-studied amyloid proteins such as Aβ. For this large β-sheet core, we found that semi-automated assignment programs[1-3] are not sufficient for spectral assignment, because these programs require a curated peak list in which resonances corresponding to individual residues have been grouped and attributed to amino acid types. This step is the most challenging one for partly congested spectra. Instead, semi-automated assignment programs are more useful for proteins with good NCA and NCO resolution, for which the ambiguous backbone walk is the most challenging step.

To assign the β-sheet core of 0N3R tau, we instead adopted an integrated strategy that augments the conventional backbone walk. We first conducted residue-type assignment using both the 3D NCACX spectrum and the 23 ms 2D CC spectrum, which shows well-resolved correlations for Ala Cα-Cβ, Asp Cβ-Cγ, Glu/Gln Cγ-Cδ, Ile Cβ-Cγ2 and Cβ-Cδ1, Pro Cα-Cβ, Ser Cα-Cβ, Thr Cα-Cβ, and Val Cβ-Cγ. These spectra resolved 7 rigid Ala, 9 Thr, and 7 Pro spin systems, which matched the number of each residue type within the rigid core (residues 262-441) **(Table S2)**. After residue-type assignment, we conducted sequential assignment by walking both forward and backward along the amino acid sequence. The typical sequential assignment process emphasizes backward walks from residue *i* peaks in the NCACX spectrum to *i*-1 peaks in the NCOCX spectrum, with validation by inter-residue correlations in the CONCA and CAN(CO)CA spectra. The forward walk from residue *i* to *i+1* is equally valid, and is helpful when a residue's signals are not well resolved in the NCACX spectrum. In this forward walk, we identify the $^{15}N_{i+1}$ chemical shift of residue *i*+1 with the $^{13}CO_i$, $^{13}C\alpha_i$, and $^{13}Cx_i$ chemical shifts of residue *i* in the NCOCX spectrum. We then match the $^{15}N_{i+1}$ and $^{13}CO_i$ shifts with the $^{13}CO_i$-$^{15}N_{i+1}$ -$^{13}C\alpha_{i+1}$ correlation in the CONCA spectrum and the $^{13}C\alpha_i$-$^{15}N_{i+1}$-$^{13}C\alpha_{i+1}$ correlation in the CAN(CO)CA spectrum. The resulting $^{13}C\alpha_{i+1}$ chemical shift is next combined with the $^{15}N_{i+1}$ chemical shift to specify an NCACX strip, which usually leads to the assignment of residue *i*+1. The joint forward and backward walks allowed us to uniquely assign ^{322}CGS^{324}, ^{326}GN^{327}, ^{341}SEK^{343}, ^{354}IGS^{356}, ^{367}GN^{368}, ^{384}AK^{385}, ^{400}SG^{401}, and ^{414}TGS^{416} (**Table S3** and **S4**).

For backbone connectivities that give a sufficient but not necessary set of assignments due to incompletely resolved peaks, we disambiguated these assignments first using an amino acid type exclusion strategy (**Table S4**). For example, there are seven Ala spin systems in the 3D NCACX spectrum, matching the presence of seven Ala residues in the trypsin-resistant core (**Table S2**). Three Ala residues, in triplets ^{383}KAK^{385}, ^{388}GAE^{390}, and ^{425}LAT^{427}, have unique backward or forward walks and thus can be unambiguously assigned. For the remaining four Ala's in ^{381}NAK^{383}, ^{428}LAD^{430}, ^{433}SAS^{435}, and ^{436}LAK^{438}, the preceding spin system must be N, L, or S while the following spin system must be K, D, or S. By eliminating assignments that do not agree with the preceding or following spin systems, we assigned ^{381}NAK^{383} and ^{428}LAD^{430}. Assignment of these two Ala's then allowed A437 in ^{436}LAK^{438} to be assigned, which in turn allowed the assignment of A434 in ^{433}SAS^{435}. Exchanging any Ala assignments caused inconsistencies with peak connectivities, thus validating the assignment of all seven Ala (**Table S4**, labeled as "type exclusion"). This strategy was similarly applied to assign the 9 Thr in the trypsin-resistant core. We also used this type exclusion strategy to assign Val and Asp residues, although not to

completion. This type exclusion strategy is less useful for Gly, which has weaker N-Cα-CO cross-peaks due to inefficient Cα-CO polarization transfer during the CORD mixing period in the NCACX spectrum and low CO-Cα transfer efficiencies in the BSH-CP mixing period of the CAN(CO)CA spectrum.

In addition to amino acid type exclusion, non-sequential inter-residue cross-peaks in various 3D spectra provided important validation of the sequential assignment. 40 of the 125 rigid residues show short-and medium-range inter-residue correlations that agreed with the sequential assignment and would not agree if alternate assignments were made. We found 90 medium-range correlations ($1 < i\text{-}j < 4$) in the 3D NCACX and CCC spectra and 200 ms 2D CC spectrum (**Table S5b**). Ten residues were sequentially assigned with the help of these inter-residue cross-peaks (**Table S4**). As our assignment neared completion, when no other assignment can be found for some of the initially tentatively assigned residues, then these residues become confidently assigned. 11 residues were assigned in this way (**Table S4,** indicated as "additional backbone walks"). Finally, peak intensities provide additional support to connectivity-based assignment. An intense N-CA-CA peak in the NCACX spectrum should best match an intense CO-N-CA peak, CA-N-CA peak, and N-CO-CO peak in the CONCA, CAN(CO)CA, and NCOCX spectra, because the backbone atoms involved in dipolar polarization transfer between two residues must have similar rigidity and thus peak intensity. Due to β-sheet interdigitation, sidechains are also rigid, thus giving intense sidechain ^{13}C signals. Taking into account peak intensities allowed us to differentiate a single intense peak from several overlapped weak peaks, and led to confident assignment of ^{306}VQIVY^{310} in the R3 hexapeptide motif and residues ^{337}VEV^{339} in the R4 hexapeptide analog (**Table S4,** indicated as "validated by intensity"). The high intensities of these peaks are consistent with the well-known rigidity of the hexapeptide motif in crystal structures of small amyloid-forming peptides[4] and the assigned solid-state NMR spectra of 0N4R tau[5].

Using this integrated assignment strategy, we assigned 104 residues out of the 149 residues between S262 and L441, which show rigid β-sheet signals (**Fig. S5**).

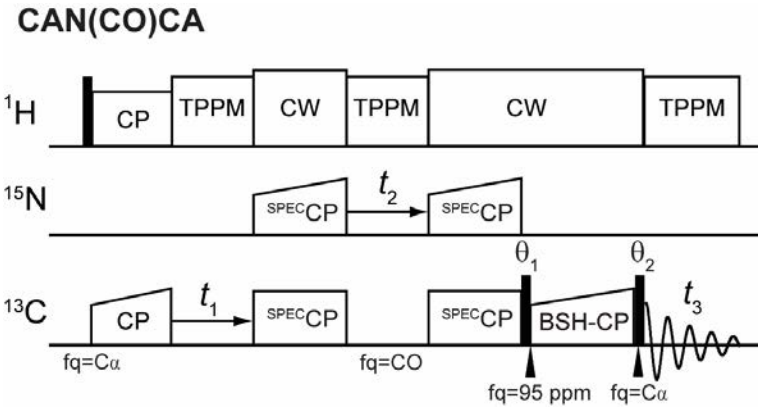


Figure S1. Pulse sequence diagram for the CAN(CO)CA experiment with modified BSH-CP[6] for CO-Cα polarization transfer. Traditionally, the BSH-CP CO-Cα transfer is accomplished with the ^{13}C carrier on resonance with either CO or Cα, and with a spin lock with effective nutation frequencies that satisfy the n=2 double quantum matching condition: $\omega_{1,eff,\,Ca} + \omega_{1,eff,\,CO} = 2\omega_r$, where ω_r is the MAS frequency. However, at 14 kHz MAS and a magnetic field of 18.8 T (800 MHz ^1H, 200 MHz ^{13}C Larmor frequencies), CO and Cα have a ~24 kHz chemical shift separation, thus an applied rf field strength of ~3.8 kHz on either CO or Cα is required. This low rf field was found to give low polarization transfer efficiency. Instead, we placed the RF carrier at 95 ppm, which is one third of the way from Cα towards CO, which allows an RF field strength of ~7 kHz to be used. Because of the weak spin locks on both Cα and CO, two trim pulses were required. An average transfer efficiency of ~45% from CO to Cα was obtained with a 3.5 ms 100-70% amplitude ramp centered on a 7 kHz RF field strength. Polarization transfer to Gly Cα often has far lower efficiency due to the larger resonance offset.

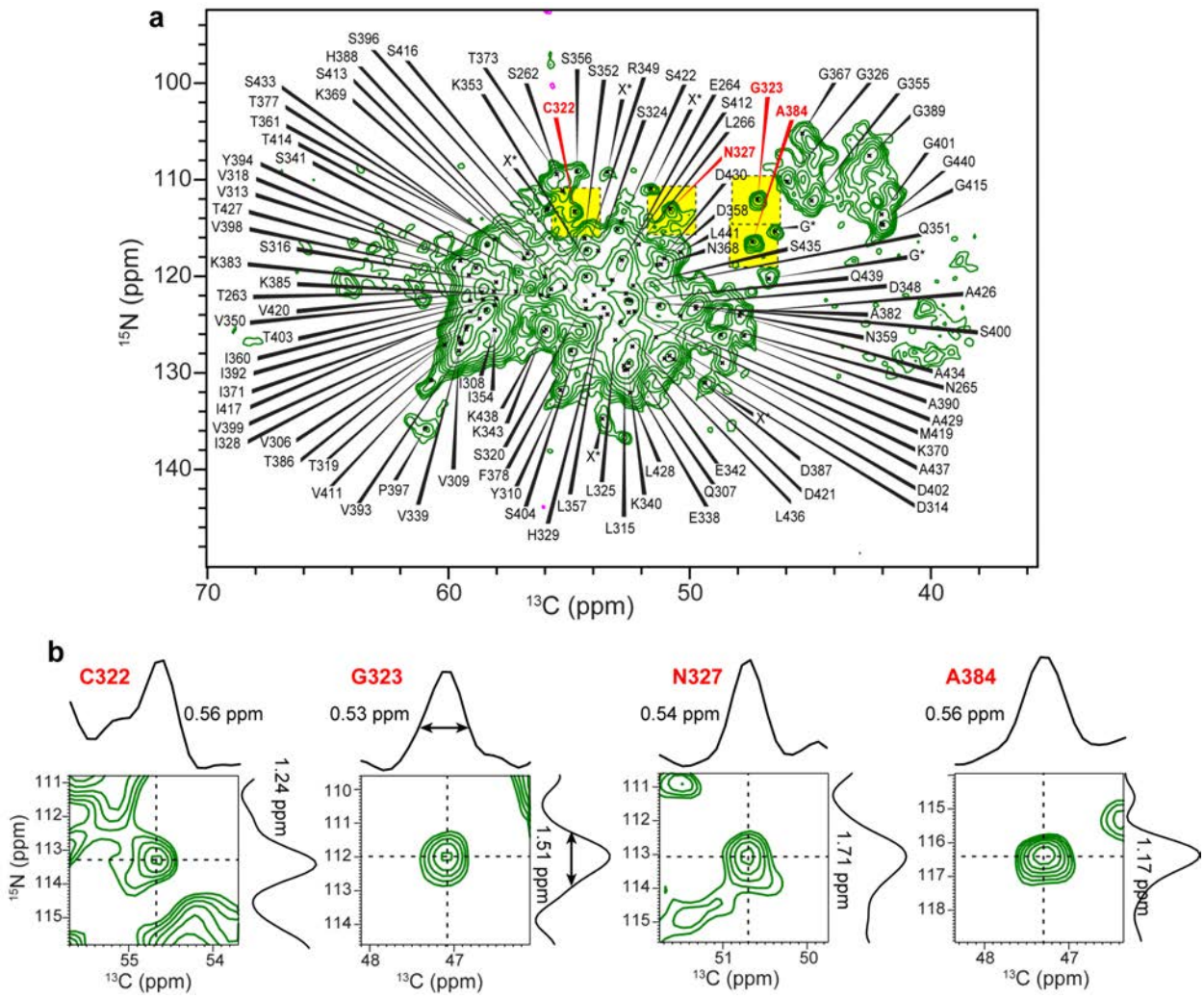


Figure S2. (a) 2D ^{15}N-^{13}Cα correlation spectrum of 0N3R tau fibrils. Resonance assignments obtained from the 3D correlation experiments are given. Type-assigned Gly cross-peaks are marked as G*. Resolved cross-peaks without an assignment are marked as X*. (b) Representative zoomed-in areas that are highlighted in yellow in (a). The ^{15}N and ^{13}C 1D cross-sections are shown to indicate the sensitivity and linewidths of the peaks. The full widths at half maximum are given, as reported from TopSpin through the "peakw" command.

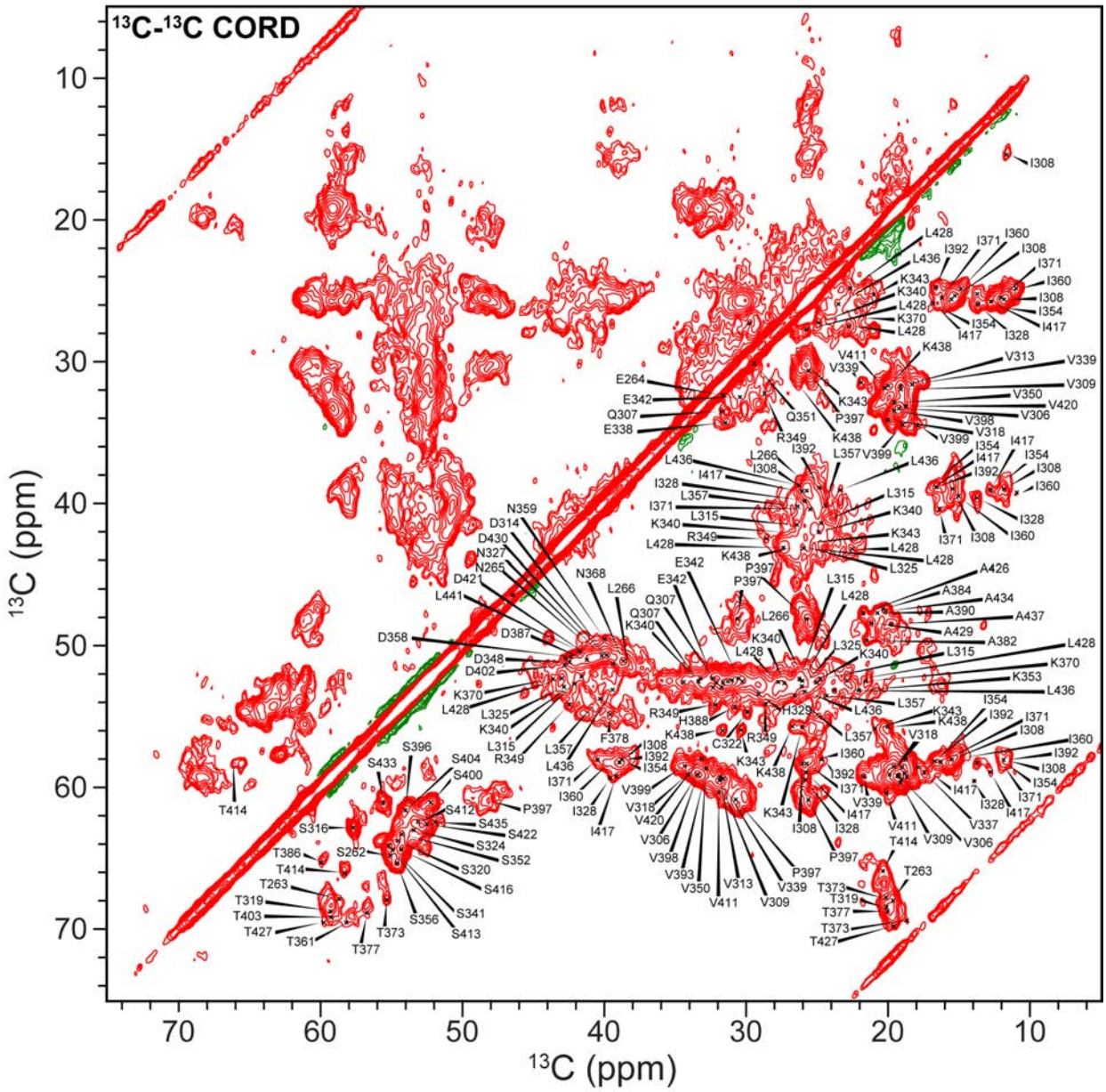


Figure S3. Aliphatic region of the 23 ms 2D ^{13}C-^{13}C CORD spectrum of 0N3R tau fibrils. Resonance assignments obtained from the 3D correlation experiments are indicated. Some cross-peak annotations in the Gln/Glu/Lys Cα-Cβ and Val/Thr Cα-Cγ regions are omitted for clarity.

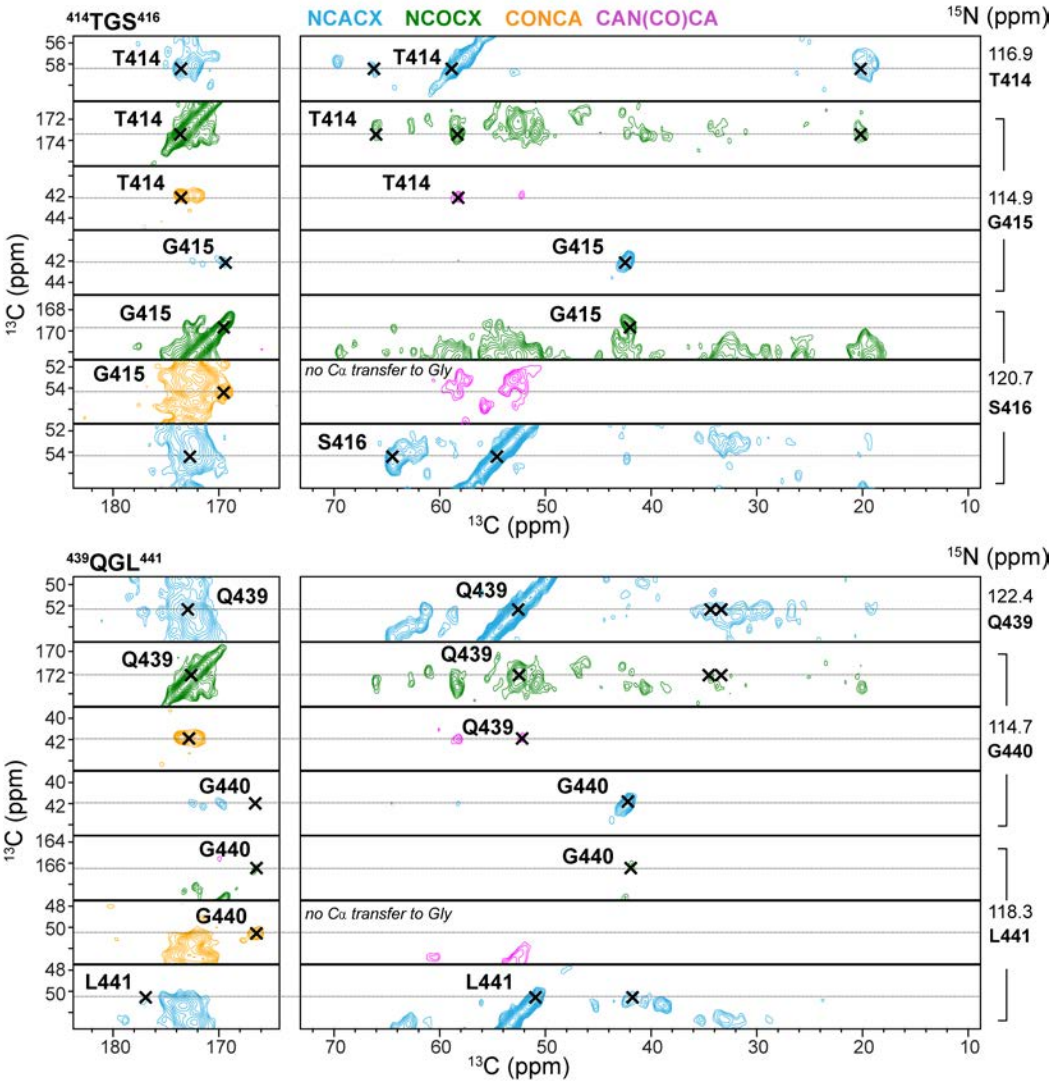


Figure S4. Additional strips of 3D correlation spectra of heparin-fibrillized 0N3R tau. (**a**) Assignment of ^{414}TGS^{416}. (**b**) Assignment of ^{439}QGL^{441}.

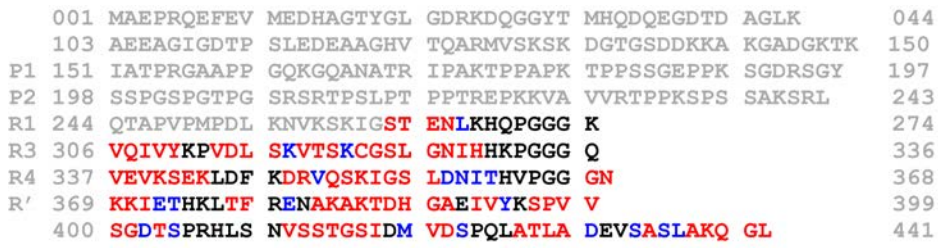


Figure S5. Summary of assignment status of 0N3R tau. Excluding residues outside the trypsin-resistant core (grey), for the trypsin-resistant core (262-441), 84 residues are confidently assigned (red), 20 are tentatively assigned (blue), and 45 are unassigned (black) due to low labeling level, dynamics, or sequence repetitiveness (e.g. PGGG).

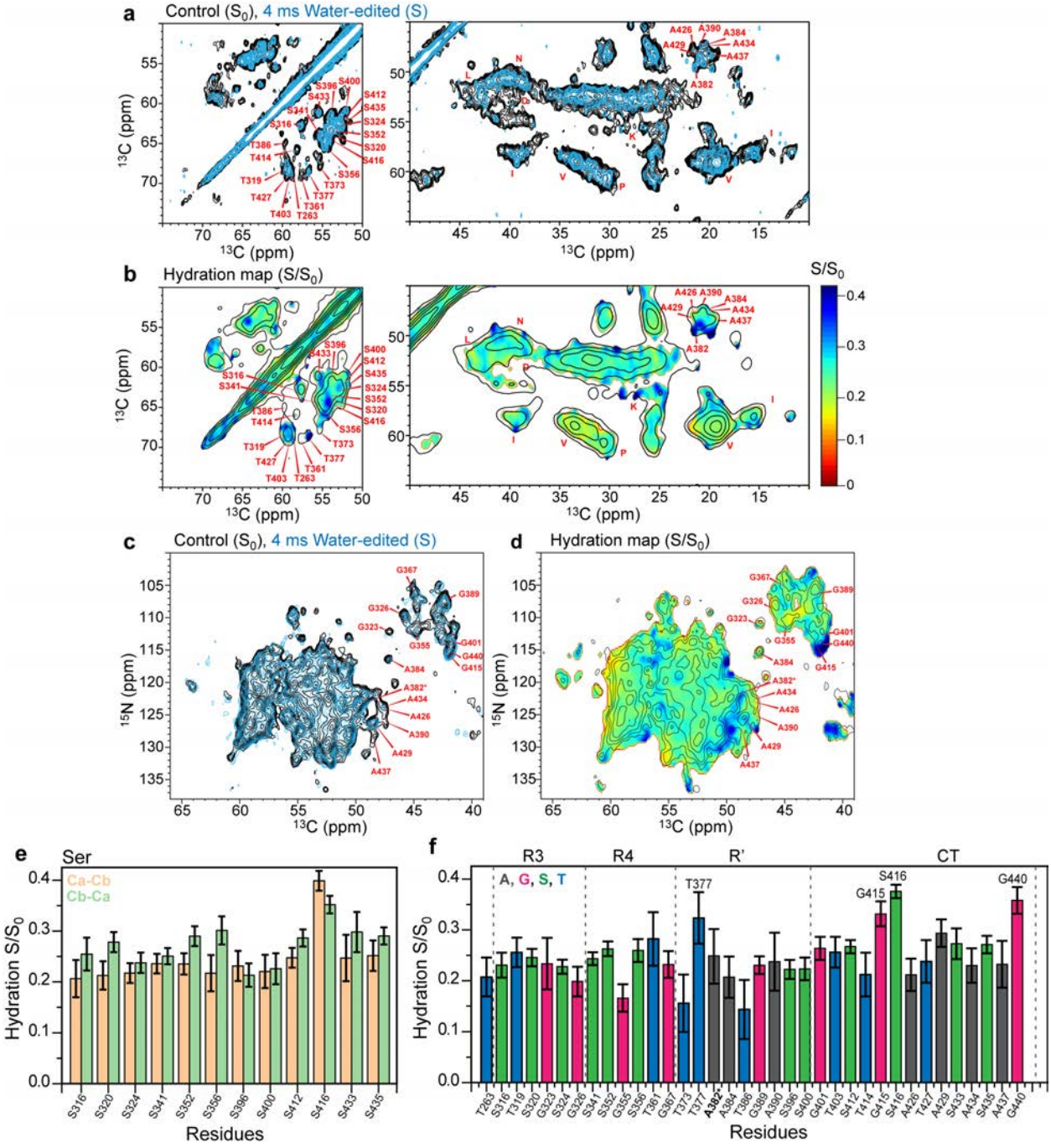


Figure S6. Water accessibility of the β-sheet core of 0N3R tau fibrils from water-edited 2D spectra, measured at a sample temperature of 268 K. (**a**) 2D water-edited ^{13}C-^{13}C correlation spectra measured with 23 ms CORD mixing. The control spectrum (black) was measured without the ^{1}H T_2 filter while the water-edited spectrum (blue) was measured with 4 ms ^{1}H mixing and a 190 μs water-selective ^{1}H T_2 filter. To report S/S_0 intensity ratios in terms of the 4 ms versus the 100 ms mixing spectra, we scaled the intensities of the unedited 2D spectra by the relative intensity between the unedited 1D spectra and the equilibrated 1D 100 ms water-edited spectra. Selected assignments from the 3D spectra are indicated. (**b**) 2D CC hydration map, obtained by dividing the intensities of the water-edited spectrum by the control spectrum in a pointwise fashion.

Additional line broadening and noise filtering were applied before pointwise S/S_0 division [7]. (**c**) 2D water-edited NCA spectrum (blue) measured with 4 ms ^1H mixing, overlaid with control spectrum measured without water-editing (black). (**d**) 2D NCA hydration map, obtained by dividing the intensities of the 4 ms water-edited NCA spectrum by the intensities of the control spectrum. (**e**) S/S_0 hydration values of Ser Cα-Cβ and Cβ-Cα signals, showing residue-specific differences in hydration. (**f**) S/S_0 hydration values of resolved residues in the rigid core. Ser and Thr values are obtained from the 2D CC spectra where their signals are better resolved, whereas the Ala and Gly values are obtained from the 2D NCA spectra where their signals are better resolved. The A382 signal is overlapped in the NCA spectrum, thus we used its better resolved Cα-Cβ and Cβ-Cα peaks in the 2D CC spectra for the hydration analysis. 2D CC derived intensities are the average of the Cα-Cβ and Cβ-Cα S/S_0 values to reduce the uncertainty. Residues S416, G415 and G440 are the best hydrated in the β-sheet core, indicating that these residues are either exposed to the surface of the β-sheet core or are located at a water pocket.

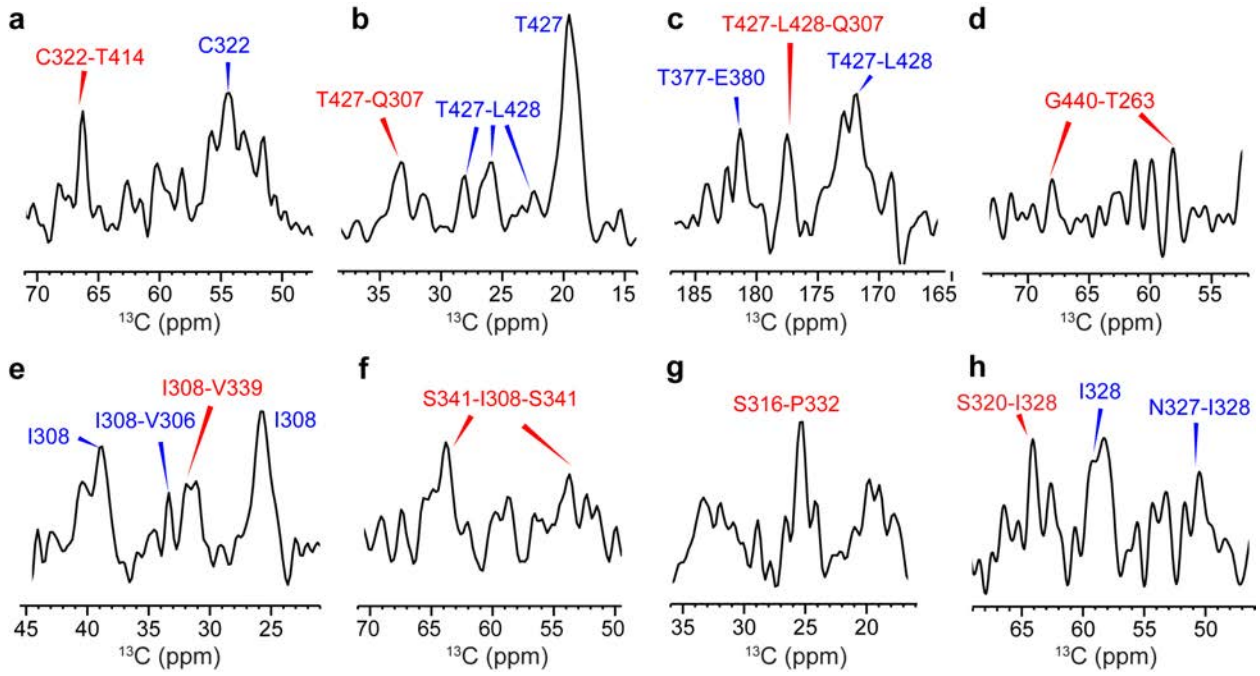


Figure S7. 1D cross-sections of the weak long-range cross-peaks shown in **Fig. 6**. Cross-sections are taken from the blue dashed lines in the parent 2D and 3D spectra.

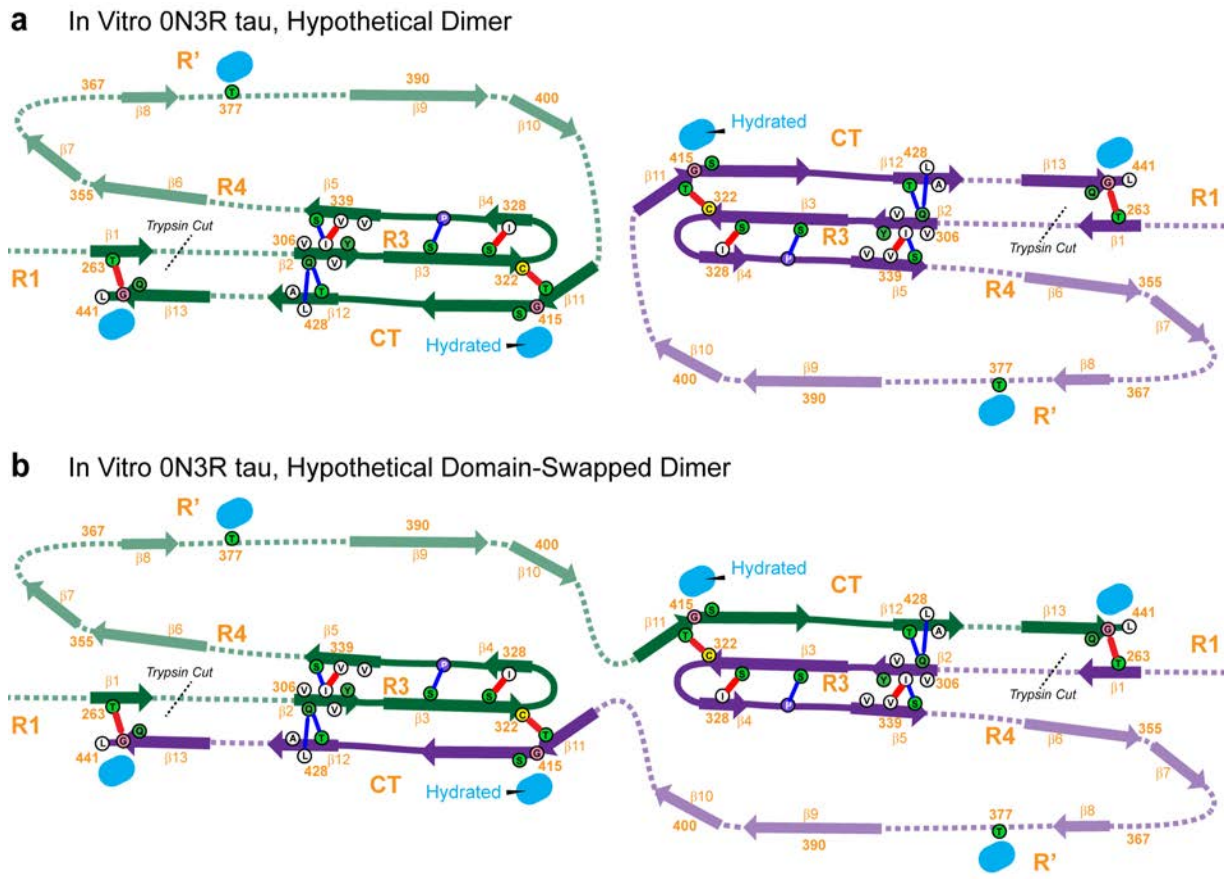


Figure S8. Hypothetical dimer models of heparin-fibrillized 0N3R tau fibrils that would agree with the observed CT-R3 contacts in the NMR spectra. (a) Dimer model where the CT-R3 interaction is intramolecular. (b) A domain-swapped dimer model where the CT-R3 interaction is intermolecular. We present no data that indicates the number of protofilaments in our fibrils, or the relative arrangement of the potential multimeric assembly. In addition, we cannot determine whether the long-range contacts we observed are intra- or inter-molecular. Other potential dimer or higher oligomer arrangement may also be consistent with either intra-molecular or domain-swapped R3-CT interactions, based on current data.

Table S1. Solid-state NMR experiments and their parameters used in this study. All experiments were conducted on an 800 MHz (18.8 T) spectrometer using a BlackFox 3.2 mm E-Free HCN probe. Reported temperatures are estimated sample temperatures based on external calibration of probe thermocouple and water chemical shift.

Experiment	**Experimental Parameters**	**Experimental Time (hrs)**
1D ^{13}C CP	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 1024, τ_{rd} = 1.7 s, τ_{dwel} = 5 µs; τ_{acq} = 15.3 ms; τ_{HC} = 70 µs; ν_{1Hacq} = 83 kHz	0.5
1D ^{13}C INEPT	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 1024, τ_{rd} = 2.0 s, τ_{dwell}= 5 µs; τ_{acq} = 24.6 ms; ν_{1Hacq} = 71 kHz	0.6
1D ^{15}N CP	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 1024, τ_{rd} = 2.0 s, τ_{dwell} = 10 µs; τ_{acq} = 20 ms; τ_{HN} = 1.0 ms; ν_{1Hacq} = 83 kHz	0.6
1D ^{15}N INEPT	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 1024, τ_{rd} = 2.0 s, τ_{dwell} = 10 µs; τ_{acq} = 25.6 ms; ν_{1Hacq} = 71 kHz	0.6
2D HN INEPT	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 64, τ_{rd} = 1.5 s, $t_{1,max}$ = 21.8 ms; $t_{1,inc}$ = 125 µs; τ_{dwell} = 20 µs; τ_{acq} = 81.9 ms; ν_{1Hacq} = 50 kHz	10
2D CC TOCSY	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 32, τ_{rd} = 1.7 s, $t_{1,max}$ = 7.5 ms; $t_{1,inc}$ = 24.9 µs; τ_{dwell} = 6 µs; τ_{acq} = 24.5 ms; τ_{TOCSY} = 12 ms; $\nu_{13C\ TOCSY,\ Waltz16}$ = 36 kHz at 50 ppm; $\nu_{1H\ TOCSY}$ = 0 kHz; ν_{1Hacq} = 50 kHz	9
2D CC: 23 ms CORD	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 56, τ_{rd} = 1.7 s, $t_{1,max}$ = 7.5 ms; $t_{1,inc}$ = 24.8 µs; τ_{dwell} = 6 µs; τ_{acq} = 15.3 ms; τ_{HC} = 70 µs; τ_{CORD} = 23 ms; ν_{1Hacq} = 83 kHz	16
2D CC: 200 ms CORD	T = 275 K; ν_{MAS} = 10.5 kHz, ns = 156, τ_{rd} = 1.5 s, $t_{1,max}$ = 7.5 ms; $t_{1,inc}$ = 24.8 µs; τ_{dwell} = 6 µs; τ_{acq} = 15.3 ms; τ_{HC} = 70 µs; τ_{CORD} = 200 ms; ν_{1Hacq} = 83 kHz	50
2D CC: 450 ms CORD	T = 293 K; ν_{MAS} = 14 kHz, ns = 208, τ_{rd} = 1.7 s, $t_{1,max}$ = 5.0 ms; $t_{1,inc}$ = 24.8 µs; τ_{dwell} = 6 µs; τ_{acq} = 18.4 ms; τ_{HC} = 70 µs; τ_{CORD} = 450 ms; ν_{1Hacq} = 71 kHz	51
2D water-edited CC: 23 ms CORD	T = 268 K; ν_{MAS} = 10.5 kHz, ns = 480, τ_{rd} = 1.7 s, $t_{1,max}$ = 6.2 ms; $t_{1,inc}$ = 24.8 µs; τ_{dwell} = 6 µs; τ_{acq} = 15.3 ms; $\tau_{1Hselpulse}$ = 952 µs; $\tau_{T2filter}$ = 95 µs x2; τ_{SD} = 4 ms; τ_{HC} = 70 µs; τ_{CORD} = 23 ms; ν_{1Hacq} = 83 kHz	117
2D CC: 23 ms CORD, Low temp	T = 268 K; ν_{MAS} = 10.5 kHz, ns = 80, τ_{rd} = 1.7 s, $t_{1,max}$ = 6.2 ms; $t_{1,inc}$ = 24.8 µs; τ_{dwell} = 6 µs; τ_{acq} = 15.3 ms; τ_{HC} = 70 µs; τ_{CORD} = 23 ms; ν_{1Hacq} = 83 kHz	19
2D NCA	T = 288 K; ν_{MAS} = 10.5 kHz, ns = 128, τ_{rd} = 1.7 s, $t_{1,max}$ = 14.2 ms; $t_{1,inc}$ = 95.2 µs; τ_{dwell} = 6 µs; τ_{acq} = 15.4 ms; τ_{HN} = 1 ms; τ_{NC} = 4 ms; $\nu_{1HspecificCP}$ = 83 kHz; ν_{1Hacq} = 83 kHz	19
2D water-edited NCA	T = 268 K; ν_{MAS} = 10.5 kHz, ns = 1824, τ_{rd} = 1.7 s, $t_{1,max}$ = 7.1 ms; $t_{1,inc}$ = 95 µs; τ_{dwell} =5 µs; τ_{acq} = 15.4 ms; $\tau_{1Hselpulse}$ = 952 µs; $\tau_{T2filter}$ = 95 µs x2; τ_{SD} = 4 ms; τ_{HN} = 1 ms; τ_{NC} = 4.5 ms; $\nu_{1HspecificCP}$ = 83 kHz; ν_{1Hacq} = 83 kHz	131

2D NCA, Low temp	T = 268 K; ν_{MAS} = 10.5 kHz, ns = 128, τ_{rd} = 1.7 s, $t_{1,max}$ = 7.1 ms; $t_{1,inc}$ = 95 μs; τ_{dwell} =5 μs; τ_{acq} = 15.4 ms; τ_{HN} = 1 ms; τ_{NC} = 4.5 ms; $\nu_{1HspecificCP}$ = 83 kHz; ν_{1Hacq} = 83 kHz	9
3D NCACX: SPECCP-CORD	T = 290 K; ν_{MAS} = 14 kHz, ns = 72, τ_{rd} = 1.7 s, $t_{1,max}$ = 6.9 ms; $t_{1,inc}$ = 143 μs; $t_{2,max}$ = 4.2 ms; $t_{2,inc}$ = 143 μs; τ_{dwell} = 6 μs; τ_{acq} = 12.3 ms; τ_{HN} = 1 ms; τ_{NC} = 5 ms; $\nu_{1HspecificCP}$ = 83 kHz; τ_{CORD} = 80 ms; ν_{1Hacq} = 71 kHz	207
3D NCOCX: SPECCP-CORD	T = 290 K; ν_{MAS} = 14 kHz, ns = 96, τ_{rd} = 1.6 s, $t_{1,max}$ = 6.9 ms; $t_{1,inc}$ = 143 μs; $t_{2,max}$ = 4.9 ms; $t_{2,inc}$ = 214 μs; τ_{dwell} = 6 μs; τ_{acq} = 12.3 ms; τ_{HN} = 1 ms; τ_{NC} = 5 ms; $\nu_{1HspecificCP}$ = 90 kHz; τ_{CORD} = 80 ms; ν_{1Hacq} = 71 kHz	211
3D CONCA: SPECCP-SPECCP	T = 290 K; ν_{MAS} = 14 kHz, ns = 96, τ_{rd} = 1.7 s, $t_{1,max}$ = 6.9 ms; $t_{1,inc}$ = 143 μs; $t_{2,max}$ = 4.9 ms; $t_{2,inc}$ = 214 μs; τ_{dwell} = 6 μs; τ_{acq} = 12.3 ms; τ_{HC} = 1.2 ms; τ_{CN} = 4 ms; τ_{NC} = 4 ms; $\nu_{1HspecificCP}$ = 83 kHz; ν_{1Hacq} = 71 kHz	203
3D CAN(co)CA: SPECCP-SPECCP-BSHCP	T = 292 K; ν_{MAS} = 14 kHz; ns = 128, τ_{rd} = 1.6 s, $t_{1,max}$ = 5.0 ms; $t_{1,inc}$ = 143 μs; $t_{2,max}$ = 7.5 ms; $t_{2,inc}$ = 214 μs; τ_{dwell} = 6 μs; τ_{acq} = 13.5 ms; τ_{HC} = 170 μs; τ_{CN} = 4 ms; τ_{NC} = 4 ms; τ_{CC} = 3.5 ms; $\nu_{1HspecificCP}$ = 83 kHz; $\nu_{1HBSHCP}$ = 83 kHz; ν_{1Hacq} = 71 kHz	287
3D long-range NCACX: SPECCP-450 ms CORD	T = 292 K; ν_{MAS} = 14 kHz, ns = 96, τ_{rd} = 1.7 s, $t_{1,max}$ = 8.6 ms; $t_{1,inc}$ = 214.2 μs; $t_{2,max}$ = 4.3 ms; $t_{2,inc}$ = 143 μs; τ_{dwell} = 6 μs; τ_{acq} = 12.3 ms; τ_{HN} = 1 ms; τ_{NC} = 4 ms; $\nu_{1HspecificCP}$ = 83 kHz; τ_{CORD} = 450 ms; ν_{1Hacq} = 83 kHz	281
3D CCC: 50 ms CORD-400 ms CORD	T = 292 K; ν_{MAS} = 14 kHz, ns = 32, τ_{rd} = 1.7 s, $t_{1,max}$ = 3.4 ms; $t_{1,inc}$ = 57.2 μs; $t_{2,max}$ = 3.4 ms; $t_{2,inc}$ = 57.2 μs; τ_{dwell} = 6 μs; τ_{acq} = 18.4 ms; τ_{HC} = 170 μs; τ_{CORD1} = 50 ms; τ_{CORD2} = 400 ms; ν_{1Hacq} = 83 kHz	278

Table S2. Amino acid (AA) composition and distribution in the primary sequence of 0N3R tau.

Amino acid	**Fuzzy Coat (1-261)**		**Rigid Core (262-441)**					Total 262-441	Total 1-441
	1-243	R1 244-261	R1 262-274	R3 306-336	R4 337-368	R' 369-399	CT 400-441		
Ala	19	1	0	0	0	3	4	7	27
Cys	0	0	0	1	0	0	0	1	1
Asp	13	1	0	1	3	1	4	9	23
Glu	11	0	1	0	2	3	1	7	18
Phe	1	0	0	0	1	1	0	2	3
Gly	23	1	3	5	4	1	3	16	40
His	3	0	1	2	1	2	1	7	10
Ile	3	1	0	2	2	2	1	7	11
Lys	17	3	2	4	4	6	1	17	37
Leu	5	1	1	2	2	1	5	11	17
Met	4	1	0	0	0	0	1	1	6
Asn	1	1	1	1	2	1	1	6	8
Pro	24	3	1	2	1	1	2	7	34
Gln	7	1	1	2	1	0	2	6	14
Arg	11	0	0	0	1	1	1	3	14
Ser	18	1	1	3	3	1	9	17	36
Thr	16	1	1	1	1	3	3	9	26
Val	6	2	0	4	4	3	3	14	22
Trp	0	0	0	0	0	0	0	0	0
Tyr	3	0	0	1	0	1	0	2	5
Total	185	18	13	31	32	31	42	149	352

Table S3. Assigned ^{13}C and ^{15}N chemical shifts of residues in the rigid core of heparin-fibrillized 0N3R tau. Chemical shifts are reported in ppm and are referenced externally to TMS for ^{13}C and to liquid ammonia for ^{15}N. Tentative assignments are marked with an asterisk (*). TALOS-N predicted torsion angles are given.[8] Italicized residues are unassigned but whose torsion angles are still predicted based on their neighboring residues. Chemical shifts have been deposited in the Biological Magnetic Resonance Bank (BMRB) with ID number 50785.

Residue	N	CO	Cα	Cβ	Cγ/γ1	Cγ2	Cδ/δ1	Cδ2/ε	Nγ2/δ2	ϕ (°)	ψ(°)
S262	111.0	172.3	55.2	64.9							
T263	121.7	171.2	58.7	67.9	19.6					-104±14	137±9
E264	120.2	172.1	53.0	31.7	32.3		177.8			-132±9	152±9
N265	124.2	172.8	50.3	40.9	170.9					-122±16	150±13
L266*	116.7		52.1	38.4	26.1					-74±10	142±13
V306	125.7	172.5	59.1	33.4	19.3					-113±11	131±5
Q307	129.4	172.3	52.6	33.4	32.0		177.6		111.9	-129±12	146±11
I308	126.3	171.9	58.0	39.0	25.7	15.3	11.8			-112±8	128±7
V309	128.4	170.9	59.7	32.0	19.2					-108±14	127±8
Y310	132.2	171.1	55.1		130.6					-99±19	137±23
V313	118.5	172.2	59.4	31.8	19.2					-112±15	130±11
D314	128.7	171.7	50.7	40.0	177.7					-97±13	124±12
L315	130.0	174.8	52.2	41.7	26.2	24.7				-95±16	123±12
S316	120.4	172.3	57.8	62.7						-88±13	-41±12
K317*			53.1							-135±26	152±16
V318	119.2	173.0	59.0	34.1	19.9					-127±7	134±12
T319	127.1	172.3	59.4	68.5	20.0					-105±8	128±5
S320	122.4		54.4	63.8						-122±15	159±13
K321*		172.5								-125±26	153±15
C322	113.8	173.2	54.8	30.0						-147±12	162±9
G323	112.5	171.6	47.2							-167±31	-173±21
S324	117.2	171.1	53.7	63.0						-129±23	151±12
L325	126.6	173.9	52.9	42.8	26.1					-120±24	141±15
G326	110.0	170.1	45.9							-172±22	-174±14
N327	113.3	173.8	50.7	40.1	174.6				110.1	-126±22	149±15
I328	124.7	172.8	59.4	39.8	25.9		13.8			-119±17	134±12
H329	123.4	171.7	53.5	29.1						-95±32	147±18
V337	125.8	172.4	59.2	31.5	19.4	17.5				-101±14	127±7
E338	129.2	172.4	52.7	31.5	34.4		180.8			-115±14	139±12
V339	126.9	171.0	59.5	31.6	21.7	18.4				-109±9	127±6
K340	130.0	174.2	52.6	33.4	24.9		27.3	42.3		-114±11	131±11
S341	120.9	171.3	55.3	64.2						-110±22	136±13
E342	127.0	173.4	52.3	30.5	32.4		178.3			-124±15	140±12
K343	126.0	173.8	56.0	30.3	21.1		25.8	42.7		-67±5	137±8
D348	123.6	171.0	51.2	42.4	178.4					-135±13	160±11

R349	117.4	172.2	54.3	32.3	28.8		42.7		85.3	-136±16	149±10
V350*	122.8	172.4	58.6	32.8	19.5					-117±16	138±11
Q351	122.2	172.6	52.6	30.3	29.3		177.1		111.4	-129±13	148±11
S352	116.3	172.1	54.4	63.5						-136±20	150±12
K353	121.3		53.3	33.2	22.0					-115±17	136±12
I354	125.7	175.4	58.3	38.5	26.2	16.8	11.9			-111±12	135±14
G355	112.7	169.6	45.0							-173±18	-170±21
S356	109.3	170.7	54.7	65.4						-142±13	162±12
L357	124.5	173.5	53.6	40.2	26.6		24.4			-71±9	137±14
D358*	118.5	171.8	50.9	42.2	178.2					-142±13	156±11
N359*	123.3	171.9	49.5	39.9						-119±19	139±14
I360*	122.4		58.0	39.4	24.8	14.9	11.0			-117±17	139±13
T361*	116.3	172.0	58.0	69.5	18.8					-127±17	151±15
G367	105.6	172.8	45.2							64±35	-142±32
N368	118.5	173.8	51.4	38.7						-96±13	-7±8
K369	117.3	171.7	56.2							-71±16	146±9
K370	123.9		52.5	34.1	21.8		26.9	43.6		-125±13	148±9
I371	123.5	171.2	58.2	40.5	25.5	16.3	12.1			-133±12	143±12
E372*		175.0	52.4							-103±20	136±12
T373*	110.0	169.3	55.4	68.1	20.4						
T377	118.2	170.4	56.9	68.8	20.0					-110±11	132±10
F378	127.9	173.4	54.8	39.8	128.7					-106±9	134±8
E380*			52.5	31.0							
A382	123.5	172.3	49.6	21.4						-143±14	154±13
K383	121.6	169.9	57.2							-71±10	-35±13
A384	116.9	174.6	47.4	20.4						-138±9	156±9
K385	122.0	173.3	55.8	33.6						-92±25	135±12
T386	127.3	170.4	60.0	65.2						-88±15	131±10
D387	128.7	172.0	51.1	41.3	177.5		177.8			-108±14	134±13
H388	123.2	172.1	54.4	30.7						-96±32	150±12
G389	107.6	170.8	42.6							-126±41	166±18
A390	126.4	173.4	47.6	20.8						-136±11	151±14
E391										-92±19	132±11
I392	123.0	171.6	58.1	38.8	24.7	16.7	11.9			-120±11	135±9
V393	127.3	171.8	59.5	32.9						-111±12	134±9
Y394*	121.9		56.2	42.1						-134±10	149±10
K395										-85±18	135±13
S396	122.1	171.7	54.0	61.6						-82±28	146±19
P397	136.1	172.9	60.9	30.5	25.7		48.3			-68±7	150±10
V398	119.9	172.7	59.1	33.4	19.4					-118±18	138±14
V399	124.3	173.6	58.6	34.5	18.9	17.9				-121±11	132±9
S400	122.9	171.8	52.4	61.2						-100±12	129±10

G401	114.0	172.2	41.9							-166±52	170±31
D402*	126.6	172.4	51.4	42.8	179.1					-119±17	145±14
T403		170.7	59.1	69.1						-111±9	129±6
S404*	125.2		54.4	62.1						-88±14	135±15
V411	130.6	172.3	60.4	32.0	20.1					-106±12	132±11
S412	117.4	171.9	53.0	62.5						-121±16	148±13
S413	121.0	172.5	55.3	64.2						-115±16	140±12
T414	116.9	173.4	58.4	66.1	20.1					-103±18	124±31
G415	114.9	169.7	42.1							-164±67	166±26
S416	120.7	172.7	54.4	64.3						-113±19	140±13
I417	124.1	172.4	59.2	39.0	25.9	17.1	12.9			-118±12	130±10
D418										-90±14	125±11
M419*	123.8	172.4	52.2	32.1	29.4					-106±13	132±11
V420	121.8	172.2	58.0	33.2	18.7					-124±10	131±13
D421	128.6	173.4	50.9	40.7	179.3					-93±15	135±16
S422*	114.3	168.9	52.5	62.7						-135±16	150±20
A426	124.0	173.9	47.8	21.6						-131±12	146±14
T427	119.3	169.2	59.7	69.6	19.7					-118±12	130±9
L428	132.3	172.5	52.5	43.1	27.6		24.9	22.7		-107±11	128±7
A429	126.5	173.6	48.6	21.2						-136±11	151±14
D430*	119.0	172.8	51.3	39.4	175.1					-90±31	144±14
S433*	120.2	173.1	55.8	61.1						-92±12	-12±16
A434	124.3	172.8	47.9	20.4						-133±16	146±11
S435*	121.0		52.2	62.6						-115±16	144±14
L436*	124.2	170.9	53.3	38.8	26.0		23.5			-79±13	134±8
A437	128.5	174.0	48.5	19.9						-121±14	142±16
K438	125.6	171.0	55.9	31.7	20.1		26.6	43.2		-77±19	136±10
Q439	122.4	172.2	52.3	34.4	33.3		177.6			-138±10	149±9
G440	114.7	166.5	41.9							-144±26	169±20
L441	118.3	176.8	50.5	41.7							

Table S4. Summary of assignment status and disambiguation strategies for the 0N3R tau fibril core.

Residue	Assignment status	Disambiguation method
S262	confident	not required
T263	confident	not required; type exclusion of T263
E264	confident	additional backbone walks
N265	confident	not required
L266	tentative	n/a
V306	confident	Inter-residue cross-peaks, validated by intensity
Q307	confident	Inter-residue cross-peaks, validated by intensity
I308	confident	Inter-residue cross-peaks, validated by intensity
V309	confident	Inter-residue cross-peaks, validated by intensity
Y310	confident	n/a
V313	confident	type exclusion of V313
D314	confident	not required
L315	confident	Inter-residue cross-peaks
S316	confident	not required; Inter-residue cross-peaks
K317	tentative	n/a
V318	confident	type exclusion of V318, additional backbone walks
T319	confident	type exclusion of T319, Inter-residue cross-peaks
S320	confident	Inter-residue cross-peaks
K321	tentative	n/a
C322	confident	not required; Inter-residue cross-peaks
G323	confident	not required; Inter-residue cross-peaks
S324	confident	additional backbone walks
L325	confident	not required; additional backbone walks
G326	confident	not required; Inter-residue cross-peaks
N327	confident	not required; Inter-residue cross-peaks
I328	confident	additional backbone walks
H329	confident	n/a
V337	confident	Inter-residue cross-peaks, validated by intensity
E338	confident	Inter-residue cross-peaks, validated by intensity
V339	confident	Inter-residue cross-peaks, validated by intensity
K340	confident	additional backbone walks
S341	confident	not required; Inter-residue cross-peaks
E342	confident	not required; Inter-residue cross-peaks
K343	confident	n/a
D348	confident	not required; type exclusion of D348
R349	confident	additional backbone walks
V350*	tentative	Inter-residue cross-peaks
Q351	confident	additional backbone walks
S352	confident	additional backbone walks
K353	confident	Inter-residue cross-peaks
I354	confident	not required; Inter-residue cross-peaks
G355	confident	not required; Inter-residue cross-peaks
S356	confident	Inter-residue cross-peaks
L357	confident	additional backbone walks
D358	tentative	n/a
N359	tentative	Inter-residue cross-peaks

I360	tentative	type exclusion of T361
T361	tentative	n/a
G367	confident	not required
N368	confident	not required
K369	confident	Inter-residue cross-peaks; additional backbone walks
K370	confident	n/a
I371	confident	not required
E372	tentative	type exclusion of T373
T373	tentative	n/a
T377	confident	not required
F378	confident	n/a
E380	tentative	Inter-residue cross-peaks
A382	confident	type exclusion of A382
K383	confident	not required
A384	confident	not required; Inter-residue cross-peaks
K385	confident	not required
T386	confident	type exclusion of T386
D387	confident	type exclusion of D387
H388	confident	not required
G389	confident	type exclusion of A390
A390	confident	n/a
I392	confident	Inter-residue cross-peaks
V393	confident	n/a
Y394	tentative	n/a
S396	confident	not required; Inter-residue cross-peaks
P397	confident	n/a
V398	confident	type exclusion of V398, additional backbone walks
V399	confident	not required; Inter-residue cross-peaks
S400	confident	not required
G401	confident	not required; Inter-residue cross-peaks
D402	tentative	type exclusion of D402
T403	confident	n/a
S404	tentative	additional backbone walks
V411	confident	Inter-residue cross-peaks
S412	confident	Inter-residue cross-peaks
S413	confident	Inter-residue cross-peaks
T414	confident	not required; Inter-residue cross-peaks
G415	confident	not required; Inter-residue cross-peaks
S416	confident	not required; Inter-residue cross-peaks
I417	confident	n/a
M419	tentative	n/a
V420	confident	Inter-residue cross-peaks
D421	confident	n/a
S422	tentative	n/a
A426	confident	not required; Inter-residue cross-peaks
T427	confident	not required; Inter-residue cross-peaks
L428	confident	additional backbone walks
A429	confident	type exclusion of A429, additional backbone walks
D430	tentative	n/a
S433	tentative	n/a

A434	confident	Inter-residue cross-peaks; type exclusion of A434
S435	tentative	n/a
L436	tentative	n/a
A437	confident	type exclusion of A437, Inter-residue cross-peaks
K438	confident	additional backbone walks
Q439	confident	not required
G440	confident	not required; Inter-residue cross-peaks
L441	confident	n/a

Table S5a. Long-range (*i-j* > 4) cross-peaks observed in the 2D and 3D ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra. Ambiguously assigned cross-peaks are marked with an asterisk.

Long-Range Cross-Peaks	Spectrum
C322Cα – T414Cβ	3D CCC CORD 400 ms
T263Cβ – G440CO	2D CC CORD 200 ms
T263Cα – G440CO	2D CC CORD 200 ms
S320Cβ – I328Cδ1	2D CC CORD 200 ms
I308Cγ2 – V339Cβ	3D CCC CORD 400 ms
T263Cβ – G440Cα*	3D CCC CORD 400 ms
Q307Cδ – L428Cα*	3D CCC CORD 400 ms
T427Cα – Q307Cβ/V306Cβ*	3D CCC CORD 400 ms
I308Cβ – S341Cβ*	3D CCC CORD 400 ms
S316Cβ – P332Cγ*	3D CCC CORD 400 ms
C322Cβ – T414Cβ*	3D CCC CORD 400 ms

Table S5b. Medium-range ($5 > i\text{-}j > 1$) contacts observed in the 2D and 3D ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra.

Medium-Range Contacts		
3D NCACX 80 ms	**3D CCC CORD 400 ms**	**2D CC CORD 200 ms**
V306Cβ – I308Cα	V306Cα – I308Cα	S262Cβ – E264Cδ
S316Cα – S320Cα	V306Cα/Cβ – I308Cβ/Cγ1	V306Cβ – I308Cβ/Cγ1/Cδ1
S316Cα – S320Cβ	V306Cα – I308Cγ1	I308Cβ – Y310Cα
V339Cα – S341Cα	V306Cα – V309Cβ	D314Cβ – S316Cβ
I354Cα – S356Co	V306Cβ – I308Cδ1	S316Cβ – V318Cα/Cβ/Cγ
P397Cα – V399Cα	Q307Cδ – Y310Cβ	V318Cβ – S320Cβ
V399Cα – G401Cα	I308Cγ1 – Y310Cα	T319Cβ – C322Cβ
V411Cα – S413Cβ	D314Cβ – S316Cα	C322Cβ – S324Cβ
T414Cα – I417Cγ1	L315Cβ – T319Cβ	C322Cβ – G326Cα
3D NCACX 450 ms	S316Cα – S320Cα	G323Cα – L325Cβ
S262Cα – E264Cα	S316Cβ – V318Cα	G323Cα – N327Cβ
Q307Cδ – V309Cα	V318Cβ – S320Cα/Cβ	S324Cβ – G326Cα
S316Cα – V318Cα	C322Cα/Cβ – S324Cβ	D348Cβ – S352Cβ
V318Cα – S320Cα	C322Cβ – L325Cγ/Cδ1	S352Cα – I354Cδ1
V339Cα – S341Cα	G323Cα – L325Cγ	S352Cβ – G355Cα
V339Cα – E342Cα	S324Cβ – N327Cβ	S352Cβ – S356Cβ
V339Cγ – E342Cα	S324Cα – I328Cγ1	I354Cα – S356Cβ
K340Cε – E342Cα	E338Cα – S341Cα	I354Cγ – S356Cβ
S396Cα – V398Cβ	K340Cε – E342Cδ	S356Cβ – D358Cα
V398Cα – S400Cβ	D348Cα/Cβ/Cγ – V350Cγ	T377Cα/Cβ – E380Cδ
V399Co – G401Cα	V350Cβ – S352Cβ	K385Cα – D387Cγ
V411Cβ – T414Cα	I354Cβ – S356Cα	I392Cγ2 – Y394Cα
S412Cα – T414Cα	I354Cγ1/Cδ1 – S356Cα/Cβ	V398Cα/Cβ – S400Cβ
T414Cα – S416Cα	I354Co – S356Cα	V411Cα – T414Cβ
L436Cα – K438Cβ/Cγ/Cδ/Cε	S356Cα – D358Cα	S413Cα – I417Cγ2
	T377Cβ – E380Cα/Cβ/Cγ/Cδ	G415Cα – I417Cβ/Cδ1
	E380Cα/Cβ/Cδ – A382Cβ	T414Cβ – S416Cα
	V399Cα/Cβ – G401Cα	G415Cα – I417Cα
	V411Cβ – T414Cβ	I417Cγ2 – M419Cγ
	S412Cα/Cβ – T414Cα	V420Cβ – I417Cδ1
	S412Cβ – T414Cβ	A426Cα – L428Cβ
	G415Cα – I417Cβ	K438Cδ – G440Co
	A426Cα/Cβ – L428Cα	
	T427Cβ – A429Cα/Co	

References

1. Hu, K. N.; Qiang, W.; Tycko, R., A general Monte Carlo/simulated annealing algorithm for resonance assignment in NMR of uniformly labeled biopolymers. *J. Biomol. NMR* **2011,** *50*, 267-276.

2. Yang, Y.; Fritzsching, K. J.; Hong, M., Resonance Assignment of Disordered Proteins Using a Multi-Objective Non-Dominated Sorting Genetic Algorithm. *J. Biomol. NMR* **2013,** *57*, 281-296.

3. Fritzsching, K. J.; Yang, Y.; Schmidt-Rohr, K.; Hong, M., Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *J. Biomol. NMR* **2013,** *56*, 155-167.

4. Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I., . . . Eisenberg, D., Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007,** *447*, 453-457.

5. Dregni, A. J.; Mandala, V. S.; Wu, H.; Elkins, M. R.; Wang, H. K.; Hung, I., . . . Hong, M., In vitro 0N4R tau fibrils contain a monomorphic β-sheet core enclosed by dynamically heterogeneous fuzzy coat segments. *Proc. Natl. Acad. Sci. U.S.A.* **2019,** *116*, 16357-16366.

6. Hoffmann, J.; Ruta, J.; Shi, C. W.; Hendriks, K.; Chevelkov, V.; Franks, W. T., . . . Lange, A., Protein resonance assignment by BSH-CP-based 3D solid-state NMR experiments: A practical guide. *Magn. Reson. Chem.* **2020,** *58*, 445-465.

7. Dregni, A. J.; Duan, P.; Hong, M., Hydration and Dynamics of Full-Length Tau Amyloid Fibrils Investigated by Solid-State Nuclear Magnetic Resonance. *Biochemistry* **2020,** *59*, 2237-2248.

8. Shen, Y.; Bax, A., Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **2013,** *56*, 227-241.