UNIVERSITY OF CALIFORNIA

Santa Barbara

Probing the aggregation transition of the disordered protein Tau

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Michael P. Vigers

Committee in charge:

Professor Songi Han, Co-chair

Professor Michelle A. O'Malley, Co-chair

Professor M. Scott Shell

Professor Cyrus R. Safinya

June 2022

The dissertation of Michael P. Vigers is approved.

Cyrus R. Safinya

M. Scott Shell

Michelle A. O'Malley, Committee co-chair

Songi Han, Committee co-chair

June, 2022

ACKNOWLEDGMENTS

Presented here is a collection of research conducted during my time in the labs of Professors Songi Han and Michelle O'Malley. I must first thank Songi for all her mentorship, guidance, and patience. Not only did Songi provide strong guidance and ideas for all the work presented here. Songi has created a welcoming atmosphere for conducting research in her group. Some of the highlights from my time in Santa Barbara include group social activities, including a ski trip to Big Bear and many beach barbeques. I spent the first 8 months of my time at UCSB in Michelle's lab before I transitioned into working on Tau, primarily in Songi's lab. I am incredibly thankful for my time in Michelle's lab. I learned a lot about how to conduct research in an organized, clean, and efficient manner. I also learned about molecular biology and protein expression that gave me a great foundation for biophysical research.

As is the case with most scientific results, this entire work was the result of collaboration with many others who contributed their expertise, time, and creativity towards my work. Without them this work would not have been possible. Much of the work I conducted contributed to larger stories created by my collaborators. I present as faithfully as possible, the work that I was responsible for, but some background information and context are required. Portions of the manuscripts associated with each work are reprinted or amended to reflect this. There are many people I would like to thank but will thank a few particularly influential people here.

First, I would like to thank Yann Fichou for his close guidance early in my research career. Yann was the driving force behind the publications presented in Chapters 2 and 3. Yann provided me with training on EPR instrumentation and experimental methods and helped me develop as an independent researcher. I had the pleasure of sharing an office space with Yann while he finished a postdoctoral fellowship. We had discussions that ranged from global politics to the physics of coffee brewing, and our informal office discussions often drove new ideas forward more effectively than any organized meeting could.

Professor Ken Kosik has been a great influence on my work throughout my time at UCSB. Ken's ability to span scientific fields from clinical pathology to biophysics is admirable, and this breadth of knowledge was crucial to keep my very fundamental work grounded in the goal of making progress towards solving Tauopathies. I also would like to thank Ken for always providing personal feedback and making time for personal meetings with me although I was not one of his many students. We had multiple individual and small group meetings that were very insightful.

Judith Steen entrusted me to work with her proprietary information about fragments of Tau and provided a lot of great discussion with regards to how fragmentation could drive aggregation. She was exceedingly patient and helpful while I learned out how to produce the fragments she identified. One of the few positive outcomes of the Covid-19 pandemic was the development of more regular meetings with Judith and her lab members Pieter Beerepoot and Long Cheng.

Zhikai (Kate) Zeng was also my office mate, and neighbor during our PhDs. Kate joined the group at the same time as me, and I think we have both shared the highs and lows associated with research together. Another highlight of being home during Covid was dog playdates with her dog, Lufi and my own dog, Suki.

Susanna Seppälä is a staff scientist in Michelle O'Malley's lab to whom I owe immense gratitude. Any molecular biology ability I have can be attributed to Susanna's guidance. Additionally, I enjoyed our discussion of literature (scientific or otherwise), and her enthusiasm for protein science.

I had the pleasure to work with many brilliant scientists during my time at UCSB including Jenny Rauch, Andrew Longhini, and Mei Zhang from the Kosik lab, Nikki Eschmann (nèe Schonenbach), Neil Eschmann, Tom Casey, ChungTa Han, Yanxian Lin, Khanh (Kendrick) Nguyen, Yuanxin (Miranda) Li, Austin Dubose, Karen Tsay and Vishnu Vijayan from the Han lab, and Pieter Beerepoot from the Steen group. All of these people had a tangible influence on me personally. Andrew Goring, Tommy Ngo, Alicia Maddy, and Jayden Taylor were all very talented undergraduate students I had the pleasure to work with.

I'm grateful for the collaborations with each person I have worked with and wish them all success in their future careers.

ABSTRACT

Probing the aggregation transition of the disordered protein Tau

by

Michael P. Vigers

In the Tauopathy subfamily of neurodegenerative diseases, the intrinsically disordered protein (IDP), Tau, undergoes a transition from its native disordered state into insoluble, pathological amyloid fibrils, such as those found in the neurofibrillary tangles of Alzheimer's disease (AD). The mechanisms of Tau aggregation are still unclear, despite decades of research. This dissertation describes how Tau interacts with its surrounding environment and the cofactors that lead to its pathological aggregation. Chapter 1 provides an overview of disordered proteins, aggregation, Tau, neurodegeneration, and the biophysical techniques used throughout this manuscript. Chapter 2 describes the fibril structures formed by recombinant Tau when incubated with heparin, a polysulphated glycosaminoglycan, and how the structures are heterogeneous and dissimilar to the fibrils observed in AD. Chapter 3 investigates the nature of the interaction between Tau and the anionic cofactors routinely used to generate fibrils *in vitro*. The cofactors were found to be integral to the fibrils and stoichiometric quantities of cofactor were required to form the fibril. In Chapter 4, fragments of Tau were produced that were found to exist in pathological brain tissue. These proteolytic products of Tau that are observed in AD patients form potent, aggregation-prone fragments of Tau that may be involved in initiating aggregation of larger species of full-length Tau. In Chapter 5,

shorter peptides of Tau are used to demonstrate the effect of a hereditary mutation (P301L) observed in chromosome 17 linked frontotemporal dementia and Parkinsonism (FTDP-17). The primary effect of the mutation, in a short peptide of Tau (HP301L), is to create a slower-moving hydration shell with a lower entropy than in the WT form. The lower entropy around the hydrophobic leucine makes the dehydration required for amyloid formation more energetically favorable, leading to higher aggregation propensity. Chapter 6 describes ongoing work to determine structure of the fibrils formed by the P301L peptide in Chapter 5. A structure is presented showing a fibril core at 4.6 Å resolution with a C_2 symmetry composed of 2 hairpin-shaped protofilaments. Chapter 7 investigates the proline rich domain of Tau and its role in liquid liquid phase separation and identifies the SH3 domain as a binding partner that may be involved in the removal of Tau from microtubules. Chapter 8 suggests future directions for the field and outstanding questions from this body of work.

Vita of Michael P. Vigers

Michael P. Vigers
michael.vigers@gmail.com
701 Bolton Walk Apt. 201 Goleta, CA 93117 • 720-220-2968

EDUCATION

PhD Chemical Engineering, University of California, Santa Barbara, Santa Barbara, CA — June 2022
PhD focus: Biophysical characterization of the intrinsically disordered protein Tau

B.S. Biological Engineering, Montana State University, Bozeman, MT — May 2016
Highest Honors

B.S. Chemical Engineering, Montana State University, Bozeman, MT — May 2016
Highest Honors

EXPERIENCE

University of California, Santa Barbara, Santa Barbara, CA — January 2017-June 2022
Graduate Researcher - Labs of Songi Han & Michelle O'Malley

Montana State University, Bozeman, MT — May 2014-May 2016
Research Assistant - Wilking Lab and Heys Lab

PUBLICATIONS

- Vigers, M., *et. al.* Hydrophobic templating of key Tau segment is a defining event in pathological Tau aggregation and seeding. In Preparation.
- Beerepoot, P., Vigers, M., Cheng, L. (2021). Neuronal immunoproteasome induction links inflammatory signaling to pathological Tau aggregation in Alzheimer's Disease. In review.
- Zhang, X., Vigers, M. et. al. (2020). The Proline-rich Domain Promotes Tau Liquid Liquid Phase Separation in Cells. J Cell Biol (2020) 219 (11): e202006054. https://doi.org/10.1083/jcb.202006054
- Nguyen, K., Vigers, M. et. al. Oligomerization of the human adenosine A_{2a} receptor is driven by the intrinsically disordered c-terminus. eLife 2021;10:e66662. DOI: 10.7554/eLife.66662
- Fichou, Y., Vigers, M., Goring, A. K., Eschmann, N. A., & Han, S. (2018). Heparin-induced Tau filaments are structurally heterogeneous and differ from Alzheimer's disease filaments. Chem. Comm., 54(4573), 8-10. https://doi.org/10.1039/c8cc01355a
- Fichou, Y., Lin, Y., Rauch, J. N., Vigers, M., Zeng, Z., Srivastava, M., Keller, T. J., Freed, J. H., Kosik, K. S., & Han, S. (2018). Cofactors are essential constituents of stable and seeding-active Tau fibrils. PNAS, 115(52), 13234-13239. https://doi.org/10.1073/pnas.1810058115

TEACHING

Chemical Engineering Thermodynamics, UC Santa Barbara — March 2019-May 2019
Teaching Assistant

Chemical Engineering Lab Course, UC Santa Barbara — January 2018-March 2018
Teaching Assistant

Introduction to Computational Applications, Montana State University — January 2015-May 2015
Teaching Assistant

Awards/Accomplishments:

- 2018 Tau Consortium Fellow
- 2016 Biophysical Society Education Committee Travel Awardee
- 2016 NSF Graduate Research Fellowship Honorable Mention
- Passed FE Exam for Chemical Engineering April, 2016

TABLE OF CONTENTS

Preface

The work presented below is the results of multiple collaborations with colleagues from UC Santa Barbara and other institutions. Works that are published elsewhere are presented with editing for clarity, but include work conducted by collaborators to keep the continuity of each story. Each chapter includes a conclusion section that summarizes the impact of the work. If the work I conducted was distinct from the rest of the publication, such as Chapters 4 and 7, my own results are summarized in the context of this dissertation.

Chapter 1: Introduction

1.1 Disordered Proteins

Intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) of proteins are stretches of amino acids that lack a permanent secondary structure such as β-sheets and α-helices. Historically, IDPs and IDRs have been understudied compared to their prevalence in the proteome. IDPs' lack of permanent structure precludes the use of many classic structural biology techniques, such as crystallography, but developing biophysical techniques for probing IDP functions have produced more studies on IDPs[1]. It is clear that IDPs and IDRs are heavily involved in many important cellular functions[2]. While IDPs generally do not have enzymatic function[3], their primary role is to introduce inter-molecular interactions including signaling cascades, allostery[4], RNA binding, DNA-chromatin complexes, and liquid-liquid phase separated (LLPS) condensate formation[2,5]. IDPs can drive these binding events through the multiple 'soft' interactions; the large exposed surface area, and the many different conformations in the unstructured chain present a vast number of possible binding interfaces that are more promiscuous and less specific than globular proteins[6].

The proportion of the proteome that is composed of disordered regions increases with organism complexity[2], but these regions have evolutionary constraints, and evolve rapidly[3]. The high rate of evolution is possible because many IDP functions have low sequence specificity; instead relying on the entropically-driven binding of the disordered chain that is dependent on chain length, and sequence complexity[3,7]. In stark contrast to their sequence-independent function, IDPs can undergo a disorder-to-order transition, that leads to dysfunctional protein behavior and is strongly associated with disease[8,9]. One well-known

example is the formation of amyloid fibrils by IDPs that are observed in many neurodegenerative diseases[8,9]. Amyloid fibrils represent the lowest free-energy and most favorable conformations possible for many proteins[9–11]. Because IDPs have a flatter energy landscape than globular proteins[9,12] and there are no distinct local minima, small perturbations to the thermodynamic properties of the protein can have large impacts on its formation. Thus, small perturbations to the global energy landscape can bias IDPs towards fibrilization.

1.2 Mechanisms of amyloid aggregation

Many IDPs, including $A\beta_{40}$ and $A\beta_{42}$, α-synuclein, TDP-43, TMEM-106, Huntingtin (HTT) and Tau are known to form amyloid filaments that are associated with a broad range of neurodegenerative diseases[8,9,11,13,14]. These diseases, including Alzheimer's disease (AD), Huntington's disease, and Parkinson's disease cause slow degradation of cognitive and motor abilities for patients and are costly to manage and treat. 6.5 million people suffer from AD in the United States alone, and the cost of treatment is estimated at over $300 billion annually[15,16]. The insoluble amyloid filaments form neurofibrillary tangles (NFT), plaques, and other toxic inclusions that spread throughout the brain, causing dysregulation of cellular pathways and leading to neurodegeneration.

Amyloid aggregation is defined by the formation of long filaments that contain a cross-β sheet structure (Figure 1.1A). Amyloid proteins are packed in extended β-sheet conformations that lay perpendicular to the fibril's axis (Figure 1.1B)[9,17,18]. It is believed that amyloids are the most stable possible protein conformation, but the transition into this structure by globular proteins is less frequent due to a partial unfolding that is required to transition into the aggregation pathway[11]. The more energetically favorable aggregation of IDPs is observed in an enrichment of amyloid-associated proteins[8,12].

Amyloid fibrils are extremely stable due to the high bond density formed with in the fibrils. The stacked cross-β sheets form hydrogen bonds between the backbones of adjacent proteins (Figure 1.1A)[17,19] and the sidechains are oriented nearly orthogonal to the axis of the filament. Frequently, these side chains will form shielded folds, steric zippers, or interdigitate across strands (Figure 1.1B)[20–22]. Most amyloid fibrils (and all known Tauopathy fibrils) orient in a parallel manner, which aligns hydrophobic, hydrophilic and aromatic sidechains directly in register with the same sidechain. These homotypic interactions create strong hydrophobic, Van Der Waals, and π-π bonding forces to stabilize the fibrils. The filamentous nature of amyloids leaves an exposed face at each end of the fibril, which present an active site for polymerization[23]. This polymerization makes aggregation even more favorable once an oligomer has formed, leading to 'prion-like' behavior. Like Prions, if amyloid fibrils are introduced to naïve monomers further aggregation occurs, leading many to hypothesize that this prion-like effect, often referred to as 'seeding' drives the progress of neurodegeneration[24–27].

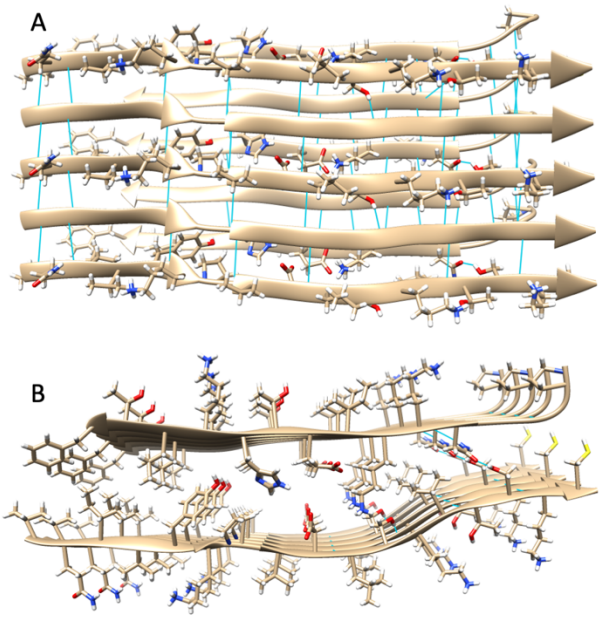


Figure 1.1: Residues V306-C322 and P364-F378 of the AD PHF fibril (PDB #5O3L). A) Side view of an amyloid fibril. Hydrogen bonds between strands are shown in teal. B) Top view of the fibril. Interdigitation of sidechains across β-sheets is shown between Phenylalanine, Isoleucine, and Leucine on lefthand side. Homotypic stacking of hydrophobic residues, and aromatic residues can be seen throughout.

1.3 Role of Tau in neurodegenerative disease

Tau is the protein produced by the *MAPT* gene of chromosome 17. There are 6 different isoforms translated due to alternative splicing[28]. The longest length isoform is 441 amino acids (referred to as hTau40 or 2N4R Tau). There are two N-terminal repeat domains that produce 0N, 1N, and 2N isoforms depending on splicing, followed by a proline-rich domain (PRD) and a microtubule binding domain (MTBD) with 4 repeats. Repeat 2 can be spliced out, forming so called 3R forms of Tau. The MTBD is followed by a positively charged C-terminal domain (Figure 1.2B).

Tau's functional role is to bind tubulin and stabilize the polymerized form in microtubules. Tau binds microtubules primarily along the MTBD and through the proline-rich domain[29]. This allows for the formation of very long microtubules found in the axons of

neurons, where Tau is found most abundantly[30]. When it becomes dysfunctional, Tau self-polymerizes into insoluble amyloid fibrils. Most famously, these insoluble fibrils cluster into the intracellular neurofibrillary tangles observed in Alzheimer's Disease (AD). Tau aggregation is observed in a whole host of other neurodegenerative diseases, referred to as Tauopathies.

There has been significant interest in the formation of Tau fibrils since the early discovery of NFTs in AD by Alois Alzheimer in 1907[31,32]. Later, electron microscopy was used to show that tangles are composed of filaments, including paired helical filaments (PHF) and straight filaments (SF) in AD[33–35]. Kosik and Selkoe later showed that Tau is the primary component of the PHF in the tangles[36,37]. Since these early studies, much work has been devoted to understanding how these fibrils are formed, their exact composition, and the molecular structure of the fibrils. Early work by Crowther and Wischik *et. al.* used negative stain EM (NS-EM) to show the PHF structure consisted of 2 C-shaped protofilaments[34]. The model generated by Wischik et. al. was later used to resolve the structure of PHF and SF to atomic (3.4 Å) resolution in a groundbreaking work by Fitzpatrick *et. al.* [38]. Since 2017, the structures of 12 Tauopathies have been solved with 8 unique structures found. All Tauopathies surveyed thus far have shown that a single fibril structure typifies the disease phenotype. There can be structural homologues, as is the case in AD, with straight filaments and paired helical filaments (PHF) that share a common protofilament structure but contain different interfaces between the protofilaments[39].

Clinically, Tauopathies are diagnosed with pathophysiological and behavioral tests, but correct diagnosis can be difficult broad range of behaviors associated with Tauopathies that overlap with other neurodegenerative diseases such as TDP-43 proteinopathies[40]. Though there

is ongoing work towards the development of diagnostics, such as PET tracers, the correct diagnosis of Tauopathy is only possible through an autopsy[40,41]. Tauopathies are primarily classified by the isoform (3R, 4R, or 3R/4R) of Tau observed in the fibrils, but recent atomistic structures of the fibrils make it clear that they can be classified more specifically into 4 major families (Figure 1.2). The structure in Pick's disease (PiD), a 3R Tauopathy, is different to any 4R Tauopathy structure as it lacks the R2 domain. AD and Chronic Traumatic Encephalopathy (CTE) (3R/4R) display a shared C shape, but have different internal structures, with an additional non-proteinaceous density in the CTE fibril. Corticobasal degeneration (CBD) (4R) displays a larger core, with multiple internal folds, and a non-proteinaceous density.

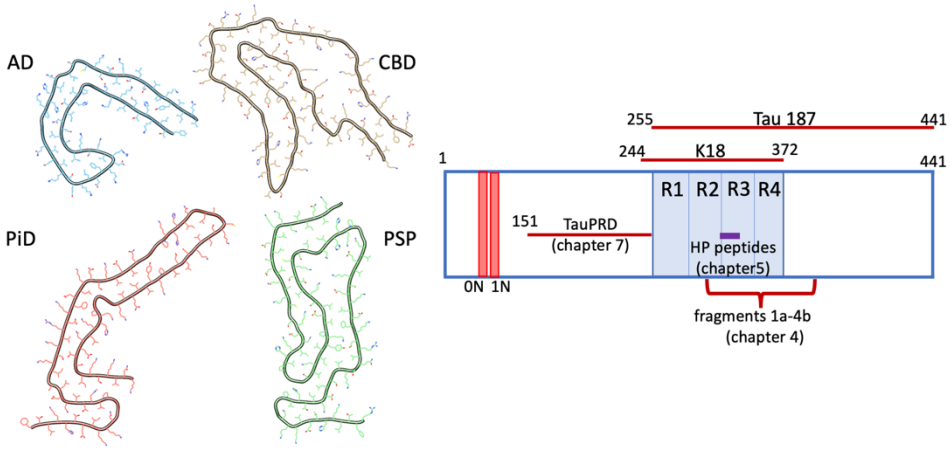


Figure 1.2: (clockwise from top left) The immobilized core of fibrils formed by Tau in Alzheimer's disease (AD), corticobasal degeneration (CBD), prefrontal supranuclear palsy, and Pick's disease (PiD). These structures represent the 4 major structural classes of Tau fibrils observed in Tauopathies. (right) Schematic of Tau's primary sequence. Highlighted are segments of Tau used and referred to in this manuscript.

Recently, structures of multiple rarer Tauopathies have been published. Some share structures with the larger Tauopathies, indicating they may have commonalities in the pathology[42]. It remains to be seen if the structural similarities can help link these diseases that were thought to be independent. For example, Familial Danish Dementia (FDD) has pathology

distinct from AD but shares an identical structure. It remains to be seen, but these structural studies may suggest shared treatment methods between previously distinct cases.

Despite the structures of Tau in most Tauopathies being resolved, there are multiple questions remaining regarding Tau pathology that still require biophysical characterization of Tau. Spreading of Tau tangles from cell to cell is of paramount importance. Tau pathology follows a distinct migration across the neurons of the brain that is Tauopathy dependent, but is characteristic of each Tauopathy[42]. When Tau spreads from neuron to neuron, what is the transport pathway? Binding partners of Tau need to be identified that may be carriers for Tau. Many hypothesize that microglia uptake Tau fibrils, but it is also suggested that cell-surface receptors may bind to Tau, and internalize toxic fibrils, or oligomers of Tau[43]. It is also unclear if there are specific forms of Tau that are more prone to aggregation *in vivo* than others. Some hypothesize that proteolytic cleavage of Tau may form aggregation-prone conformations (see Chapter 4 for more discussion), whereas others posit that cleaved forms of Tau found postmortem are the result of stalled proteolysis, wherein the most stable core of the fibril is unable to be digested, leaving only the core residues in the cell[34,44].

Finally, Tau's interactions with non-proteinaceous cofactors require more study. As will be detailed later (Chapters 1 & 2), multiple cofactors have been identified *in vitro* to enhance or promote aggregation. Cryo-EM structures have shown that in some Tauopathies (CTE and CBD most notably), cofactors are integral to the fibrils, but most of the previously identified cofactors do not match up with the predicted size and properties of the densities observed in the cryo-EM structures[45,46]. With the complexity of the cellular environment, the homogeneity of the Tau fibrils observed, and the disordered nature of native Tau, it is easy to envision a cofactor acting as a scaffold along which Tau can align and polymerize.

Alternatively, one could argue that many oligomers, or protofibrils of Tau form, and only the most stable ones survive, that then seed the aggregation of identical fibrils.

1.4 *In Vitro* Tau aggregation

Work by Mandelkow and coworkers showed that fibrils of similar morphology to PHF and SF could be formed with the addition of anionic cofactors to recombinant Tau purified from bacterial cultures[47]. This discovery prompted a significant increase in biophysical work to describe Tau aggregation. Cofactors such as Heparin, RNA, and arachidonic acid have been widely used to form Tau fibrils for *in vitro* study. With the advent of atomic models, the goal posts for *in vitro* work were moved, and the goal of recreating the structure with recombinant material became paramount[48].

When we began the work presented below, the atomic structures of the fibrils produced *in vitro* (or in any Tauopathies beyond AD) were unknown, but morphological analysis of the fibers with negative stain TEM indicated that fibrils formed with heparin were like the PHF fibrils in AD. There is, of course, significant interest in the *in vitro* reproduction of Tau filaments for therapeutic developments. For compound screening assays and other drug development pipelines to be useful, the starting material (fibrils in some cases) must faithfully reproduce the structures of the 4 classes of Tau structures that will eventually be targeted in patients. We used double electron electron resonance DEER to determine the structure formed by Tau in heparin-made Tau fibrils (Chapter 2), and we determined that not only did we fail to produce PHF fibrils of Tau, but heterogeneous fibril populations were forming. This result spurred on much of the work that follows: we explored multiple pathways towards generating a representative fibril, and characterized the aggregation using a variety of electron paramagnetic resonance (EPR) and microscopy techniques described below.

1.5 Experimental Methods

1.5.1 Electron paramagnetic resonance for biological systems

Structural determination of proteins has made immense gains in the understanding of protein function, but we are often limited in our understanding of their dynamics. To understand the true function of proteins and how to tune their function, we must understand the dynamics of the system. Electron paramagnetic resonance (EPR) provides a unique method for measuring and describing the dynamics and structure of proteins quantitatively across length scales between 2 Å and 7 nm. EPR requires small amounts of samples relative to other biophysical methods like NMR, on the order of 100s of µg instead of the milligrams required for NMR.

EPR relies on the behavior of unpaired electrons in the presence of a magnetic field when irradiated with a microwave. When placed in a magnetic field, the up and down spin-states of an electron split in energy. If a microwave of that energy is applied to the electron, there is an excitation that occurs that can be observed as an absorption of energy[49].

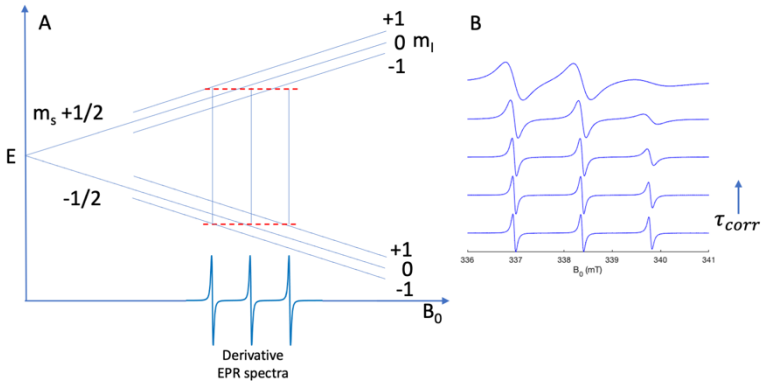


Figure 1.3: A) An energy diagram shows the splitting in energy by the spin state (m_s) of an electron. Additional hyperfine splitting of the ^{14}N nuclei quantum number (m_I) the nitroxide radical. Dashed red lines indicate the energy of the microwave source. On the x-axis, a representative derivative spectrum that would be acquired by

sweeping the magnetic field (B_0) with CW EPR. Figure adapted from Bordignon, 2017[50]. B) CW EPR spectra of a nitroxide radical with different tumbling rates. Changes in spectra are observed due to partial averaging of anisotropies. Spectra were simulated with Easyspin[51].

In a nitroxide radical (used in all studies presented) the spin Hamiltonian depends on the electron Zeeman, hyperfine and nuclear Zeeman terms (Eq. 1). The electron Zeeman (term 1, Eq.1) defines the center-field splitting due to the spin-state (m_s). Hyperfine interactions between the electron and nuclei lead to additional splitting in the A tensor (term 2, Eq.2). The ^{14}N nuclei has a +1, 0, and -1 quantum number with 3 associated energy gaps give the characteristic 3 distinct peaks of the nitroxide spectrum (Figure 1.3A).

$$\boldsymbol{Eq.1} \qquad \hat{H} = \mu_B \boldsymbol{B} g \hat{\boldsymbol{S}} + \hat{\boldsymbol{S}} \boldsymbol{A} \hat{\boldsymbol{I}} + \mu_N g_n \boldsymbol{B} \hat{\boldsymbol{I}}$$

1.5.2 CW EPR

In continuous-wave EPR (CW EPR), the sample is irradiated with a continuous microwave. The frequency of the microwave is held constant, the field (B_0) is swept through the range of interest, and the absorption is measured in a resonant cavity containing the sample. In practice absorption is low, and the signal-to-noise ratio (SNR) is too low for typical radical concentrations. To enhance SNR, an oscillating field (usually 100 kHz) is imposed in the direction of the static magnetic field. The field oscillation creates an absorbance signal that can be amplified. The amplitude of the signal wave that is produced is equivalent to the derivative of the EPR spectra, so the derivative of the absorption peak is recorded (Figure 1.3A).

In the isotropic limit, wherein spins are tumbling more rapidly than the relevant spectral width of the nitroxide radical (10s of picoseconds or faster), *g* anisotropies are averaged out, but as the tumbling approaches similar timescales to the spectral width, anisotropic

components become visible. By fitting the CW EPR spectra to simulations of the nitroxide, a rotational correlation time (τ_{corr}) can be calculated. τ_{corr} provides a useful parameter for estimating complex size in the context of fibrils. As a spin label is embedded in a larger fibril, its tumbling rate decreases. Thus, fibrillar species can clearly be delineated from soluble species. This technique is used in Chapters 3, 5, and 8 to quantify the degree of aggregation.

Radicals in proximity to another, experience dipolar interactions and exchange coupling. Dipolar broadening and exchange can be detected at distances below 1.5 nm in CW. EPR. Dipolar broadening is a predominantly anisotropic effect, and needs low temperature glassed samples in order to quantitatively extract distances between spin labels. In this manuscript, a spin exchanged component is fit to solution CW EPR spectra to obtain information about how many spins are within 1.5 nm. The heterogeneity of amyloid fibril samples makes quantitative distance estimation from CW EPR difficult, but fitting provides good estimates of the fibril state in Chapter 3, and 5.

1.5.3 Double electron-electron resonance

Double electron-electron resonance (DEER), also referred to as pulsed electron double resonance (PELDOR), is a pulsed EPR technique that relies on measuring the distribution of dipolar couplings between intra-molecular pairs of nitroxide free-radical spin labels. At distances larger than 1.5 nm, through-space exchange interactions are negligible, so DEER is detects only through-space dipolar coupling. Dipolar coupling is dependent on the distance, *r*, between spins, so DEER can extract a distribution of end-to-end distances between the pair of labeled sites. DEER is capturing distance information over at least 10^{14} different spin pairs, a much larger number than single molecule FRET and does not rely on a predetermined distance distribution model. Furthermore, all conformations are equally weighted in DEER, so close-range contacts are not overly weighted as might be observed in cross-linking and FRET

experiments, rendering DEER a more quantitative approach. Most studies of Tau measure intramolecular distances, although DEER has been used to measure intermolecular distances in ordered oligomeric complexes for example. Intramolecular distance measurements use two cysteine mutations to provide defined spin label sites. Measurement of soluble Tau is straightforward, but the close intermolecular packing of Tau in amyloids presents a significant number of intermolecular distances at relevant DEER distances. Spin-labeled Tau is mixed with unlabeled Tau by a 10-20x dilution factor so that random spacing along the axis of the fibril so that intermolecular interactions are minimized. Dipolar coupling is an anisotropic effect, so the samples are frozen in a glassed state to make the anisotropic effect observable.

The Han lab uses a 4-pulse sequence to record DEER. Two frequencies are selected, an observe frequency, and a pump frequency each of which excites separate spin populations. First, a Hahn echo sequence is applied to the 'observe' channel with a 90˚ followed by a 180˚ refocusing pulse after time (τ_1) flips the spins to produce an echo. Next, a 180˚ pump pulse is applied to the probe channel at an incremented time delay (T). This pulse flips the spins in the 'pump' population of spins, which exert a change in the magnetic field experience by the 'observe' spins. After a time delay ($\tau_1+\tau_2$), the probe spins are refocused with a second 180˚ pulse, and the echo is recorded as a function of the time delay (T) (Figure 1.4). The recorded signal is a convolution of oscillating decay signals. The signal S(t) can be decomposed into the characteristic decay frequencies, and a distance distribution P(r) can be calculated.

Calculating the P(r) from S(t) is an ill-posed problem, so there are multiple methods for calculating it, all relying on different approximations or assumptions. In Chapter 2 the signal was fit to a single gaussian P(r) distribution. As numerical methods and confidence and acceptance of the DEER technique grew, later work in Chapters 3 and 5 uses Tikhonov

regularization, or single value decomposition (SVD), algorithms to generate model-free distributions that is more representative of the true P(r) distribution being measured.

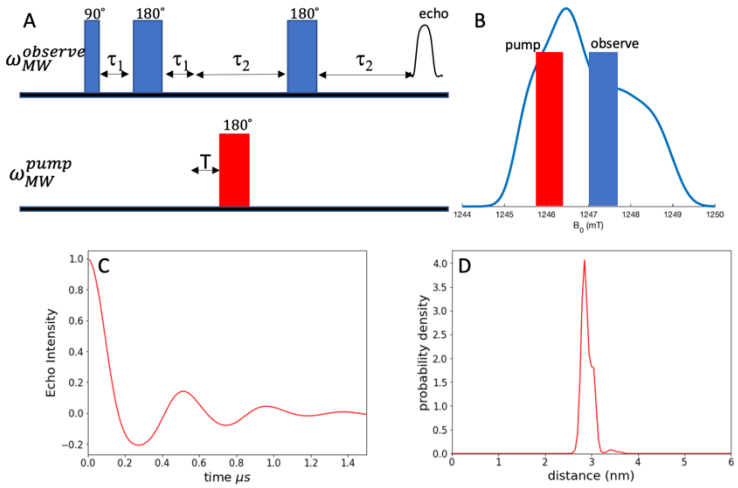


Figure 1.4: A) A schematic of the 4-pulse sequence used to record DEER. An observe channel undergoes 3 pumps, and a pump population experiences 1 pulse. An echo is detected and the dependence on the time delay of the pump pulse is recorded. The pump pulse flips the associated spins and creates a change in the magnetic field experienced by the observe spins, altering the magnitude of the echo detected. B) EPR spectra of a nitroxide spin experiencing dipolar coupling while in a solid state. Representative frequencies of the pump and observe pulses are shown. Spectra were simulated in Easyspin[51]. C) Echo decay signal that is recorded by DEER of the P(r) distribution shown in (D). The time domain signal (C) of a narrow distribution such as the P(r) shown in (D) shows a clear decay frequency. More complex, or broad P(r) shown throughout the manuscript have a time domain that is a convolution of all distances. S(t) and P(r) were simulated using DEER-PREdict to describe the expected outcome of labeling sites 351 and 373 in the fibrils from PSP.

1.5.4 Overhauser dynamic nuclear polarization

Overhauser dynamic nuclear polarization (ODNP) combines NMR and EPR techniques to measure water dynamics surrounding a spin label. ODNP relies on cross relaxation of H^1 nuclei of water molecules induced by a free electron spin of a nitroxide spin label. Because the electron-^{1}H cross relaxation falls off with the electron-^{1}H distance r following $1/r^6$, ODNP is sensitive to the dynamics of water within 8 Å of the spin probe. Hence, if the spin probe is localized to a specific site, ODNP can report on local hydration

water dynamics near that site. The efficiencies of electron-^{1}H coupling in solution state that leads to cross relaxation is directly dependent on the speed of movement of water near the electron spin probe and are reflected in ^{1}H NMR signal enhancement. Through the measurement of the ^{1}H NMR signal enhancement and ^{1}H T_1 relaxation times, the electron-^{1}H cross-relaxivity, k_s, can be determined. In the dynamic regime of relevance in this study, k_σ monotonically decreases with slower water diffusion dynamics. ^{1}H NMR ODNP measurements performed at 0.35 Tesla are sensitive to translational diffusion dynamics of water with correlation time in the 10's ps to 100 ps timescale, which corresponds to freely diffusing hydration water (HW) bound to the protein surface with similar or stronger hydrogen bond strength than to surrounding bulk water molecules.

1.5.5 Electron Paramagnetic Resonance in Biophysics

EPR fills a unique role in the biophysical community with its ability to provide quantitative ensemble measurements of parameters of interest. In biological systems, the free radical source is often a paramagnetic spin label attached to select locations to provide local information. For proteins, site-directed spin labeling (SDSL) utilizes single amino acid mutation to place a cysteine at the site of interest. A nitroxide based spin label, with a free thiol group MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate) is attached via a disulfide bond to the cysteine[52].In this work, EPR is used to accomplish three main tasks: **I)** Generate intramolecular distance ensembles that describe the conformation of the protein by measuring the distance between two labeled amino acids. We utilized DEER to form low-resolution reconstructions of Tau fibril structures, and to describe the effect of disease-relevant single amino acid substitutions to peptides of Tau. **II)** continuous wave EPR (CW EPR) is used to explore local conditions near labelled sites on

Tau, and its binding partners. CW EPR reports on many parameters, of most interest in this study is the local mobility of the spin label. When in a fibril, the tumbling mobility of the spin-label is significantly reduced. Similarly, protein-protein binding can reduce the mobility at locations near the binding site. This technique was used to quantify the binding of Tau to an SH3 domain, as well as quantifying the dependence of *in vitro* Tau fibrils on anionic cofactors that are integral to the fibril. **III)** The excited free electron can also be utilized to polarize H^1 protons in the surrounding water with Overhauser dynamic nuclear polarization (ODNP). ODNP was used to quantify the local hydration dynamics surrounding Tau prior to fibrilization in Chapter 5. We found that single amino acid substitutions disrupt the hydration shell surrounding Tau peptides, leading to a more entropically favorable dewetting of the protein.

Chapter 2: Heparin-induced Tau filaments are structurally heterogeneous and differ from Alzheimer's disease filaments

Text and Figures reprinted and adapted from Fichou, Y. et. al. (2018). Heparin-induced Tau filaments are structurally heterogeneous and differ from Alzheimer's disease filaments. doi.org/10.1039/C8CC01355A *published in Chemical Communications by The Royal Society of Chemistry*

2.1 Introduction

Because brain-extracted material is rare, expensive and often impure, in-vitro Tau aggregate models have been developed and widely used for common biophysical characterizations. The most commonly used filament model is made with the help of heparin, as it was proposed to make Alzheimer-like aggregates when mixed with recombinant Tau protein[53]. Subsequently many structural, dynamic and thermodynamic studies of Tau aggregation have been carried out using heparin-induced aggregation.

The question of biological relevance in using heparin-induced filaments has become critical in the last few years, as the notion of Tau strains gains in popularity. Strains are defined as distinct types of aggregates that display unique toxicity properties, are specific to a disease phenotype and can faithfully propagate themselves through biological tissues[14]. Several studies have shown, using antibody specificity or limited proteolysis[27,54–56], that different Tau conformations are involved in different strains, but precise structures of individual strains have remained elusive. Thus, the question arises which strain does *in-vitro* heparin-induced filaments represent most closely, and whether or not they form a single structure or multiple polymorphs.

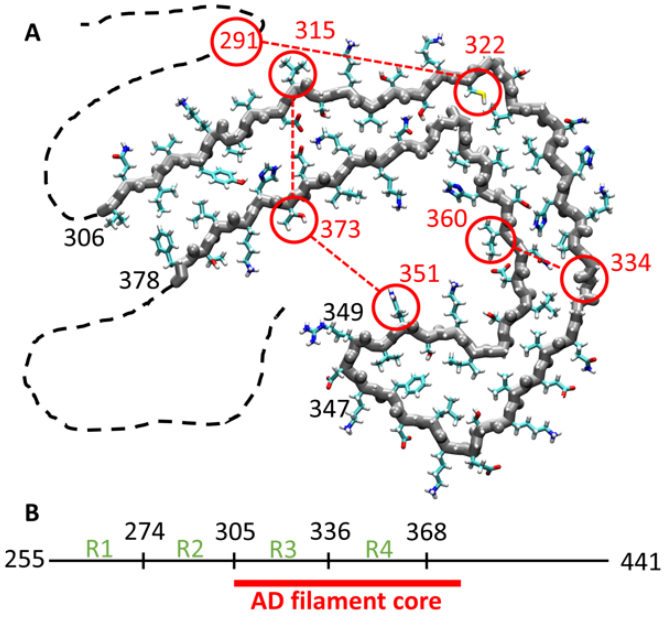


Figure 2.1: A) Schematic view of the AD Tau protofilament structure (PDB ID 5O3L). It features a C shape and anti-parallel cross-β structure between residues 306-347 and 349-378. The red circles connected with a dashed-line represent pair distances that have been measured by DEER. B) Representation of the 187-residues truncation of Tau used in this work, which comprises most of the 4 repeat domains (R1-R4) and the C terminus.

In this work, we aim at revealing whether or not heparin-induced filaments, often used as an AD Tau filament model, retain the main structural features discovered from AD brain-

extracted Tau filaments. In particular, we tested the existence of (i) the C shape in the filament core, (ii) the anti-parallel cross-β structure between 303-347 and 349-378, and (iii) the interaction between R3 and R2. We also tested if a stronger packing exists in the region R3 and R4 as compared to R1 and R2.

We used pulsed electron paramagnetic resonance spectroscopy (EPR), including double electron-electron resonance spectroscopy (DEER) at Q-band to probe local packing and intramolecular distances, respectively. DEER has the unique ability to measure distance distributions among the sample in the 1.5-8 nm range between two electron spin labels tethered to the protein[57]. Because DEER and EPR are solely sensitive to paramagnetic elements, they are particularly well suited for tracking protein aggregation[58], where protein crowding and sample heterogeneity makes the use of NMR or crystallography methods challenging, especially for samples generated *in situ* without further workup.

We used a truncated Tau variant that contains residues 255-441, *i.e.* all of the four repeat domains and the C terminus but not the N terminus (see Figure 2.1B). This construct is referred to as Tau throughout this manuscript. MTSL electron spin labels were tethered to cysteine residues, which were added to the Tau protein by site-directed mutagenesis of the indigenous residue. Aggregation was triggered by the addition of heparin. Measurements denoted as before and after aggregation were performed before and 24h after mixing Tau and heparin, respectively (see supplementary material and methods for more details). Aggregation kinetics were measured by thioflavin T fluorescence (Figure S2.1A) and mature filaments were imaged by transmission electron microscopy (Figure S2.1B). DEER data were analyzed using a Gaussian distance distribution model because it gave data fitting of sufficient quality (see Figure S2.9) to provide a reliable assessment of the distance distribution mean and width.

2.2 Results

2.2.1 Choice of spin-label sites

The locations of the spin label pairs were rationally chosen using a solution state structure ensemble composed of 1356 Tau conformers published by Schwalbe and coworkers[59]. We used this ensemble to map intra-molecular pair-wise distances and retained distances that fall in the DEER measurable range (Figure 2.2). DEER excels at measuring distances between 1.5 nm and 5 nm, so by computing residue pair distances from a Tau (hTau40 isoform) solution ensemble[59] (Figure 2.2), we could eliminate pairs that would be too close, or far apart for accurate measurement.

The agreement between the DEER-measured distances and ensemble-calculated distances was good (Figure S5 and S6), offering confidence in the solution ensemble. We then computed the expected distance shift between the solution ensemble and the AD protofilament structure. Finally, we picked the label pair positions that led to a significant distance shift between solution and aggregated states through the following filament structural features: (i) C-shape, (ii) anti-parallel cross-β packing, (iii) interaction between R2 and R3 regions. Additionally, to give the best resolution possible, we estimated the distance changes expected from solution state to an AD protofilament (PDB ID 5O3L; Figure 2.2). After considering constraints on amino acids deemed to be detrimental to mutate, such as residues internal to the β –sheets, we designed mutations that would provide the largest change in distance.

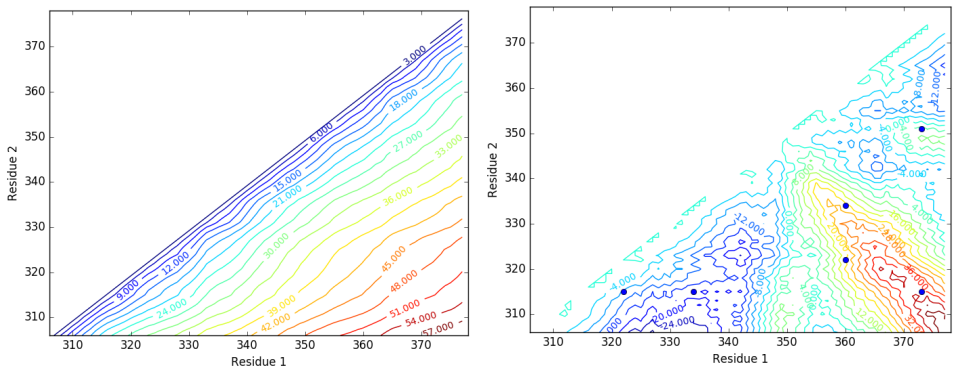


Figure 2.2: (Left) Average residue pairwise distances (C_α-C_α in the range 303-378 computed from a solution state ensemble of the Tau isoform hTau40[59]. Only pairs consisting of residue 1 > residue 2 are shown for simplicity. Contour map represents the pairwise distance in Angstroms with a step of 3Å between each line. Cold to warm colors depict small to large distances, respectively. (Right) Pairwise distance change from solution state to AD filaments. The pairwise changes were calculated by subtracting the PHF distances (C_α-C_α) from the average solution distances shown in (left). Only pairs consisting of residue 1 > residue 2 are shown for simplicity. Contour map shows the distance shift in Angstrom with a step of 4 Å between each line. Positive numbers (bluer contour) reflect a decrease of the distance from solution state to PHF. Blue dots show pairwise distances we have measured with DEER in this study.

2.2.2 P(r) distributions of Tau filaments

Figure 2.3 presents DEER-derived distances for heparin-induced Tau filaments compared with the simulated DEER distance distribution expected from AD Tau filaments[38]) using the RotamerConvolveMD (see below for discussion of simulations)[60,61]. The DEER distributions were simulated assuming a 1:1 mixture of PHF:SF (see Figure S7). Residues 351 and 373 were used to probe the existence of a C-shape in the Tau protofilament, which should bring these residues in close proximity. As shown in Figure 2C, the measured distance between residues 351 and 373 is much larger than expected from the AD-filament structure, demonstrating that heparin-induced filaments do not adopt the C-shape shown in Figure 2.1. The β-solenoid structure in the region G355-E338 allows the formation of intra-molecular anti-parallel β sheets between the segments V306-K347 and F378-R349. We probed the existence

of this feature by measuring distances between residues 334 and 360, and between residues 315 and 373. The finding that these residue pairs have an average separation of 4.3 nm and 4.7 nm, respectively, (Figure 2.3B and 2.3A) shows that the anti-parallel β-sheet structure does not predominantly exist in the heparin-induced filaments. Although not found in the filament core, residual electronic densities suggested that residues 290-305 of R2 (in the case of 4R Tau) fold back along R3 [Figure 2 in Fitzpatrick, 2017[38]], bringing K290 approximately in register with L315. We measured the distance between the two natural cysteines 291 and 322 to probe for this structural feature. The measured average distance of 4 nm (Figure 2.3D) remains significantly greater than the approximate expected distance of 3.4 nm in the case where R2 would fold onto R3 (see SI for details on the distance estimation). We thus conclude that heparin-induced filaments do not reproduce the key structural features found in AD filaments.

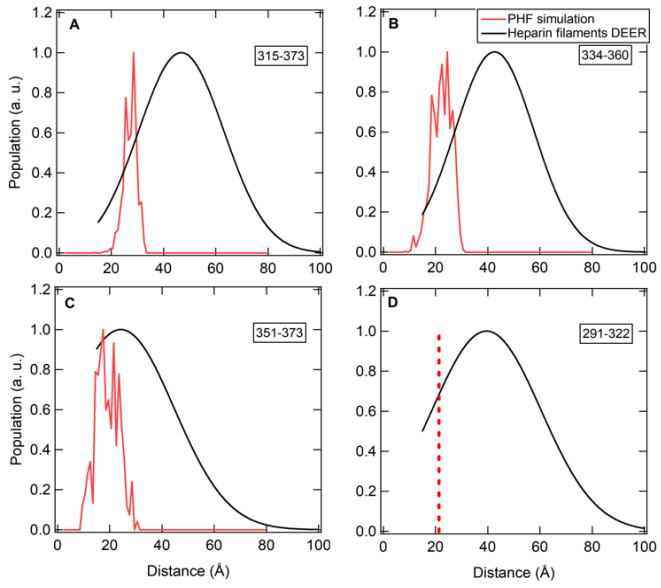


Figure 2.3: Intra-molecular DEER distance distributions measured for heparin-induced Tau filaments (black) and simulated from AD-filament structure (PDB ID 5O3L, red). The presence of intra-molecular anti-parallel β-sheet in the AD filament results in an

expected proximity of residue pairs 315-373 and 334-360 (A,B), the C shape of the filament and the refolding of R2 on R3 should bring residues 351-373 and 291-322 (C,D) close to each other, respectively. None of these features are measured in the heparin-induced filaments as distances are much larger. The broad distance distributions reveal heterogeneity of heparin-induced filament structure.

2.2.3 Heparin induced conformational stretching of Tau into β-sheets

Considering our hypothesis where heparin templates Tau aggregation, i.e. induces Tau conformational changes necessary to adopt an amyloid structure (e.g. opening of PHF6(*) fragments), the drastic difference between heparin-induced filaments and the AD filaments can be reconciled with the physical properties of heparin. Heparin has been found to be fairly extended (see for instance molecular models in [62]), and an experimental persistence length of ~2 nm measured[63]. It follows that high-curvature structural features, such as β-solenoid or refolding of R2 cannot easily be imprinted when using a generic heparin molecule as a template, without additional chaperones. Rather, one could expect heparin to transfer its generally extended state to the protein and thus result in an overall stretching of Tau protein segments.

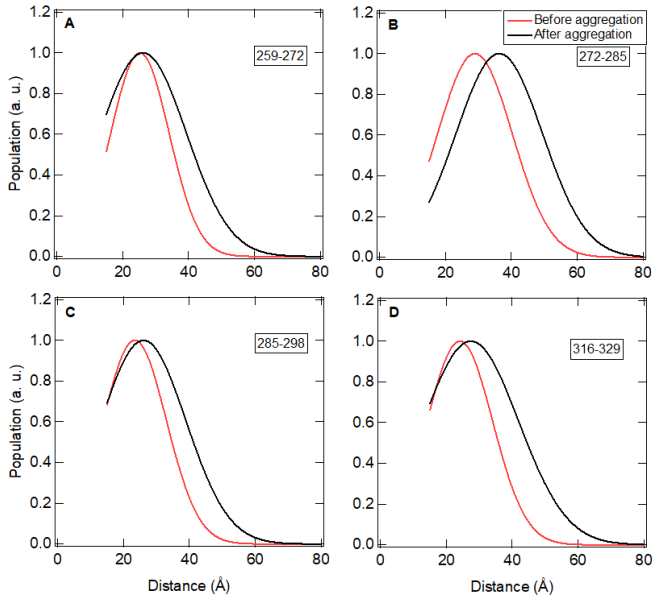


Figure 2.4: Intra-molecular distance distribution of 13-residue-apart pairs before (red) and after (black) aggregation. All these segments extend upon aggregation, revealing an overall stretching of Tau conformers in the aggregating state.

We have previously shown an opening, immediately after heparin addition, of the aggregation-prone segments ^{275}VQIINK^{280} and ^{306}VQIVYK^{311} by measuring distances between resides 272-285 and 303-316, respectively[64]. Here we reproduced the measurements for residues 272-285 (following the same conditions as for the data shown in Figure 2.3), in addition to three other distances, 259-272, 285-298 and 316-329 (Figure 2.4), spanning over R1, R2 and R3. All of these distances, which are measured between residues that are relatively close to each other (13 amino acids apart) exhibit a distinct increase in distance upon aggregation (Figure 2.4A, B, C and D). This increase is compatible with the formation of a local c structure that is typically shorter than 10 amino acids and result in a local "flattening", and thus an extension of local residue pair distances. Therefore, the extensions visible in Figure 2.4 suggest that heparin-induced filaments possess an overall stretched structure due to cross-β structures, and minimal large-domain structuring. This stretching will expose the hydrophobic PHF6(*) region to solvent, and so enhance the aggregation propensity of Tau.

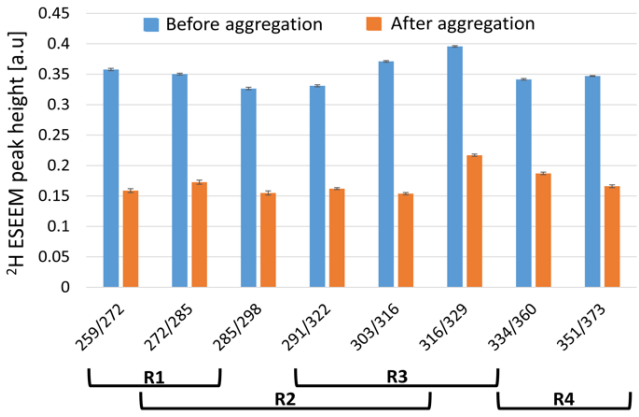


Figure 2.5: ESEEM deuterium peak height in the frequency domain for different label positions. The peak height reflects the average quantity of solvent (D_2O) molecules present around the two spin labels tethered to the protein. The position of the labels relative to the repeat domains R1-R4 are represented in the lower part. Overall, all labels similarly dehydrate upon aggregation, revealing a homogenous packing of the 4 repeat domains and suggesting that they are all part of the filament core.

2.2.4 Fibrils are dehydrated around spin-labels

In order to further measure structural arrangements of the repeat domains in heparin-induced filaments, we apply electron spin echo envelope modulation (ESEEM) on various mutants. Deuterium ESEEM signal is proportional to the quantity of deuterium atoms present in the vicinity of the spin labels. Because the proteins are placed in deuterated buffer, we used this method to quantify the hydration of the labelled sites before and after aggregation, thus assessing their degree of packing in the fibrillated state. Figure 2.5 reports on the height of ESEEM's deuterium peaks measured for 8 mutants labelled at different position throughout the repeat domains. All sites show a decrease of the ESEEM signal after aggregation, reflecting a significant dehydration of the spin labels. This dehydration originates from the local protein packing around the spin label. The measured dehydration upon aggregation is fairly homogenous across the four repeat domains, strongly suggesting that all four repeat domains are part of the structured filament core. This is in good agreement with previous structural characterization reports of heparin-induced Tau filaments[65,66].

2.2.5 Comparison between DEER measurement and simulations using RotamerConvolveMD

The distributions of all distances measured by DEER in Figure 2.3 and 2.4 remain markedly broad after aggregation, spanning distances between 3.0 nm and 4.9 nm (full width half maximum (FWHM)). The use of MTSL as a spin-label for DEER measurements

introduces multiple additional rotational degrees of freedom inherent to the MTSL structure. The flexibility of MTSL artificially broadens the distance distribution measured by DEER. We used a python package, RotamerConvolveMD[60,61], to simulate the DEER signal of Tau with MTSL spin labels in both a PHF conformation[38], and in solution based on an ensemble of conformations generated for hTau40[59]. RotamerConvolveMD appends a library of MTSL onto label positions and calculates a distribution function based on all possible distance pairs that do not contain any steric hindrance between the spin label and the protein. This simulation allows us to compare structural data from cryo-EM and other measurement techniques to DEER data directly. For each spin label pair we measured with DEER, we also simulated the expected distance distribution for the solution state. We found good agreement between DEER measurements and the calculation from the solution ensemble (Figure S4 and S5). Only distances between 315 and 373 show a significant discrepancy, which can be explained by long expected distances that cannot be measured by DEER. In addition to the measured and simulated solution state in Figure S4 and S5, we also overlaid on Figure S5 the expected and measured distances after aggregation (same data shown in Figure 2.3 and 2.4). The good agreement between simulations and measurements in the solution state but not in the aggregated state reinforce the confidence in the DEER measured distances and thus in the conclusion that heparin-induced filaments are significantly different from the AD filament.

2.2.6 Broadening due to MTSL rotations

Additionally, the simulation of the MTSL rotations tethered to the PHF and SF enable the evaluation of the distance broadening due to MTSL flexibility. Indeed, with no rotational contributions from MTSL, any pairwise distance distribution on a PHF would be a Dirac distance. However, the calculated distance distributions on the PHF plotted in Figure 2.3 show distribution broadening of 2.1 nm for a probability value greater than 1% of the most probable

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

Although the flexibility of the spin labels tethered to the cysteine residues broadens our measured distribution, it cannot account for such broad distributions. This means that even after aggregation, the protein exists in several different conformations that are very distinct from each other, giving rise to broad pair-wise inter-residue distance distributions. Thus, our results demonstrate that heparin-induced (proto)filaments are structurally heterogeneous and exist in several different molecular arrangements adopting a wide range of protein folds.

When filtering out species smaller than tetramer, the distance distribution did not change (Figure S2.8), showing that a broad distribution cannot be explained by a mix of unfibrillated monomers and aggregates in the sample. Most of the established aggregation models suggest that Tau oligomers are transient and evolve toward mature filaments[67–69], suggesting that we are probing several arrangements of mature Tau filaments. Polyanion chains such as heparin or RNA are thought to induce aggregation by establishing electrostatic interactions and triggering conformational rearrangements. However, because these cofactors are mostly disordered chains, it is not surprising that they cannot induce a well-define single structure of Tau filaments. This argument becomes even more relevant in light of recent findings that Tau aggregates display prion-like behavior[70]. This means that Tau aggregates have different strains (which have been shown to be disease specific[54]), and that each strain is faithfully propagated through seeding. In other words, a structurally-defined seed can template

assembly of naïve monomers into the same structure. In this view, because heparin lacks this well-defined structure, it is expected to form polymorphic Tau aggregates rather than a unique Tau strain.

2.3 Discussion

In summary, we have established that heparin-induced filaments are not only different from *in-vivo* AD filaments, but also are structurally heterogeneous. This finding should warn us against using polyanion-induced aggregation for any structure-based analysis such as drug design or antibody development. Although triggering aggregation with a cofactor remains interesting to understand general aggregation mechanisms, it is likely that the use of a structurally-defined template (using biologically-extracted seeds for instance), potentially in addition to polyanion cofactors, will prove more applicable to biological phenomena.

2.4 Conclusion

This work described the structure of Tau187 fibrils formed with heparin and found that there was a heterogeneous population of fibrils formed under these *in vitro* aggregation conditions. This result stands in stark comparison to the findings of postmortem Tauopathy material wherein a single structure (or polymorphs of the same structure) are observed. This work was supported and validated by cryo-EM studies shortly after its publication[71]. Zhang et. al. found that heparin-formed fibrils formed structures separate to those in AD, and that there are multiple species present in each sample. Our work made it clear that a renewed focus on creating recombinant Tau fibrils that recapitulate structures in Tauopathies is necessary. However, this was a puzzling result in that the highly controlled, homogenous conditions of *in vitro* aggregation of recombinant Tau produces a heterogenous mix of fibril, while the complex media of the neuron can produce homogenous populations of fibrils of one structure type. We

hypothesized that this result indicated that there is some chaperone, cofactor, or templating agent that interacts with Tau during fibrilization *in vivo* to direct Tau into one singular structure. With this hypothesis in mind, we turned attention to cofactors, including heparin and RNA to determine the nature of the of interaction between Tau and these non-proteinaceous cofactors. What follows in Chapter 3 describes the result of this study.

Chapter 3: Recombinant Tau, aggregated with Ionic cofactors forms fibrils with cofactors as an integral internal constituent

Text and Figures reprinted and adapted from Fichou, Y. *et. al*. (2018). Cofactors are essential constituents of stable and seeding-active Tau fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(52), 13234–13239. https://doi.org/10.1073/pnas.1810058115

3.1 Abstract

Amyloid fibrils are cross-beta rich aggregates that are thought to be the most stable form of protein assembly. Accumulation of Tau amyloid fibrils is involved in many neurodegenerative diseases, including Alzheimer's disease, where neurofibrillary tangles (NFT) are composed of Tau amyloid fibrils. Heparin-induced aggregates have been widely used and assumed to be a good Tau amyloid fibril model for most biophysical studies. Here we show that mature fibrils made of 4R Tau variants, prepared with heparin or RNA, spontaneously depolymerize and release monomers when their cofactors are removed. We demonstrate that the cross-β sheet assembly formed in vitro with polyanion addition is unstable at room temperature, counter to common assumptions. We furthermore show that the high seeding capacity observed with mouse brain-extracted Tau fibrils gets exhausted after one generation, while supplementation with RNA cofactor resulted in sustained seeding over multiple generations. We suggest that Tau fibrils formed in brains are supported by cofactors, which may be critical in the pathological seeding capacity of Tau aggregates and strain differentiation. Our study shows that more resources should be invested into understanding the role of cofactors in Tauopathies that may be viable targets for diagnosis and therapeutics.

3.2 Introduction

Amyloid aggregates are structured aggregates that are characterized by a high cross-β sheet content. They rely on an extended intermolecular hydrogen bond network that provides high stability, and allow even hydrophilic and charged domains to be dehydrated and tightly packed in a protein assembly[72]. Amyloid aggregates have been conjectured to be the most stable form of protein assembly[73]. However, direct measurements of their thermodynamic parameters are scarcely reported in the literature, and the notion of high stability tends to be generalized to all cross-β-rich assemblies, even when direct measurements have not been carried out.

The stability of amyloid fibrils is an important field of research as their clearance is a central question in many therapeutic approaches. Indeed, accumulation of amyloid fibrils are involved in several diseases including Parkinson's and Alzheimer's disease (AD)[11]. In many cases, the clearance of deposited amyloid fibrils is the target of drug developments strategies (see for instance [74]).

The Tau protein is an intrinsically disordered protein that is mostly present in neurons and can form amyloid fibrils in several neurodegenerative diseases including AD, frontotemporal dementia and Pick's disease[75]. Tau is highly charged and hydrophilic, making it highly soluble and stable in aqueous environments across a range of pH and temperature. Yet under pathological conditions, it can assemble into amyloid fibrils in which parts of its microtubule binding domain, predominantly positively charged, densely pack into a cross- β sheet arrangement[38]. However, the triggering factors and driving forces of fibrillation of the highly soluble Tau protein remain unknown, and consequently neither the mechanism that

achieves convergence to a unique fibril structure, nor the main fibril feature that contribute to their stability, is understood.

Aggregation mechanisms are most often studied with recombinant, truncated, Tau proteins. In vitro fibrillation of Tau is typically triggered with the help of cofactors, most commonly heparin[76], but also other cofactors such as RNA[77] or arachidonic acid[78]. In the last few years, seeding aggregation has been shown to be possible by adding pre-made fibrils (seeds) to fresh monomers. The seeding process is still highly improved by the presence of cofactors (RNA or heparin) in solution[79,80], yet the roles of cofactors in the fibrillation process remain controversial, and their effect on mature fibril properties essentially unstudied. Although it has been shown that heparin is a limiting factor in fibril formation[69,81], the exact nature of association of heparin with the mature fibrils remains unclear, as heparin has been reported to be part[82] or not be part of mature fibrils[81,83]. Heparin-induced fibrils were shown to be more stable than AD fibrils, using chemical denaturation[56,84], but it is not clear if heparin contributes stability to the fibrils, by catalyzing a structuration of Tau amyloid fibers that are more stable than those in AD fibrils, or if the heparin contributes a direct stabilization. In this study we assess whether cofactors are crucial constituents of mature fibrils and contribute to their stability, or whether they only promote aggregation toward a self-sustained protein assembly.

Tau aggregates have been shown to propagate from neuron to neuron, and to be able to seed aggregation[56,85], i.e. convert naïve Tau monomers into aggregates. This led to the hypothesis that in vivo, monomeric Tau can spontaneously polymerize into amyloid filaments when an appropriate seed template is provided. For that reason, the pathological origins of Tau aggregation have mostly been searched in the properties of Tau itself, such as

hyperphosphorylation, cleavage, high local concentrations, alternative splicing, but marginally in abnormal interactions with other cofactors. Paradoxically, cofactors are always used for in vitro aggregation of Tau and even assumed to be biologically relevant. Therefore, gaining an understanding about the influence of cofactors on mature fibril properties will (i) provide key insight into the role of cofactors in in vitro fibrils stability and conformation, and (ii) guide the search for cofactors that assist in seeding and spreading of Tau aggregation in vivo.

Here, we used a set of biochemical tools together with electron paramagnetic resonance (EPR) to characterize the consequences of cofactor removal after fibril formation. We used both heparin and RNA cofactors to trigger Tau fibrillation. Upon cofactor enzymatic digestion, the fibrils destabilized, and released monomers that have lost their aggregation-prone conformation. We conclude that Tau fibril models made in vitro require cofactors to be stable and ask how this finding relates to aggregates formed in vivo. We propose that abnormal interactions of Tau with cofactors in neurons should be considered as potential critical events that could provide new directions to interfere with pathology development, and to differentiate among different Tauopathies.

3.3 Results

3.3.1 Fibrils assemble when the cofactors are present

We used a truncated version of the longest human Tau isoform 2N4R that contains the four repeat domains (R1-R4), as well as the entire C-terminal region (residues 255-441, named here as Tau187), from which one of the two cysteines was mutated (C291S) in order to perform site-directed spin labeling for EPR[50,67]. The construct, with the addition of the aggregation-promoting disease mutation P301L[86] is referred to as Tau throughout the manuscript (Fig, S1A). Polydisperse heparin (average MW 15 kDa) and polyU (RNA, average MW 900 kDa) were incubated with Tau to induce fibrillation. The addition of heparin or RNA to Tau resulted

in amyloid fibril formation as verified by significant increase of Thioflavin T (ThT) fluorescence intensity (Figure S2) and the presence of fibrillar structures captured by transmission electron microscopy (TEM, Figure S1B, C). We referred to the heparin-induced and RNA-induced Tau amyloid fibrils as heparin fibrils and RNA fibrils, respectively.

3.3.2 Fibrils depolymerize when the cofactors are digested

We tested whether these cofactors act as catalysts that assist fibril formation and subsequently dissociate from the product, or whether they are reactants that are part of the fibril scaffold and are necessary to ensure the stability of mature fibrils. To address this question, we investigated the amyloid fibril quantity by preparing heparin fibrils and RNA fibrils, and then degrading the cofactors via enzymatic digestion.

Heparin and RNA are cofactors that can be digested using heparinase and RNase, respectively (see Methods and Materials). In their digested form, both cofactors become incapable of triggering fibril formation (Figure S4). ThT fluorescence, which provides an in-situ measure of cross-β sheet structures, was used to quantify the amyloid fibrils present in the sample. Heparinase and RNase were added to heparin fibrils and RNA fibrils after maximal ThT fluorescence was reached, and were incubated for 7 hours, resulted in a 20-30% and 60-70% decrease of ThT fluorescence, respectively (Figure 3.1A). This reduction of ThT fluorescence was 2-3 times more significant than the effect of protein dilution (by enzyme addition). The control samples where buffer without enzymes was added showed a signal decrease of about ~10% and ~30%, respectively (Figure 3.1A). These results show a significant loss of cross-β sheet when cofactors were digested.

The loss of Tau amyloid fibrils was independently verified by continuous wave EPR (CW EPR). CW EPR of spin labelled Tau has been previously used to assess the packing and

mobility of Tau fibrils[67]. The CW EPR spectra of amyloid fibrils can be decomposed into 3 components, a mobile component corresponding to soluble species, an immobile component corresponding to large molecular weight species (oligomers, aggregates, etc) or interfacial proteins, and a spin-exchange component resulting from spin labels that are within 5 Å proximity due to parallel, in register, amyloid cross-β stacking of Tau[67]. Here we spin-labeled the native cysteine 322 and acquired CW EPR spectra after aggregation with heparin, as well as after digestion with heparinase. We fit the spectra using an established simulation protocol for EPR lineshapes into mobile, immobile and spin exchange components (Figure 3.1B). Fitting results revealed that 4.8% of the tethered spins participated in the mobile component in the fibrils, which increased to 17.7% after heparinase treatment, and 86.6% of spin-exchange component decreased to 78.6% (Figure 3.1C). This result showed that the loss of cross-β stacking is consistent with the loss of ThT fluorescence, which correlates with an increase of mobile Tau species, previously associated with soluble Tau[67].

We further showed that the soluble Tau fraction that increased upon heparinase treatment of the fibrils, were mainly composed of monomeric Tau. Both heparin and RNA fibrils before and after digestion were analyzed using blue native polyacrylamide gel electrophoresis (BNPAGE). The results showed that Tau fibrils after digestion release a significant amount of solubilized monomers, in contrast to Tau fibrils before digestion, where no monomeric Tau could be observed (Figure 3.1D).

These results, together, show that the decrease of ThT fluorescence observed upon addition of the cofactor enzyme (Figure 3.1A) reflects a depolymerization of Tau fibrils, which releases soluble Tau monomers, as determined by an increase of mobile component in CW

EPR spectra (Figure 3.1B, C) and by the appearance of a Tau monomer band on a BNPAGE gel (Figure 3.1D).

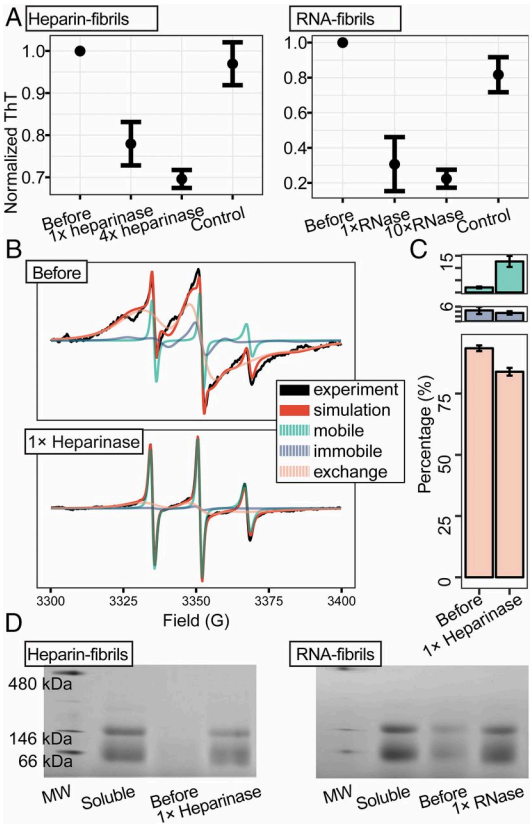


Figure 3.1: Fibrils depolymerize upon cofactor digestion. (A) ThT fluorescence of heparin-induced Tau fibrils and RNA-induced Tau fibrils before and after incubation with different concentrations of heparinase/RNase. ThT fluorescence value measured before enzyme addition was used as the normalization value. (B) cw-EPR spectra (black) of Tau heparin fibrils spin-labeled at site 322, before and after incubation with heparinase. The simulation spectra (red) are composed of mobile (green), immobile (purple), and spin-exchange components (orange). (C) Population of each component extracted from cw-EPR spectrum analysis. (D) BNPAGE of heparin and RNA fibrils before and after incubation. Freshly prepared Tau monomers were loaded as reference (soluble). In all panels, 1× heparinase: 1 U enzyme per 1 μg heparin; 1× RNase: 2.5 μg/mL. Error bars show SD ($n \geq 3$).

3.3.3 Bound cofactors are required to stabilize Tau fibrils

We learned that digesting cofactors depolymerized Tau fibrils. However, extending the digestion for longer times (Figure S2) or increasing heparinase and RNase quantity by 4 times and 10 times respectively (Figure 3.1A) did not significantly decrease the remaining ThT fluorescence, which suggest that the maximal digestion had been reached and the remaining ThT fluorescence came from species that are not sensitive to heparinase/RNase digestion. These species can originate from either (i) fibril populations that are stable without cofactor, or (ii) fibril populations stabilized by cofactors that are undigestable due to steric inhibition. To answer this question, we quantified the amount of undigested cofactor remaining after enzyme addition.

Heparin fibrils were prepared using spin labeled heparin (heparin-SL), which can be quantified by CW EPR. The double integral of a CW EPR spectrum is directly proportional to the quantity of spin-label, from which the concentration of heparin-SL can be obtained. We first confirmed using a ThT assay that heparin-SL triggers Tau fibrillation (Figure S6). The mature fibrils were then subjected to heparinase treatment, and subsequently subjected to dialysis (12kDa cutoff) that allowed the digested heparin to be exchanged out through the membrane, while retaining the fibrils and associated, undigested, heparin (Figure 3.2A). The concentrations of heparin inside the dialysis tube before and after dialysis were determined by measuring the double integral of the CW EPR spectra. The change in heparin-SL quantity is plotted next to the ThT fluorescence measured for the same sample in Figure 3.2B. It shows that the addition of heparinase digests about 16% of the heparin, matching closely the 19% loss of ThT fluorescence observed upon heparinase incubation. Furthermore, lineshape analysis of the heparin-SL CW EPR spectra shows that (i) heparin is bound to fibril before digestion, (ii) digested heparin detaches from the fibril before being dialyzed out, and (iii)

undigested heparin remaining after dialysis is bound to Tau fibrils (see supplementary Figure S5 and discussion).

Using a similar approach of cofactor quantification, we measured RNA concentration before and after RNase digestion using UV absorbance at 260 nm. RNA fibrils were filtered before and after aggregation using a centrifugal filter (MWCO 10 kDa; Amicon), to let digested RNA flow through and retained associated, undigested, RNA. The concentration of uric acid in the flow-through was determined from absorbance at wavelength 260 nm, and the amount of RNA remaining with the fibrils was calculated by subtracting this value from the total amount. Figure 3.2B shows the calculated concentration of undigested RNA before and after addition of RNase, and when adding buffer (control). Next to these values are shown the level of ThT fluorescence measured for the same samples replotted from Figure 3.1A. The fluorescence and RNA concentrations measured before digestion was used for normalization, showing a good agreement between the 60% of digested RNA and the 68% loss of ThT fluorescence.

These quantitative agreements between ThT fluorescence and undigested cofactor, along with the CW EPR lineshape analysis of heparin-SL, strongly suggest that all fibrils remaining after enzymatic digestion are stabilized by cofactors that are still associated with the fibrils and could not be digested (the second of the two posited scenarios). We hypothesize that RNA/heparin molecules were not completely digested due to steric hindrance from the fibrils that prevents the enzymatic cleavage of their target molecules.

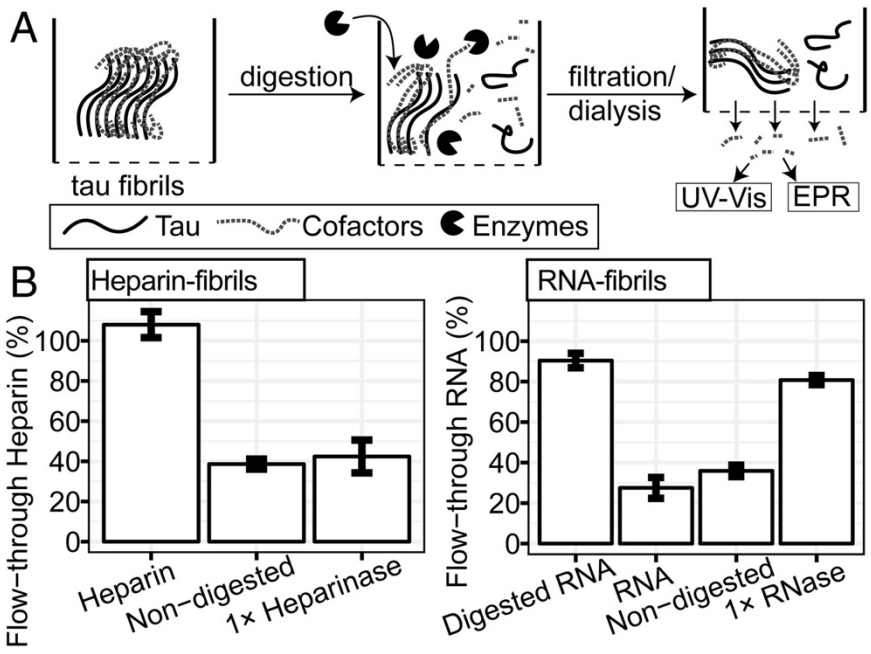


Figure 3.2: Cofactors are bound to fibrils and partially digested. (A) Schematic diagram of experimental procedures. Digested soluble cofactors were separated from Tau fibrils via dialysis (RNA fibrils) or filtration (heparin fibrils) before measuring their concentration in the flow-through by cw-EPR (heparin) or UV absorbance (RNA). (B, Left) The concentration of spin-labeled heparin in the flow-through of a 0.2-μm filter for heparin and heparin fibrils before and after digestion. The concentration is given as the percentage of the concentration before filtration. (B, Right) RNA concentration outside a dialysis bag (12 kDa molecular mass cutoff) at equilibrium was measured and the percentage of RNA that flowed through is given on the y axis. Error bars show SD of three independent repeats.

3.3.4 The depolymerized Tau monomers have no "memory" of the fibril state

Tau undergoes drastic conformational rearrangements upon aggregation. We tested whether the released monomers after digestion recovered their original properties by measuring their conformations, their capacity to reform fiber, as well as their capacity to seed aggregation.

We first tested whether or not released Tau monomers were able to re-fibrillate. After heparin digestion, we inactivated the heparinase by heating up the sample at 65 °C for 15 min. We then added fresh heparin at a molar ratio Tau:heparin 1:4 to both digested and non-digested

heparin-fibrils (labelled “pre-digested fibrils” and “Non-digested fibrils”, respectively, in Figure 3.3A), and subsequently recorded the ThT fluorescence. The fluorescence of the digested sample was, as expected, initially lower than the non-digested sample due to fibril depolymerization. Upon addition of heparin in non-digested and digested fibrils, the fluorescence of both samples reached the same level, showing that the monomers released upon heparin digestion is able to reform fibril when new heparin is provided. Similarly, RNase was inhibited by adding RNase inhibitor, then fresh RNA was added to both digested and non-digested RNA fibrils while monitoring ThT fluorescence (Figure 3.3A). The fluorescence of the digested sample was initially lower than the non-digested sample—note that this difference is smaller than in Figure 1A because 0.02× RNase was added in order to reach complete inhibition by adding a reasonable amount of inhibitor. After the addition of new RNA, both samples recovered comparable fluorescence levels, confirming that released Tau monomer upon RNA digestion can reform fibril upon addition of fresh cofactor.

We have previously identified a signature of aggregation-prone Tau conformations represented by a dramatic distance extension around the amyloidogenic hexa-peptide PHF6 (306VQIVYK311) and PHF6* (275VQIINK280)[64]. Following a similar procedure as used by Eschmann et al.[64], we measured the distribution of intra-Tau spin label distances spanning residues 272 and 285 by double electron electron resonance spectroscopy (DEER). We confirmed the expected distance extension around PHF6* transitioning from monomer to heparin fibrils (Figure 3.3B). The release of Tau monomers upon heparinase treatment resulted in the partial reversal of the PHF6* distance extension (Figure 3.3B), showing that the released soluble Tau lost their aggregation-prone conformational signature. The distance distribution of

the digested sample is broader than that of undigested sample, indicating a heterogeneous sample that likely originates from the cohabitation of both fibrils and released monomers.

It was recently proposed that after aggregation, Tau, even in its monomeric form, maintains aggregate-templated conformations that provide high seeding capacities[87,88]. This hypothesis, if true, would lead to a much higher aggregation propensity from the digested fibrils as each released monomer would represent an aggregation nucleus, drastically increasing the number of nucleation sites in the sample. We tested this hypothesis by measuring the seeding propensity of heparin and RNA fibrils before and after cofactor digestion, both in vitro and in vivo. We used a cellular seeding assay similar to previously established assays[25,54]. This cellular assay is based on Tau187 overexpressed by H4 neuroglioma cells, not K18 (fragment 244-373) in HEK cells, as in the assays established in the literature. When the overexpressed proteins aggregate, they form fluorescent puncta[54] due to the fluorescent protein mCerulean tethered to Tau. As shown in Figure S7, heparin and RNA fibrils showed poor seeding capacities (Figure S7C and E, respectively), which are not significantly improved by fibril digestion (Figure S7D and F). In contrast, mouse-extracted fibrils trigger significant aggregation (Figure S7A). In addition, we tested the seeding capacity of heparin fibrils in an in vitro assay where the premade fibrils were added to fresh monomers, in the presence or absence of a cofactor (see next section for precise assay description). Figure 3.4A shows that digested fibers have lower seeding capacity compared to intact fibers, as demonstrated with a smaller increase in ThT fluorescence. This result is in good agreement with the notion that depolymerized monomers cannot seed aggregation, and that the loss of seeding capability is due to the loss fibrils after cofactor digestion. The monomers derived from fibrils have no "memory" of the fibril state, neither in conformation, nor in property.

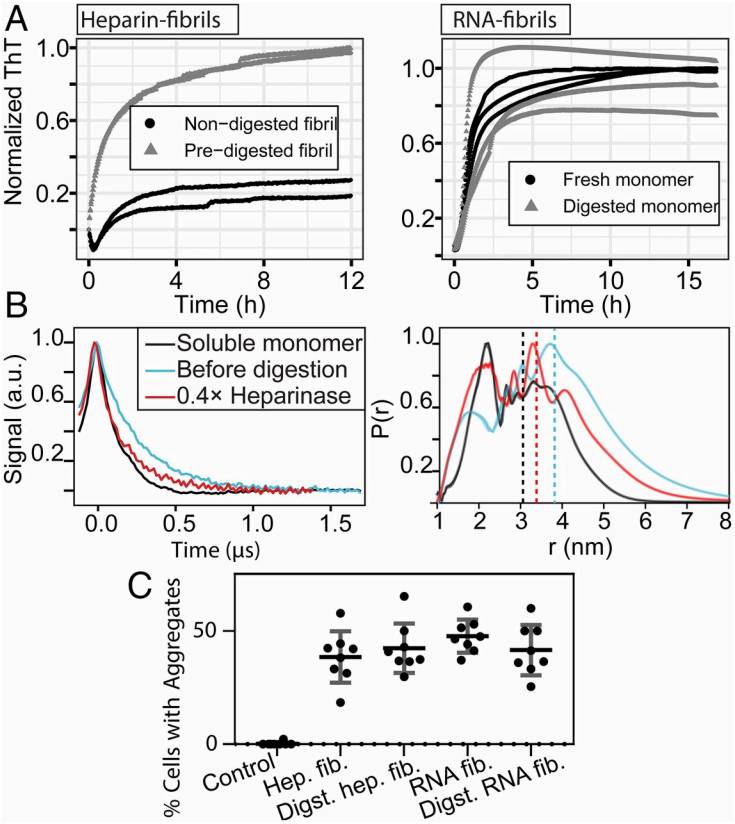


Figure 3.3: Re-fibrillization and conformation recovery of depolymerized Tau monomers. (A, Left) ThT fluorescence of non/predigested heparin fibrils after heparin addition at time t = 0 h. (A, Right) Aggregation of monomers purified from digested RNA fibrils compared with fresh monomers. Curves of the same color show independent repeats. (B) DEER time-domain signal (Left) and corresponding distance distributions (Right), extracted with the SVD method, of Tau labeled at residues 272 and 285 before and after digestion, compared with soluble monomer. (C) Quantification of in vivo seeding experiments. The ratios of cells that contain fluorescent puncta are reported. Digst, digested; fib, fibrils; Hep, heparin.

3.3.5 The presence of cofactor sustains fibrils seeding

To test the role of cofactors in aggregation seeding/spreading, we performed in vitro seeding assays over two generations with and without cofactors, using either mouse-derived fibrils or heparin fibrils as seeds. Tau187C291S, (referred to as TauWT), which lacks the aggregation-promoting P301L mutation was used for these tests. We made this choice because this Tau variant is incapable of spontaneous fibril formation with RNA (see Figure S8). 5%

(mass) of heparin fibrils were added as seeds to fresh TauWT in the presence or absence of an RNA cofactor (polyU), while ThT fluorescence was measured (Figure 3.4A). The presence of the RNA cofactor significantly increased the seeding capacity of both digested and non-digested heparin fibrils seeds. In the second-generation experiment, 10% of the end product of the first generation was added to fresh monomers, with or without cofactor (Figure 3.4B). In line with the first generation, the presence of cofactor enabled significant seeding activity. The same experiments were carried out using mouse-derived fibril (rTg4510 mice expressing 2N4R-P301L) seeds (Figure 3.4C, D). The presence of cofactor did not change the seeding capacity of mouse-derived fibrils in the first generation (Figure 3.4C). During the second generation seeding, where 10% end product of the first generation was used as the new seed, aggregation was observed only when a cofactor was supplied (Figure 3.4D). The aggregation kinetic appeared to be slower compared to the other seeding experiment as the ThT signal had not plateaued after 10 h. Altogether, the results presented in Figure 3.4 showed that cofactors promote seeding activity, and suggest that an unknown cofactor species is present in mouse-derived material, which allow first generation seeding without providing additional cofactor, while the second generation fibrils are incapable of seeding, unless cofactors are added externally.

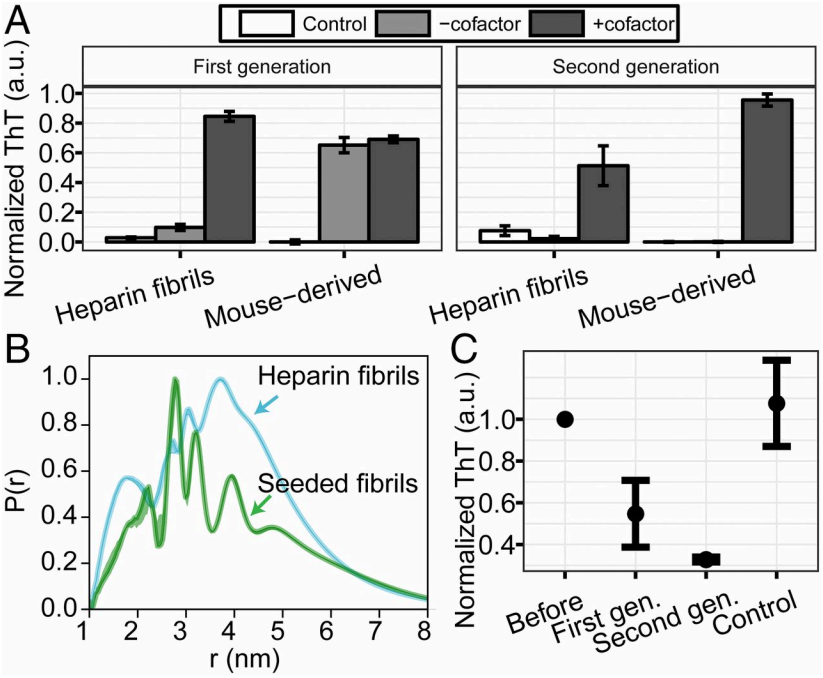


Figure 3.4: *In vitro* seeding of Tau fibrils. (A) ThT fluorescence of recombinant Tau187C291S seeded with heparin fibrils or mouse-derived fibrils (mouse-derived) in the presence or absence of cofactor (polyU RNA). In first-generation seeding (Left), 5% of pre-aggregated fibrils were added to fresh recombinant Tau and incubated at 37 °C while shaking. In the second generation (Right), 10% of the end products of first-generation samples were added to fresh monomers. Control refers to the seeds with cofactor incubated without fresh Tau. Error bars shows SD (n = 3). (B) DEER distance distribution between residues 272 and 285 of heparin fibrils and mouse-derived seeded fibrils (seeded fibrils) obtained from the SVD method. (C) ThT fluorescence before and after RNase addition to the first and second generations of mouse-derived seeded fibrils. Error bars shows SD (n ≥ 3). gen, generation.

3.4 Discussion

We have demonstrated that mature recombinant Tau fibrils require cofactors to be stable. Upon cleavage of cofactors through enzymatic digestion, the fibrils depolymerize, releasing soluble monomeric Tau. We showed that 16% of heparin and about 60% of RNA can be digested after aggregation (Figure 3.2), which matched closely the loss of cross-β sheet structure in the fibrils, as reflected from the ThT fluorescence and CW EPR spectra (Figure 3.2). The incomplete digestion of the target molecule (heparin or RNA) is likely due to

allosteric hindrance that prevents the enzyme to access and cleave its targets. The finding that in both RNA and heparin fibrils a defined and reproducible amount of cofactor can be digested suggests that the fibrils are polymorphic, where in some type of aggregates the cofactors are protected from digestion, while in others they are accessible to the enzyme. This is in good agreement with previous finding that heparin-induced fibrils are structurally heterogeneous[89]. Even single point mutations change the extent to which ThT fluorescence is lost upon cofactor digestion (Figure S3), suggesting that the mutations affect the fibril molecular structure and the accessibility of the cofactors. In addition, the significant difference seen in Figure 3.1A between the maximum digestible RNA (60-70%) and heparin (~20%) suggests that, although both aggregates are ThT active and filamentous, they have different molecular arrangements with respect to the different cofactor locations, likely because they generate different polymorphic populations. The nature of the enzymes likely contributes also to this difference in the extent of digestion, e.g where RNase A (14 kDa) would be less sensitive to steric hindrance than heparinase I (43 kDa) to process its target.

Previous work have shown that heparin-induced fibrils are more stable to Gdn-HCl denaturation than AD brain-extracted fibrils[56,90]. We show that upon removal of heparin, the fibrils are unstable and depolymerize spontaneously at room temperature. Thus, it demonstrates that the high stability of heparin-induced fibrils observed in ([56,90]) is largely due to Tau-cofactor interactions, and not due to Tau amyloid packing. This highlights the need for clarification whether the measured properties of Tau fibrils (stability, structure and seeding capacity) reflect a pure Tau aggregate or a hybrid Tau-cofactor assembly.

The finding that a polyelectrolyte cofactor is essential for the stability of fibrils prepared in vitro has two major potential implications with respect to in vivo aggregates, as

discussed below. (i) There are unknown cofactor(s) in the fibrils formed in neurons, whose roles have then been largely underestimated, and (ii) polyanion-induced recombinant Tau fibrils make limited models (of in vivo aggregates), partially owing to the strong effect of cofactors in the mature fibrils that, to date, have been mostly ignored.

We have recently shown that heparin-induced fibrils are structurally different from AD fibrils and are highly heterogeneous[91]. The structural heterogeneity combined with the poor stability without a cofactor observed here suggest that heparin-induced fibrils are in fact mispacked, i.e. they possess imperfect cross-β interactions (dense intermolecular hydrogen bond network) that are responsible for the high stability of amyloid aggregates. This lack of defined structure in the model fibrils would be compatible with a poor seeding/spreading capacity, observed here for cellular assays (Figure S3.7) and elsewhere[56,92], as the fibril would represent a poor structural template to pile up more surrounding monomers. While there is no reason why this mispacking should not happen in vivo, there is a hard selection process produced by the proteasome through its constantly clearing of, and protecting against protein aggregates. Therefore, it is likely that only the most resistant or seeding-potent fibrils survived. These might be the most perfectly packed species, which are then extracted in the brain homogenates and measured by cryo-EM[38] and other methods. In light of this analysis, one way of improving in vitro fibril models would be to aggregate in rough conditions, such as high temperature with shearing constrains or/and mild denaturing compounds, in order to promote highly stable fibril isoforms and reduce their dependence on cofactors.

We have performed in vitro seeding of recombinant protein with either heparin-induced fibrils or mouse-extracted aggregates. Under our conditions, without providing additional cofactor, no significant seeding was achieved with heparin-made seeds, whereas seeding was

observed with brain-extracted seeds (Figure 3.4A,C), in agreement with previous work[56,92]. In the latter case, when aggregates extracted from this successful seeding experiment were used as seeds for a second generation seeding, no fibrils were formed (Figure 3.4D), showing a limited propagation of mouse-extracted fibril through seeding. However, when a cofactor (polyU) was provided in the reaction buffer, heparin-induced seeds (Figure 3.4A,B) and mouse-derive seeds could be made competent over multiple generations. This result is in good agreement with the hypothesis that an excess of an unknown cofactor is present in the mouse-extracted aggregates (but not in the heparin-induced fibrils) that allows a first seeding generation, but not subsequent generations, as the cofactor present in the original seeds is consumed. In contrast, when a cofactor is provided in the buffer, seeding can be sustainably carried out over successive generations, as previously reported[80].

The aggregation kinetics when using mouse-derived seeds is significantly different between the rapid aggregation of the first generation, where the "mystery" cofactor originating from the brain-extracted seed would be assisting seeding, and the slow kinetics of the second generation, where the provided cofactor (polyU) starts assisting aggregation. This suggests that the nature of the cofactor influences the aggregation mechanisms/pathways, resulting in different kinetics. This hypothesis could also explain why in the particular cellular assay carried out in this work (Figure S3.7), the *in vitro* fibrils have poor seeding capacity. Indeed, if the couple fibrils-cofactor complex (as opposed to fibrils alone) is what determines the aggregation properties, it might be that the particular H4 cells used in this assay do not provide the cofactor compatible with an efficient seeding of the polyanion-induced fibrils conformers. There is a large variety of bioelectrolytes that could interact with Tau in vivo, including DNA, RNA, glycosaminoglycans (GAG), and ATP. Although Tau is mostly present in axons[93], it is

also found in neuronal nuclei[94] and is suspected to traffic in the extracellular matrix[95]. The complexity of the cellular environment and trafficking make it very hard to unambiguously assess in situ the role of a given cofactor in Tau aggregation and the pathogenicity. GAG, in particular heparan sulfate (HS), have been the most studied interaction partners to Tau (see for instance recent review[96]), likely because (i) it was found co-localized with Tau neurofibrillary tangles (NFTs) in AD brain[53], (ii) heparin is a very efficient cofactor to promote aggregation in vitro[53] and (iii) HS have been shown to play an important role in Tau internalization[97]. Although it is unclear when and how HS interact with Tau, as the former is exclusively present on the extracellular cell surface while the latter is mainly found in the cytoplasm, the fact that the NFTs co-localize with HS[53] suggests that HS might be incorporated into in vivo formed Tau fibrils. Similarly, RNA has also been found to specifically associate with Tau in neurons[98], to be sequestered in Tau pathological assemblies, not only in AD brains, but also in Pick bodies[99], suggesting that RNA might also be part of the final Tau fibrils. In the present study, we highlight the potential role of cofactors in the mature fibril properties, and in particular their stability. Thus, we put forward the hypothesis that identifying and targeting the fibril cofactors might provide an opportunity to destabilize Tau aggregates and clear the NFT burden from the brain.

Tau have been shown to form different aggregate polymorphs, referred to as strains that have different spreading patterns and toxicity, and are disease specific[27,54]. Strikingly, recent advances in cryo-EM has allowed the identification of two very different aggregate structures in AD brains[100] and Pick's disease[101], suggesting that different strains have very distinct atomistic fibril structures. If polyelectrolyte cofactors were present in the mature Tau aggregates, they would not necessarily be visible on cryo-EM electron density map as their

high flexibility would likely compromise their resolution. The finding that Tau can adopt different aggregate structures raises the question of the molecular origins of this differentiation toward a disease-specific structure. We propose that interactions with different cofactors, potentially modulated by pathology homeostasis, might dictate the choice of a given structure and strain. Following this hypothesis, the identification and detection of cofactors in Tau amyloid deposition might be a viable approach to differentiate between different Tauopathies.

3.5 Conclusion

This study highlighted the integral nature of cofactors to recombinant fibrils of Tau. Before this study, there had been multiple reports of seeding competent forms of Tau, but it became clear as we studied Tau, that the seeding being reported was verry sensitive to the conditions used. In fact, all the studies that had previously reported seed-competent material were using shorter forms of Tau (often K18) and/or the protein had been exposed, if temporarily, to additional cofactors. Tau was exposed to cofactors *in vitro* and the cofactors were removed, or the products were passaged through a cell line and the endogenous Tau was aggregated before purification from cell lysates[85,102]. The demonstration of the integral nature of cofactors was a call for more work to be done to produce pure, homogenous fibrils populations. We hypothesized that to form such fibrils, required the degrees of freedom in the IDP to be reduced, and investigated this hypothesis in Chapters 4, 5, and 6.

Chapter 4: Proteolytic cleavage products of Tau are Aggregation prone, and form a toxic seed

4.1 Introduction

In collaboration with Judith Steen (Harvard Medical School) we investigated the role of proteolytic cleavage to Tau's aggregation. It has long been suggested that Tau undergoes cleavage in Tauopathies. In particular, caspases have been identified to cleave Tau[103] but it is not clear if the cleavage events occur prior to, or after, aggregation has occurred[104–106]. The Steen group used Mass Spectrometry to identify common cleavage sites that correlate with AD patient pathology and seeding competence of Tau extracts (see Dujardin *et. al*[107] for associated methods). Many of the cleavage of the specific sites correlated with activation of multiple inflammation pathways, and activation of the immunoproteasome (IP). Additionally, the cleavage sites that correlated with IP activity matched very closely with the core region resolved in AD filaments.

The immunoproteasome shares the catalytic subunit of the 20S proteasome, and a 19S subunit, but the catalytic β1, β2, and β5 subunits are replaced by inducible subunits LMP2, MECL-1 and LMP7 respectively[108]. The LMP2, MECL-1 and LMP7 are replaced in the IP through induction of genes *PSMB9*, *PSMB10*, and *PSMB8* respectively by interferon-γ (INF-γ) and other cytokines[108,109]. Previously, the 20S proteasome was identified to effectively degrade Tau in the MTBD[110], but the cleavage sites associated with the IP were unique, and did not degrade the MTBD as frequently. The mass spec data indicated that patients with more Tau, and more toxic Tau species correlated with a higher presence of PSMB8/9. As is detailed in the following sections, we studied the Tau fragments produced and found them to be aggregation-prone. We used the sites identified by mass-spec to create new, recombinant forms

of Tau composed mainly of a region of the MTBD and a small portion of the C-terminus between residues 297 and 396.

4.2 Results

4.2.1 Tau fragments identified in AD subjects can fibrillize *in vitro* and seed aggregation in cells

A rational scoring scheme was developed to prioritize Tau cleavage sites for further evaluation based on specificity to seed-competent Tau, correlation with IP subunit abundance, and *in vitro* cleavage by purified IP (Table S2). Interestingly, four of the top N-terminal side cleavage sites were located immediately prior to the $_{306}$VQIVYK$_{311}$ amyloid motif in the R3 domain, exposure of which was previously shown to promote Tau self-assembly[111]. For the C-terminal side, the two most prominent sites were selected. We assign the following nomenclature for clarity: N-terminal side cleavages (**1**) N296-I297 (**2**) G303-G304, (**3**) G304-S305, (**4**) S305-V306 and C-terminal side proteolytic cleavages (**a**) N368-K369 and (**b**) Y394-K395. Combination of the prioritized cleavage sites resulted in 8 unique Tau fragments **1a-4a** and **1b-4b** that, interestingly, span the AD Tau filament core (a.a. 306-378) resolved by cryo-EM[100] (Fig 3A), although the **a** series lack the most C-terminal 10 residues of the core structure.

We produced recombinant peptides representing Tau fragments **1a-4a and 1b-4b** for characterization of fibrillization and seeding properties. Due to the size of these peptides, and the apparent aggregation propensity initial purification trials showed, we cloned a chimera of a hexahistidine tag followed by eGFP, followed by a tobacco etch virus (TEV) enzymatic cleavage site attached to the desired fragment of Tau. The protein was then purified by Nickel-NTA chromatography. The TEV site was cleaved by TEV enzyme produced in house, and the desired Tau fragment was purified from the eGFP and TEV with cation exchange chromatography. These fragments were first tested for their ability to form fibrils *in vitro* in

the presence of heparin using the Thioflavin T (ThT) fluorescence assay. All fragments were capable of forming fibrils, with the **b**-series fragments, containing the longer C-terminus, all showing higher fibrillization propensity then any of the **a**-series fragments. N-terminal side cleavage closer to the hexapeptide increased fibrillization for **b** series peptides, with a rank order of **1b**=**2b**<**3b**<**4b** (Fig 4.1A). TEM with negative staining showed that all fibrils formed by these fragment peptides display characteristic PHF-like structures (Fig 4.2).

Next, we tested the ability of the fragment fibrils to seed aggregation in mCerulean-TauRD cells. The ranked seeding activity of peptide fibrils based on % of cells exhibiting distinct puncta showed excellent agreement with the *in vitro* ThT assay-based measurement of their aggregation propensity: fibrils formed by **b**-peptides with the longer C-terminus were more effective at seeding (Fig 4.1C).

We then tested whether the fibrilized Tau fragments were capable of seeding naive Tau without heparin, using **2b** as a representative peptide (Table S2). A catalytic amount (1:40 molar ratio) of **2b** fibril was capable of inducing the formation of 2nd generation fibrils (Figure 4.1E) of both **2b**. This effect was not attributable to remaining heparin from the 1st generation fibrillization, because an equimolar effective concentration (160:1) of heparin as contained in the seeds did not produce fibrils (not shown). Surprisingly, 2b fibril could also seed monomeric Tau187 (a.a. 255-441), while Tau187 fibril seed could not produce any appreciable amounts of 2nd generation fibrils (Figure 4.1E). When normalized by the total mass of the seed (or heparin cofactor), **2b** seeds were found to be the most effective at generating fibrils, even more effective than Tau seeded with the standard proportion of heparin(4:1) (Fig 4.1E). These data demonstrate that specific Tau fragments such as **2b** can self-aggregate, as well as recruit longer, naïve, monomeric Tau in the seeding process.

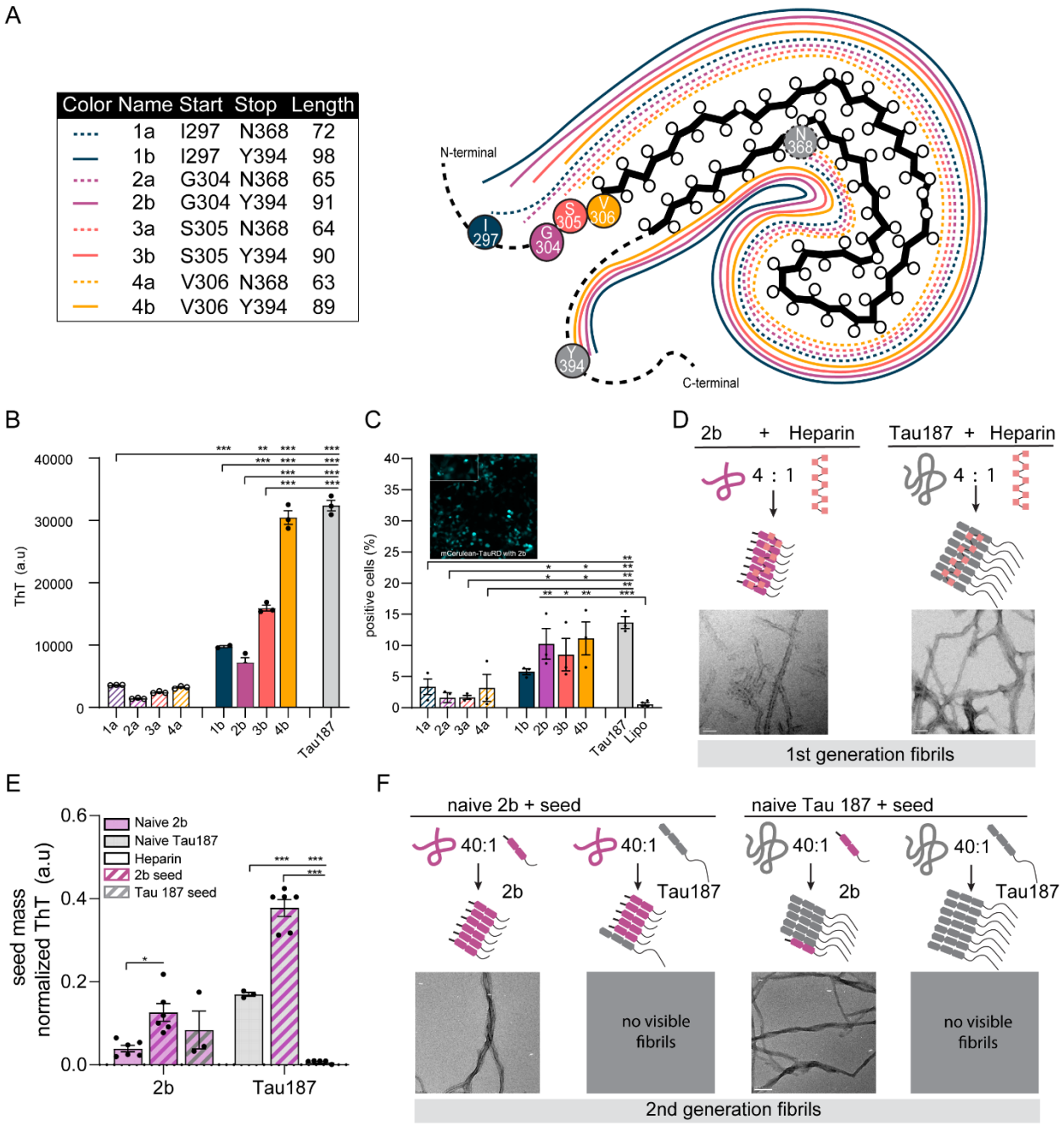


Figure 4.1. Recombinant Tau peptides defined by cleavages in post-mortem AD tissue can fibrilize *in vitro* and resultant fibrils can seed aggregation in HEK293T cells expressing mCerulean-TauRD. A) Putative core cleavage fragments mapped to the cryo-EM structure of the protofilament, showing that cleavage at N368 falls within the core structure. B) In vitro ThT aggregation assay of fragment peptides. Aggregation shown as maximum ThT after 15 hours. Fibrils were confirmed by TEM. Mean ±SEM. One-way ANOVA ($F(8, 17) = 428.0$, $p < 0.001$); stars indicate Tukey post-hoc test p-values ** $p<0.01$, *** $p<0.001$ C) Preformed fibrils of Tau fragments trigger aggregation in HEK293T cells expressing mCerulean-TauRD. Seeding assays show varying degrees of efficacy for each fragment shown as % seeded cells. It is noted,

peptides cleaved less at the C-terminal side (b type) show higher potency of fibrilization and seeding over the other (a type). Mean ±SEM. One-way ANOVA ($F(9, 13) = 9.765$, $p < 0.001$); stars indicate Tukey post-hoc test p-values * $p<0.05$, ** $p<0.01$, *** $p<0.001$ D) Schematic of *in vitro* seeding tests showing 1st generation fibrils were formed by addition of heparin to cause aggregation. E) Second generation seeding of naïve Tau fragment 2b or Tau187 with 2b or Tau187 seeds (40:1 Tau to seed), compared to seeding with heparin (4:1 Tau to heparin). ThT values were normalized by the mass of cofactor and protein used. Mean ±SEM. One-way ANOVA (2b group - $F(2, 12) = 4.954$, $p = 0.0270$, Tau187 group $F(2, 12) = 199.9$, $p < 0.001$); stars indicate Tukey post-hoc test p-values * $p<0.05$, *** $p<0.001$ F) Schematic of 2nd generation fibril seeding by 1st generation fibrils with TEM images of generated fibrils

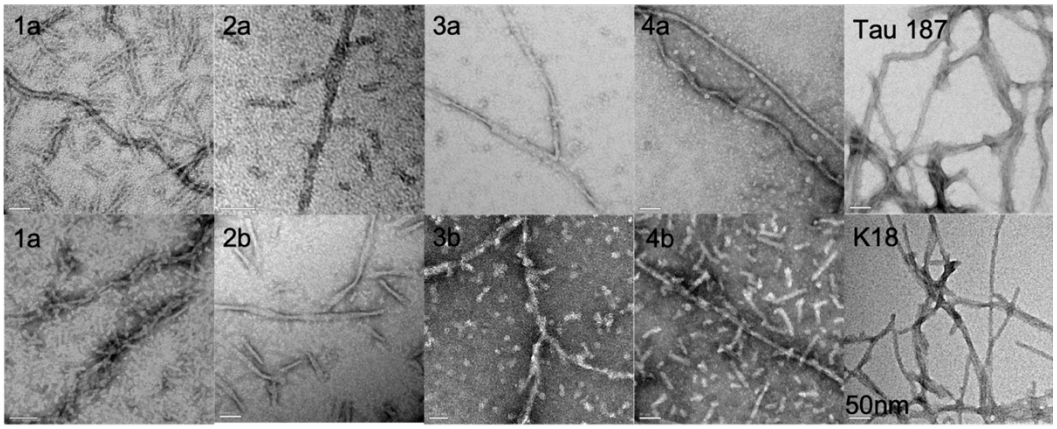


Figure 4.2: Negative stain TEM of 8 Tau fragments after addition of heparin. All eight fragments were capable of forming amyloid fibrils of varied morphology. Also shown: firbils formed from recombinant Tau187 or K18 with the addition of heparin.

4.3 Discussion

Understanding the vast diversity of Tau proteoforms, and how this landscape is altered by the cellular environment under disease conditions is key to the development of effective therapeutics for AD and other Tauopathies. Our characterization of Tau cleavage in a large cohort of AD patients combined with functional analysis provides comprehensive insight into the disease and establishes a convincing model where inflammatory cytokines and cellular stress result in the generation of Tau seeds. We identify a mechanism of IP-mediated Tau processing in the generation of aggregation-prone fragments, directly linking neuroinflammation to initiation of Tau pathology in AD. Through this mechanism, AD risk

genes expressed in cytokine-releasing microglia and macrophages[112,113], and inflammation-associated AD modifiable risk factors[114] such as smoking, diabetes, TBI (Traumatic Brain Injury), obesity, could increase formation of Tau seeds. Moreover, oxidative stress, exposure to protein aggregates, and proposed viral/bacterial risk factors can all increase IP expression in neurons. Neuronal exposure to amyloid beta also increases IP expression[115], which could in part mediate the observed synergy between amyloid-β and Tau pathology[116]. This model represents an early step in the initiation of Tau pathology and provides multiple potential treatment opportunities, such as modulation of neuroinflammation, cleaved Tau neoepitope antibodies, and IP inhibition. Indeed, inhibitors for PSMB9/6 have demonstrated efficacy in AD mouse models, although not directly in models of Tauopathy[117]. Tau PTMs such as phosphorylation have also been implicated in inflammation-mediated Tau aggregation in mouse models[118]. Given the extensive posttranslational modification of Tau[110], it is highly likely that PTMs can affect cleavage of Tau that need to be characterized in future studies. Lastly, the data presented in this manuscript has relevance for ongoing Tau antibody development projects, as it emphasizes the importance of epitope selection in light of the extensive cleavage observed in seed-competent Tau.

4.4 Conclusion

This work contributed new understanding to how proteolysis may cause Tau to become more aggregation prone. Fragments of Tau that contained the majority of the core region of Tauopathy fibrils aggregated readily. When seeds of the fragment fibrils were added to longer forms of Tau, the full-length protein was recruited into fibrils while maintaining a morphologically similar fibril structure to the initial seeds. The identified fragments present a

toxic species of that may be responsible for initiating the aggregation of Tau and its spread throughout the brain.

The Steen group hypothesized that inflammation was causing an increase in IP-based digestion of Tau, generating these toxic fragments and initiating the aggregation cascade of Tau across the brain. Indeed, when cytokines were introduced to cells, the same cleavage patterns were observed in Tau as are observed in AD patients. The Cleavages of Tau would precede detectable aggregation, but amore agregation was observed in cells treated with cytokines. The lysate of these treated cells were purified, and imaged with NS-EM and immunogold labeled TEM (see supplement) and were found to also contain amyloid-like aggregates.

It must be noted that concurrent work by Al-Hilaly, Lovestam, Wischik, Serpell, Scheres *et. al.* has come to very similar conclusions to work presented here. Wischik *et. al.* identified a pronase-resistant fragment of Tau in the core of PHF fibrils spanning residues 297-391 a fragment of Tau referred to as dGAE[119] that was produced by Al-Hilaly *in vitro* and found to aggregate without additional cofactor[120,121]. In contrast the above work, dGAE was not identified to exist at high frequency in AD patients as a natural cleavage. However, the ability to form fibrils without additional cofactor is novel and was not observed in our work. Lövestam *et. al.* published their findings recently that recombinant dGAE can be aggregated into the exact same structures observed in AD without addition of any cofactor[122]. The Tauopathy replicas were not formed with full length forms of Tau, or other commonly used truncated forms of such as K18. The groundbreaking work by these groups gives support to our findings of truncated Tau fragments being very aggregation prone, and the cause of formation of the Tauopathy disease structures.

Chapter 5: Hydrophobically driven Aggregation of Tau Peptides

5.1 Abstract

Operating under the hypothesis that short fragments of Tau have fewer degrees of freedom and are more likely to aggregate into a single conformation, a peptide of Tau was studied to understand if it is possible to define its aggregation pathway, and to understand how that aggregation is occurring. Presented in Chapter 5 is work to define the aggregation pathway of this fragment, and the effect of a hydrophobic P301L mutation to the aggregation. This Chapter is being prepared for submission as a stand-alone paper. Chapter 6 describes ongoing work to solve the structure of the fibrils formed by the fragment. In order to compare test the hypothesis that short fragments form more homorgenous fibril structures, solving the atomistic structure of the fibrils will be crucial.

5.2 Introduction

Pathological Tau proteins are prion-like in that their misfolded shape defines the disease that can recruit naïve Tau to misfold and induce aggregation. However, the defining property of a Tau prion is unknown. This study presents a 19-residue Tau fragment that can stack to form neat paired helical filaments (PHFs). This potent Tau fragment contains the hydrophobic VQIVYK hexapeptide included in the core of every pathological Tau fibril structure solved to date and adopts hairpin (HP) conformations within the folds of corticobasal degeneration (CBD) and prefrontal supranuclear palsy (PSP).[45,123,124] This HP fragment is rendered more aggregation prone and seeding competent by a single-nucleotide mutation P301L. While P301L is by far the most ubiquitous disease mutations used to model Tauopathies in vitro, in cell, and in mouse models, the molecular basis of its effect is unknown.

We present our discovery that the P301L mutation induces ordering of hydration water around the mutation site that renders the HP fragments locally more hydrophobic, and so facilitates their assembly to longer, more stable fibrils compared to the wild type. We propose that hydrophobic templating and entropy-driven dehydration of key Tau segments are a defining event in prion-like propagation of pathological Tau aggregation.5.2 Introduction

The microtubule associated protein Tau is a pathological protein that is present in insoluble neurofibrillary tangles associated with Alzheimer's Disease (AD), frontotemporal lobar degeneration (FTLD-Tau) and a large family of other neurodegenerative diseases, collectively referred to as Tauopathies. Tau exists in solution as an intrinsically disordered protein (IDP) containing a large ensemble of disordered conformations[125]. In Tauopathies, Tau proteins fold and stack into insoluble amyloid fibrils such as the paired helical filaments (PHFs) observed in the neurofibrillary tangles (NFTs) of AD. Pathological Tau fibrils can move from cell to cell and induce the misfolding and aggregation of naïve Tau, hence earning the designation of a Tau "prion"[26,126,127] . Here, a prion is defined as a protein whose misfolded shape defines the disease, and can induce pathological shape changes in naïve prions upon contact and induce aggregation. The structure of the folded Tau within the NFTs and their location in the brain tissue is diagnostic of Tauopathies, yet the molecular properties and mechanism that induce or enhance prion-like propagation of Tau aggregation are unknown. The development of therapeutic or diagnostic strategies, e.g. small molecule drugs, antibodies or positron emission tomography imaging agents, requires cellular and mouse models that respond with spontaneous aggregation of its naïve Tau upon infection with pathological Tau prions, typically added in the form of insoluble, fibrillar Tau. Remarkably, all current Tauopathy models that are used for developing and testing therapeutic or diagnostic strategies

rely on disease mutations to achieve prion-like propagation of pathological Tau. In our quest to identify the minimal molecular requirements for generating a competent Tau prion, we focus on the role of the most ubiquitous Tauopathy disease mutation used in vitro, in cell and in mouse models, P301L, in modulating and mimicking pathological Tau aggregation.

Recent advances in Cryo-EM have enabled structure determination of the amyloid fibrils of Tau in a number of Tauopathies, including AD[100], Pick's Disease (PD)[128], chronic traumatic encephalopathy (CTE)[46], corticobasal degeneration (CBD)[45], prefrontal supranuclear palsy (PSP)[124], argyrophilic grain disease (AGD), aging-related Tau astrogliopathy (ARTAG), globular glial Tauopathy (GGT), familial British dementia (FBD) and familial Danish dementia (FDD)[42]. These structures are all from post-mortem tissue of Tauopathy patients, and do not include any mutations to Tau. Yet, seeded aggregation to propagate CBD, AD and other Tauopathy-like fibrils *in vitro* nearly exclusively relies on Tau substrates that contain at least the P301L/S mutations, often together with the V337M or other mutations[85,102,129].

There are 112 identified coding single-nucleotide polymorphisms (SNPs) to the *MAPT* gene in FTDP-17 diseases, with at least 41 of these known to be pathogenic.[106,130] Many of the SNPs lead to amino acid substitutions that correlate with faster disease progression, including P301L, P301S, R406W, ΔK280, and V337M[131]. Because of their increased fibrillation propensity, such Tau mutations provide a platform for viable detection of disease pathology within an experimentally accessible timeframe. Among them, the P301 site is mutated in the vast majority of models of Tau aggregation including mice (3xTg[132], hTau.P301S[133] JNPL3[134], PS19[135], PLB1-triple[136], TauP301L[137], among others), cell lines designed to assess seeding of aggregation[85,138], and *in vitro* studies of Tau aggregation. Strikingly, P301L is not a familial

mutation of AD, yet is used in the vast majority of mice and cell models of AD. It is clear the aggregation and seeded propagation of fibrils is enhanced by these mutations, but the difference in the aggregation mechanism imparted by the single P301L mutation is still unclear. Fundamentally, it is unknown whether P301L enhances the kinetics of aggregation through a more favorable transition state conformation and/or by altering the free energy of the soluble monomer or the final structured assembly, and hence increasing the thermodynamic driving force for Tau aggregation. In the latter case, the P301L mutation would change the quantity, and potentially the structure of the aggregated product, but this question has never been pursued. Given the current state of art of Tauopathy research, the study of the Tau prion mechanism requires understanding of the mechanism of this prominent P301L Tau mutation.

It is widely observed that the P301L mutation enhances the aggregation kinetics of Tau fibrilization and the primary nucleation rate compared WT or P301S[139]. Yet, it is still an open question whether there are molecular level differences between the wild type (WT) and P301L Tau in its intrinsically disordered monomer state. Studies have concluded that the major differences in solution-state properties lie in the transient, extended conformations of the population with enhanced aggregation propensity or long-range order, and suggested that P301L leads to a slight R_g compaction, correlated motions with the N-terminus, and/or increased exposure of the PHF6 hexapeptide motif by disrupting a local hairpin conformation[64,88,111,140–142]. Note that a slight R_g compaction and extended conformations are principally contradictory properties. These differences may be attributed to the different length scales probed in these reports. Some of these studies used crosslinking mass spectroscopy to observe a hairpin-like contact surrounding the 301L site of the WT Tau in solution that forms a protective barrier to aggregation.

In this study, we set out to investigate the role of P301L mutation on the conformational distribution and hydration state of Tau monomers in their IDP state, the Tau aggregation process and the fibrillar product. We designed a minimal Tau peptide that retained enhanced aggregation from P301L to focus on the local effect of P301L and to reduce the influence of long-range interactions enhanced by other factors.. The peptide was constructed to mimic a prominent hairpin (HP) motif formed in the core of the CBD fibril; residues 295-313, and henceforth referred to as the HP peptide (Figure 5.1A). This same region (residues 295 to 311) in WT Tau was reported by Chen *et. al.* to contain a protective hairpin motif that enhances the aggregation barrier in the solution-state. The question we ask is whether and by what mechanism P301L affects the aggregation kinetics of the HP Tau fragment, as well as fibril quantity, quality and stability.

5.3 Results

5.3.1 The PHF6 region is crucial for fibril formation

Six different variants and fragments of the HP Tau peptide were produced (Figure 1A,B). Two contained the entire segment (HP301P/L), two contained the N-terminal half (HP301P/LΔ(304-313)), and two contained the C-terminal half (HP301P/LΔ(295-299)). All peptides contained the 301P/L site. To assess the aggregation ability of each peptide, thioflavin T (ThT) assays were used to measure the increasing b-sheet content of the fibril formed of each peptide over 36 hours. To induce aggregation, multiple conditions were tested. Peptides HP301P and HP301L were incubated at 37˚C at multiple concentrations in order to test their ability to self-aggregate. HP301L was found to self-aggregate significantly more readily than HP301P at all concentrations. At concentrations below 500 μM, self-aggregation was not detected with HP301P by ThT fluorescence, whereas ThT fluorescence increase was observed at concentrations as low as 100 μM. Negative stain transmission electron microscopy (TEM)

of the self-aggregated HP301P/L protein, stained with uranyl acetate, showed highly tortuous fibrils (Figure S5.2).

To facilitate aggregation at lower peptide concentrations, heparin (Galen Labs, HEP001) was added to all 6 HP peptides at a peptide:heparin molar ratio of 4:1. Heparin is a widely used cofactor to induce aggregation of Tau. It has been shown in the literature, including our own study, that heparin-induced Tau fibril structures are typically distinct from that found in Tauopathies[71,89]. However, the objective of this study is to understand the effect of P301L on the overall energy landscape of the Tau peptide monomer and not to necessarily recapitulate exact disease structures. With the addition of heparin, all fragments containing the PHF6 VQIVYK motif (residues 306-311) formed fibrils with b-sheet content (Figure 1E). Neither of the N-terminal half constructs (HP301PΔ(304-313) and HP301LΔ(304-313)), both lacking the PHF6 motif, showed ThT fluorescence, confirming that the PHF6 motif is required for fibrillization of these peptides.

5.3.2 The N-terminal residues are impacted by P301L and enhance aggregation

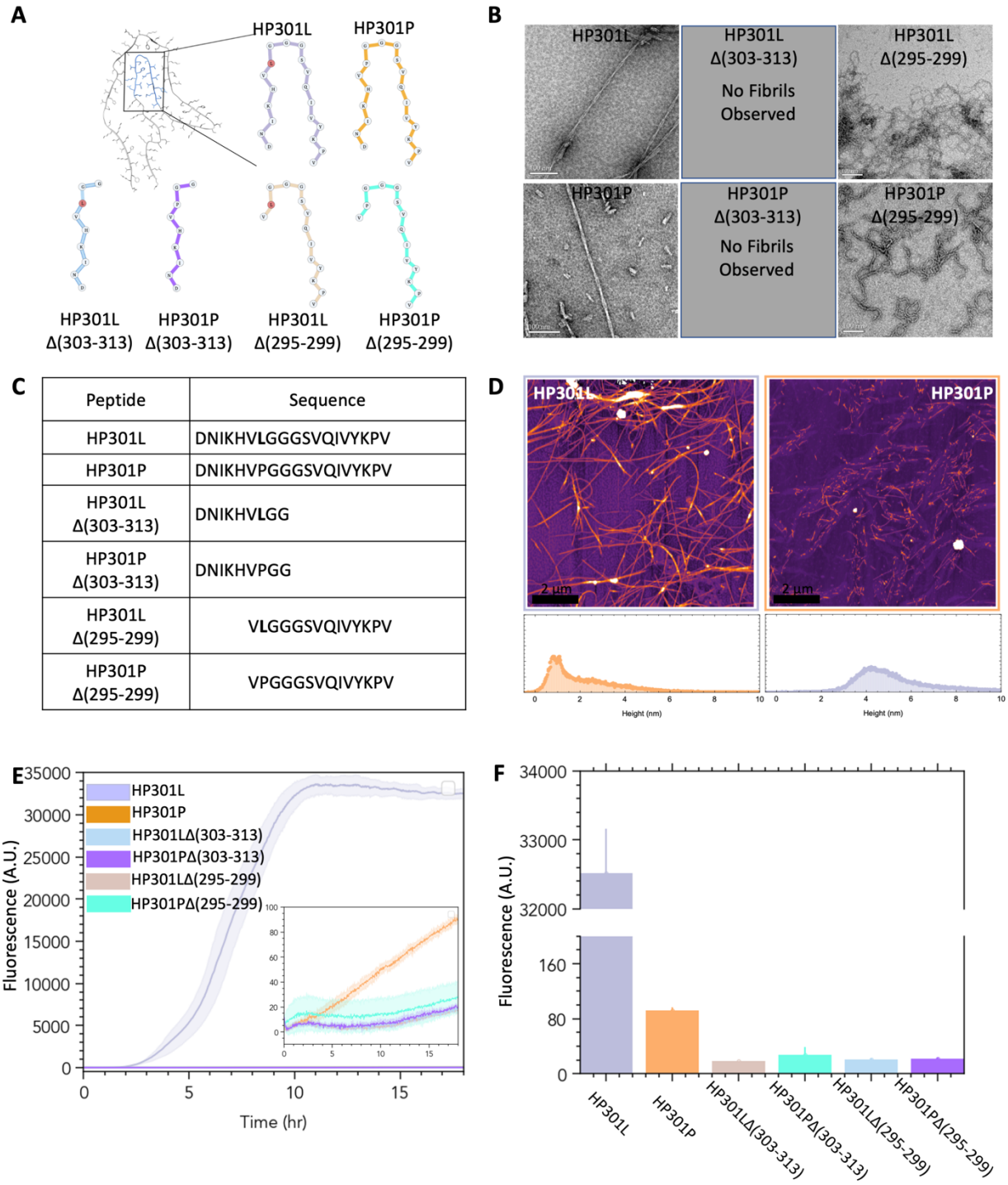
The fibrils made of full length HP301P and HP301L showed significant differences in β-sheet content as measured by ThT fluorescent intensity, while there was no significant difference in the total β-sheet content of the P301L mutant in the two C-terminal half peptides HP301PΔ(295-299) and HP301LΔ(295-299). The pronounced enhancement of aggregation by the P301L mutation, observed only in full length HP301P or HP301L compared to the two C-terminal half peptides is notable as the only difference is the additional N terminal segment 295-303 that does not itself show any inherent tendency to aggregate. This suggests that one effect of the P301L mutation is to alter the relationship between the aggregation inducing C-terminal segment 300-313 and N-terminal segment 295-303. The Cryo-EM structure of

pathological Tau fibril in CBD and PSP patients offer a potent hypothesis: a hairpin can be formed by the HP fragment, with a dry interface formed between the PHF6 region and the N-terminal residues 295-300. Potentially, the HP301L more readily folds into an aggregation-competent hairpin structure compared to HP301P, which would then manifest itself in differences in fibril morphology.

5.3.3 HP301P displays PHF-like morphologies

To qualitatively evaluate the amyloid fibrils formed, negative stain transmission electron microscopy (TEM) was first used to visualize the produced fibrils. Filamentous aggregates were observed for peptides HP301P, HP301L, HP301PΔ(295-299) and HP301LΔ(295-299) after incubation with heparin for 18 hours. The HP301P and HP301L constructs formed distinctly different fibril populations in that HP301L formed longer and more well-defined fibrils than all other peptide variants. Upon closer inspection, multiple fibril morphologies were observed (Figure S1), with the most prevalent one being paired helical filaments (PHFs) with a width of 40-80 Å and a crossover length of approximately 700Å (Figure 5.1B,D,F). Other morphologies observed of HP301L fibrils include straight, ribbon-like, and bundled filaments (Figure S5.1). Occasionally, a PHF-like periodicity was observed in the WT HP301P peptide (Figure 5.1B), but the frequency of such an observation was rare (Figure 1D). Aside from HP301L, no other peptide formed periodic twisted filaments with notable frequency. The PHF-like character of the filaments made of HP301L is reminiscent of the filaments observed in Tauopathies (Figure 5.1B). Notably, the fibrils of both of the shorter, C-terminal, peptides HP301P-C and HP301L-C formed fibrils exhibiting a shared morphology of narrow, highly tortuous fibrils (Figure 5.1B), indicating that the P301L mutation does not influence fibril morphology in the absence of the N-terminal half of the HP peptide.

While TEM provides a detailed depiction of individual fibrils, it offers at best qualitative population information. In order to obtain statistical information about the morphology of the fibril population, Atomic Force Microscopy (AFM) was used concurrently to negative stain TEM. Visually, the fibril morphologies appear very similar between the negative stain TEM and AFM images, with similar fibrils dimensions and periodicity observed by both imaging techniques (Figure S5.1). The height of the fibril population was measured and counted across an image field spanning 4 μm by 4 μm for both HP301P and HP301L, and the result presented in a histogram in Figure 5.1F. The HP301L fibrils exhibit significant differences to HP301P fibrils in terms of fibril quantity, thickness and structural features. The HP301L fibrils have on average greater heights, indicating that fibril structures with larger cross-sections are formed. Across all samples, HP301L also produce much longer continuous fibrils than HP301P, suggesting a more stable structure is formed, and pack in register to b-sheets that are less prone to fragmentation or termination. In particular, HP301L seems much more likely to form fibrils with a pair of protofilaments as observed in many Tau fibrils, consistent with the larger height of fibrils found (Figure 5.1F).



Peptide	Sequence
HP301L	DNIKHVLGGGSVQIVYKPV
HP301P	DNIKHVPGGGSVQIVYKPV
HP301L Δ(303-313)	DNIKHVLGG
HP301P Δ(303-313)	DNIKHVPGG
HP301L Δ(295-299)	VLGGGSVQIVYKPV
HP301P Δ(295-299)	VPGGGSVQIVYKPV

Figure 5.1: Fibril formation and stability of HP peptides: A) Peptides used in this study shown in the conformation observed in the CBD fibril structure (PDB #: GTJX). B) Negative stain TEM of all HP peptide fibrils formed with heparin. C) Sequence of each peptide illustrated in (A) D) AFM images of HP301P and HP301L fibrils. E) ThioflavinT fluorescence assay measures the total amount of b-sheet content after incubation with heparin at 37˚C for 18 hours. Samples contained 50µM protein with

12.5μM heparin (n=3). F) The distribution of fibrils heights from HP301P and HP301L AFM images. The HP301L population forms overall larger fibrils with greater heights than HP301P. G) ThT aggregation kinetics of HP peptides. Inset shows all peptides except HP301L.

5.3.4 Fibrils display different stabilities

To compare the stability of the final fibrils, the filaments were subjected to a denaturation assay using guanidinium hydrochloride (GdnHCl). Preformed fibrils were incubated with varying concentrations of GdnHCl ranging from 100 mM to 3 M, and the ThT fluorescence measured after equilibration. A decrease in fluorescence intensity was observed under denaturing conditions in a GdnHCl concentration-dependent manner with all peptide fibrils. A more stable fibril would resist denaturation under more aggressive conditions[143]. Hence, a more rapid decrease in ThT fluorescence intensity upon incubation with increasing GdnHCl concentration indicates a less stable fibril. All values are normalized to the fluorescence of the corresponding fibrils before denaturation (Figure 5.2A). At the highest tested GdnHCl concentration (3M), all peptides exhibited comparable extent of degradation, reflected in ThT fluorescence intensity at 21.7±2.5% of the original fluorescence before denaturation. AFM of HP301P and HP301L fibrils subjected to 3 M GdnHCl treatment show a monomer-like peptide film, indicating complete dissolution of the filaments. The AFM of HP peptide fibrils subjected to lower 1 M GdnHCl concentration show a lower total fibril density and a reduction in fibril length. The HP peptide fibrils subjected to 1 M GdnHCl also show differences in their stability: HP301L fibrils show a net loss of fluorescence of 32%, while HP301P fibrils show a greater loss of 61% of fluorescence under the same condition. This implies that HP301L produces more stable fibrils than HP301P. The fibrils of HP301PΔ(295-299) and HP301LΔ(295-299) have similar stabilities to each other, but are less stable compared to the fibril made of the corresponding longer peptide form. This result, again,

indicates that the P301L mutation stabilizes the full-length HP fibrils, but not the fibrils made of the C-terminal peptide lacking the N-termini, suggesting that the N-terminus serves as a stabilizing flanking region. If there was a distinct local effect originating directly from the 301 site due to mutation to leucine, for example from additional H-bonding to the amine group of leucine, one would expect the enhancement effect to be observed, again, in the shorter peptides that contain the P301L mutation. Because such an effect was not observed, we suggest that the major effect on the fibril structure and stability by the P301L mutation is caused by changes in protein conformation and folding involving the residues preceding the P301 site. Clearly the end state filaments are different, suggesting that the inherent solution-state property of HP301P and HP301L that make them destined to pack to different fibril structures must also be different.

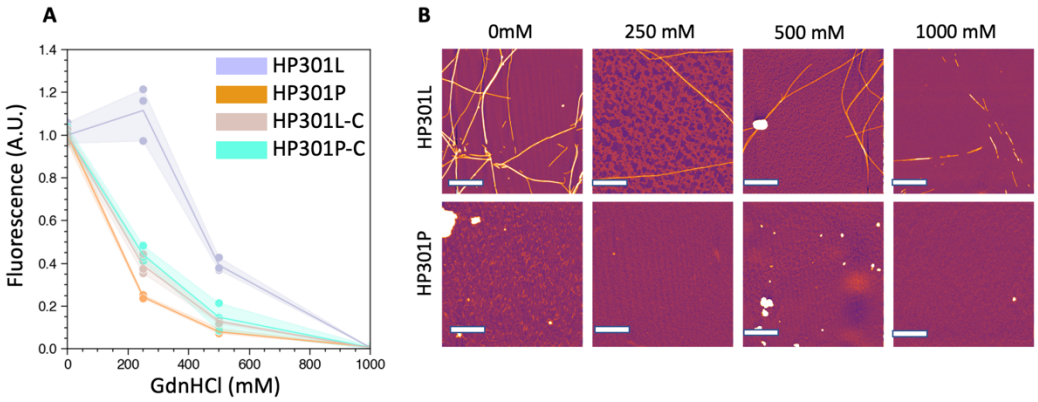


Figure 5.2: Stability of HP peptides: A) Thioflavin T fluorescence assay after incubation with guanidinium hydrochloride and variable concentration at 37˚C for 18 hours. Samples were aggregated for 18 hours prior to denaturation. (n=3) B) AFM images of resultant denatured samples.

5.3.5 Tau peptides display similar solution state conformations

A plausible hypothesis is that HP301P and HP301L peptides have distinctly different solution-state conformation ensembles. In fact, previous simulations suggested that the P301 residue helps form a protective hairpin in solution that inhibits aggregation in the WT form[111].

The proposed hairpin was observed in simulations of a small peptide of residues 295-311 and supported by cross-linking mass spectrometry. However, direct experimental comparison of the statistically significant distribution of the ensemble protein conformations of WT and P301L Tau peptides are missing in the literature.

To measure the conformational ensemble of the Tau peptides in solution-state, double electron-electron resonance (DEER) was used to generate a probability distribution, P(r), of the segment end-to-end distance, *r*, of the HP Tau peptides. DEER is a pulsed EPR technique that relies on measuring the distribution of dipolar couplings between intra-molecular pairs of nitroxide free-radical spin labels to extract a distribution of intra-molecular end-to-end distances between the pair of labeled sites.

MTSL spin-labels were attached to cysteine residues placed at both ends of HP301P and HP301L, generating the doubly spin labeled dSL-HP301P/L (See Methods). Upon vitrification from solution state, HP301P and HP301L exhibited very similar end-to-end P(r) distributions with a maximum likelihood end-to-end distance of 26.5 Å for both HP301P and HP301L (Figure 5.3A). The HP301L populated more conformations with extended distributions of around 5 nm. However, the populations are within 2 s.d. of each other at all distances. This suggests that any differences in the conformational ensemble of the peptides are subtle, and are due to sparsely populated conformations, or conformational differences that are more local than the end-to-end P(r) can describe.

The P(r) of the end-to-end distances of HP301P and HP301L were also simulated by replica exchange molecular dynamics (REMD), a method which we used in the past for proper sampling of the conformational space of disordered neuroproteins, including Tau[144–147]. The resulting distributions agreed closely with the experimental P(r) measured by DEER. In both

the experimental and computed P(r), the mean R_{ee} were very similar between the two peptides. The most probable conformations, as obtained from the REMD simulations of the peptide monomers, were predominantly hairpin-like. We performed some sort of clustering or classification scheme to characterize the most prominent conformations of the peptides. Figure 4F shows the 6 most populated conformations observed in the REMD simulation. There is a bias towards hairpin-like shapes, which is borne out in the increased contact frequency between residues on opposite ends of the protein (Example Residues Here)…

Next, P(r) was measured of aggregated samples of HP301P and HP301L to characterize changes in their distribution of end-to-end distances of their fibrilized state. To ensure that DEER is capturing the intra-molecular end-to-end distances, the peptide fibrils were formed with a mixture of 95% unlabeled peptide and 5% of the respective end-to-end labeled construct. The final spin label concentration was 20 μM and the total protein concentration 500 μM. Surprisingly, the measured P(r) of the fibril state of both peptides was similar to that in solution state, except for an enhanced long-distance population around 6 nm in the fibril state of HP301L. These results suggest that the majority of solution-state conformations are already close to hairpin folds found in the majority of populations packed into fibrils. This is further supported by the simulated P(r) profile of a homogenous hairpin conformation, calculated using RotamerConvolveMD[60]. The structure of the HP peptide in the published Cryo-EM structure of CBD fibril (PDB: #6TJX) was used to simulate the expected P(r) once the associated spin-labels were attached (See methods). The result was very similar to the experimental P(r) of the solution state peptides, with the majority of distances falling between 2 and 3 nm (Figure 5.3B). The emergence of longer distance populations in the experimental P(r) of the fibril state, however, also suggests that some of the peptide stack to fibrils in more

extended conformations. This finding is not contradictory to the published cryo-EM data of CBD fibril states, given that EPR data reports quantitatively on the entire population in the sample unlike Cryo-EM which selects homogenous classes of fibrils.

Interestingly, we observed when analyzing the time-domain DEER traces that the phase memory time, Tm, was dramatically shortened in the fibril state with Tm~X compared to the monomer state with Tm ~Y, suggesting stronger spin-spin interactions between the end-to-end labels in the fibril. These results are consistent with an increased compactness and/or rigidity of the hairpin in the fibril state. While the shortened T_m of the DEER trace is indicative of enhanced spin-spin coupling in the fibril state, DEER cannot directly resolve distances of 1.5 nm and below, among others due to spectral broadening of the spins at these distances[148,149]. DEER is limited at short distances due to multiple constraints and cannot resolve distances <1.5 nm. At short distances, the length of the pump pulse is longer than the dipolar oscillations, and there are significant contributions to the signal by exchange coupling. Additionally, if a large portion of the spins are in this limited regime, there can be bandwidth limitations on the signal[150]. Instead, continuous wave (CW) EPR lineshape analysis can be employed to identify line broadening due to short range distances between dipolar and exchange coupled spin labels. The dipolar broadening effect between two spins falls off with $\sim 1/r^3$ and is sensitive to distances under 15 Å[151]. Longer distances show negligible broadening compared to the spectral linewidth of the MTSL and can be considered as a single component during fitting. CW EPR line shape analysis was fit to aggregated HP301L-dSL and HP301P-dSL samples to a 3-component fit using MultiComponent software (See Methods). The CW EPR spectra is a linear combination of each spin label measured, so the relative populations of spin-states can be calculated from a least square regression fitting. A mobile component describes labeled,

soluble monomers of HP301L/P. An immobile component describes spin-labeled HP301L/P embedded in a fibril, and a spin-exchange component describes spin-labels in a fibril, and within 15 Å of another label. The CW EPR lineshape fitting results show that for the monomer samples of HP301L and HP301P, 86 % ± 7 % were in a mobile state. However, for the HP301L and HP301P fibrils, greater than 95% of the peptide population is in a fibril state, defined by either the immobile, or spin exchanged components of the fit. Of the labels incorporated into fibrils, 50% of HP301L and 22% of HP301P are in a spin-exchanged state. These results suggest that a greater population of HP301L adopting a closer end-to-end spin label distance compared to HP301P. The DEER P(r) distributions are not capturing these close, end-to-end spin label distances, and the potential differences between HP301L and HP301P fibrils because distances below 15 Å evade direct detection by DEER. Still, the similarity in P(r) between HP301L and HP301P in solution state holds, given that the CW EPR lineshape analysis showed no meaningful difference between the two states, with the majority population found in the monomer states (Figure 5.1D,E).

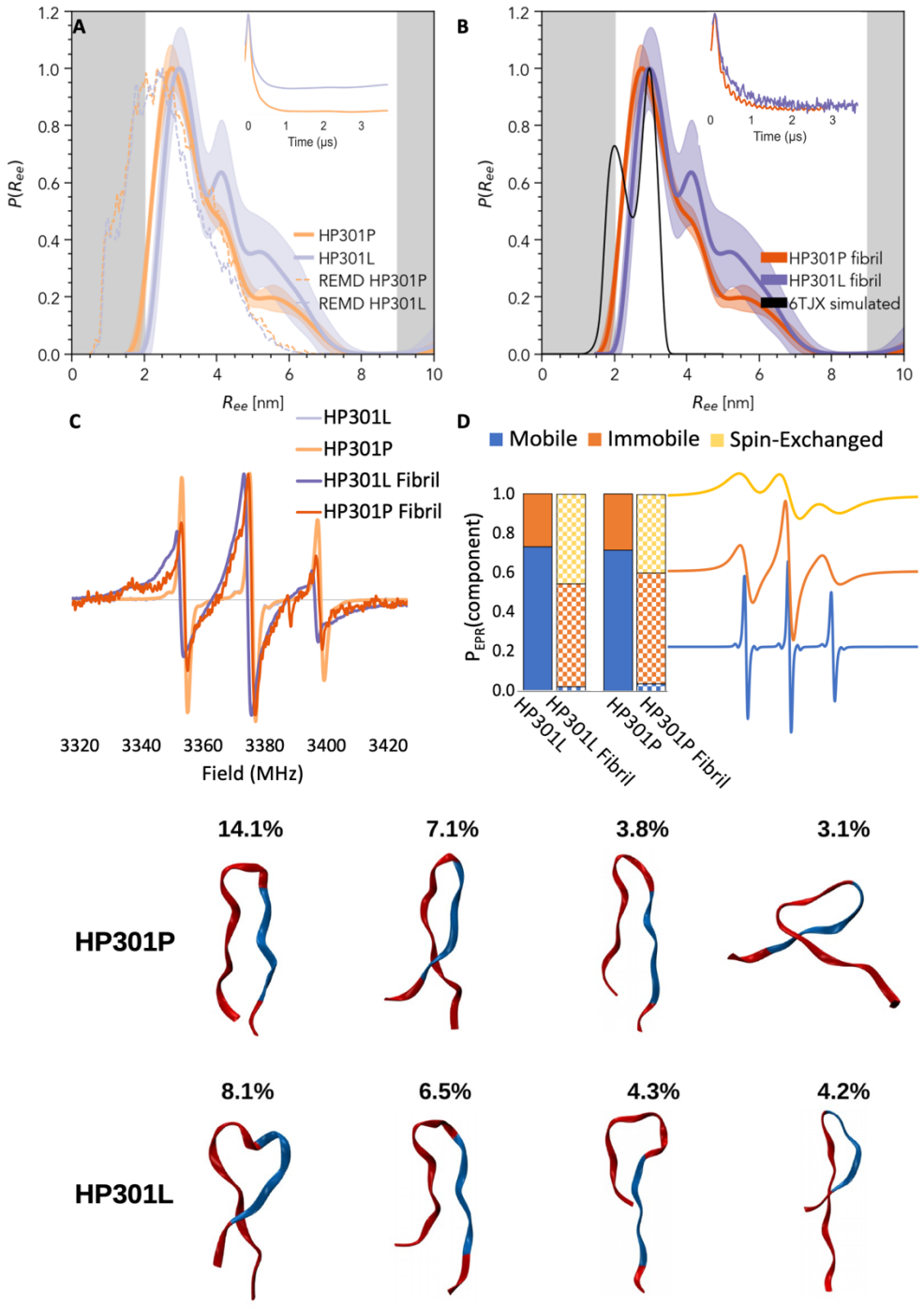


Figure 5.3: A-D) DEER probability distribution P(r) of HP301P and HP301L in solution state and predicted P(r) distributions from REMD simulations (dashed line). (B) fibril state of HP301P and HP301L. (A,B) raw signal is shown in inset. D) CW EPR spectra of HP301P and HP301L before and after fibrilization. D) The proportion of the spectra attributed to mobile (blue), immobile (orange) and immobile, spin-exchanging, species (yellow). D) Simulated EPR spectra of each component used to

fit the fibril populations. E) the four most probable conformations of the HP301P and the HP301L peptides, along with their probabilities, as obtained from the REMD simulations.

5.3.6 The local water structure is perturbed near P301L mutation

The experimental and computed P(r) of the solution state peptides reveal a consistent picture that the major conformational populations of HP301P and HP301L are similar in solution state, and that a minor extended population may be in excess in HP301L compared to HP301P, if at all. So, the key question is what are the characteristics of HP301P and HP301L in solution state that ultimately result in these peptides packing into highly different fibrils in terms of morphology and quantity according to ThT fluorescence analysis, TEM and AFM, and in terms of protein fold and packing within fibril according to a combined DEER and CW EPR analysis? The differences imparted by the mutation must affect minor conformational populations that are obscured in the P(r) by the more dominant conformations, or there must be major differences between HP301P and HP301L that do not manifest themselves in differences in their conformational population. Given the breadth of conditions and length-scales (from peptides to 441 amino acid proteins) that the P301L mutation confers aggregation enhancement, we hypothesized that there is likely a difference that is not accounted for by the end-to-end distance distribution.

When looking for major determining factors for aggregation other than differences in global conformations, differences in local hydration water structure and dynamics, manifested by differences in protein surface hydrophobicity are likely factors. Given that a hydrophobic surface would be one for which the free energy of hydration, $\Delta G_{hydration}$, is positive, we hypothesize that there are differences in the hydration properties between HP301P and HP301L. However, given that the global conformation between them is comparable and both

peptides are highly soluble in water, the difference may lie in hyper-local, residue-level differences that alter the local water structure around the 301 site. Previous work, including our own study, showed that differences in the local hydrogen bond structure of hydration water can lead to differences in the local water dynamics and results in differences in solvation thermodynamic properties[152,153]. We used Overhauser dynamic nuclear polarization (ODNP) to measure the local hydration dynamics near site 301 of the Tau peptides. ODNP relies on cross relaxation of H^1 nuclei of water molecules induced by a free electron spin of a nitroxide spin label. Because the electron-^{1}H cross relaxation falls off with the electron-^{1}H distance r following $1/ r^6$, ODNP is sensitive to the dynamics of water within 8 Å of the spin probe. Hence, if the spin probe is localized to a specific site, ODNP can report on local hydration water dynamics near that site. The efficiencies of electron-^{1}H coupling in solution state that leads to cross relaxation is directly dependent on the speed of movement of water near the electron spin probe and are reflected in ^{1}H NMR signal enhancement. Through the measurement of the ^{1}H NMR signal enhancement and ^{1}H T_1 relaxation times, the electron-^{1}H cross-relaxivity, k_s, can be determined. In the dynamic regime of relevance in this study, k_σ monotonically decreases with slower water diffusion dynamics. ^{1}H NMR ODNP measurements performed at 0.35 Tesla are sensitive to translational diffusion dynamics of water with correlation time in the 10's ps to 100 ps timescale, which corresponds to freely diffusing hydration water (HW) bound to the protein surface with similar or stronger hydrogen bond strength than to surrounding bulk water molecules.

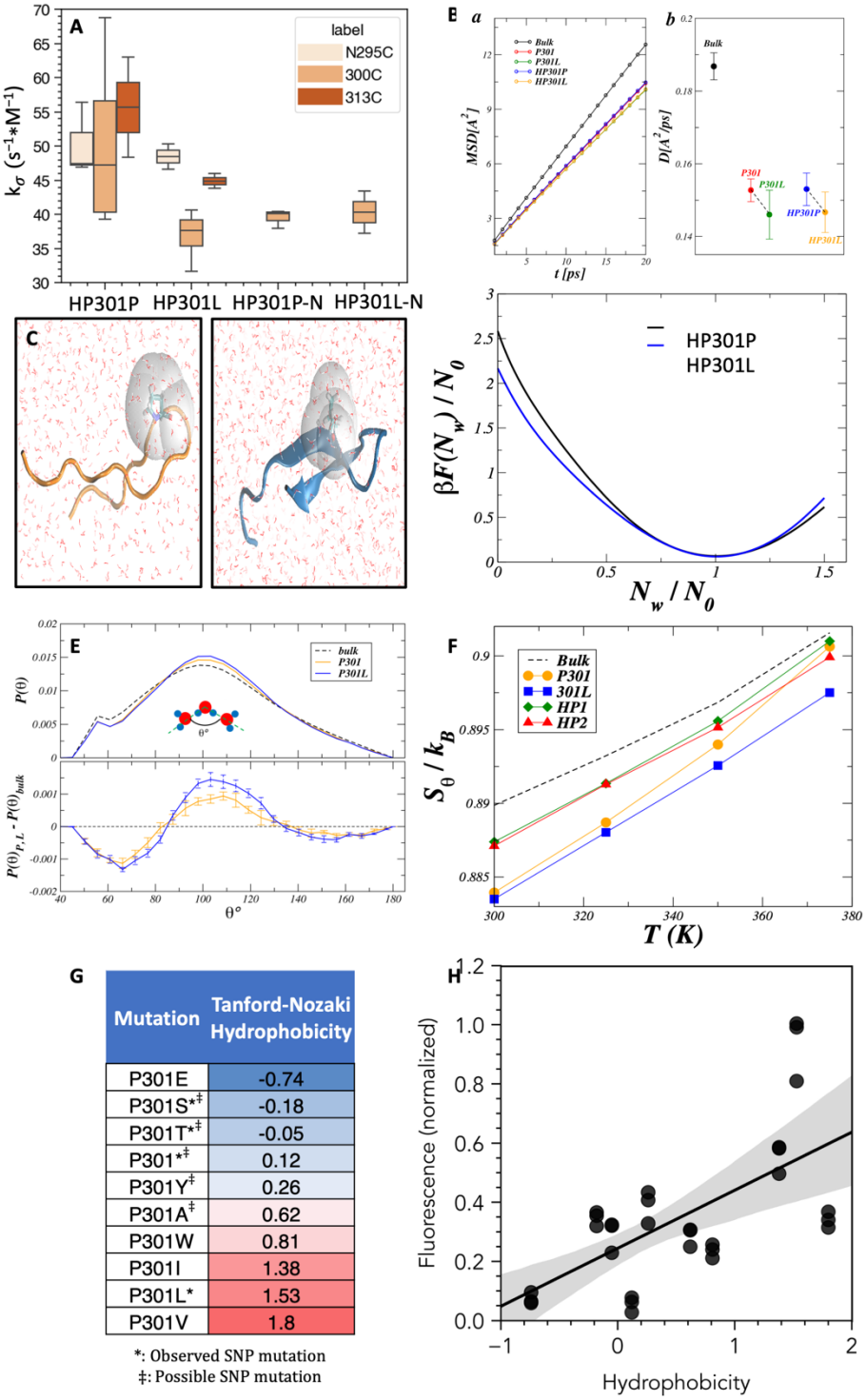
To apply ODNP for the study of local hydrophobicity of HP301P/L, a single spin-label was placed adjacent to residue 301 using a V300C substitution to which a spin-label was attached. The technique does not allow for the 301 site to be directly measured, and G302 is

thought to be important to the PGGGS motif that may be responsible for promoting a turn if a hairpin exists. ODNP-derived measurements of the electron-^{1}H cross-relaxivity show that the P301L mutation has a significant retardation effect on the dynamics of hydration water dynamics measured. For HP301P, a k_σ of 57.9 $\pm$ 32 $s^{-1}*M^{-1}$ and for HP301L a k_σ of 27.3 $\pm$ 13.1 $s^{-1}*M^{-1}$ was found surrounding the V300C site. These are in the range of normal $k_\sigma$values for water on protein or peptide surfaces, while k_σ for bulk water at the same room temperature is 95.4 $s^{-1}*M^{-1}$ [153,154]. This shows that hydration water near the V300C site is slowed, as is the case for all peptide surfaces, but the water near the V300C site is significantly slower in HP301L than in HP301P.

To understand if the hydration dynamics are globally slowed across the entire HP surface or only slowed locally near the 300 site, ODNP measurements were performed with HP301P and HP301L labelled at the N- or C- termini. We found there was no significant difference observed between any of these sites. All sites (HP301P-294C, HP301P-314C, HP301L-294C, and HP301L-314C) displayed a k_σ of 50 $\pm$ 5 $s^{-1}*M^{-1}$, and are 8 and 13 amino acids away from the 301 site. These sites do not display the same retarded hydration dynamics in the P301L mutant as observed near the V300C site, indicating that the change in water dynamics is a local effect that alters waters directly around the 301 site. Upon mutation from proline to leucine, the local hydration water dynamics are slowed, which implies that the local water structure near the 300 site is more ordered in HP301L compared to the HP301P peptide.

The structural and dynamical properties of hydration water are dependent on surface geometry and chemistry. To determine if conformational changes contributed to the retarded water surrounding P301L, the truncated HP301LΔ(304-313)and HP301PΔ(304-313)peptides were measured at the V300C site. No significant difference in k_σ was observed between either

of the truncated peptides, suggesting that local structural properties of the peptide contribute to the local structuring of water seen near site 300 in HP301L, and that the observed change in water dynamics is not due to the effect of a single proline or leucine residue.



Mutation	Tanford-Nozaki Hydrophobicity
P301E	-0.74
P301S*‡	-0.18
P301T*‡	-0.05
P301*‡	0.12
P301Y‡	0.26
P301A‡	0.62
P301W	0.81
P301I	1.38
P301L*	1.53
P301V	1.8

*: Observed SNP mutation
‡: Possible SNP mutation

Figure 5.4: HP301L displays altered hydration dynamics surrounding the 301 site. A) k_σ calculated from ODNP measurements of HP301P and HP301L was measured at three sites, the N-terminus (294C), directly adjacent to the P301(L) site (V300C), and at the C-terminus (314C) of each peptide. A significant reduction in dynamics was observed at the V300C of HP301L in comparison the V300C of HP301P and 294C of HP301L. All other comparisons were insignificant ($p>0.05$) with an independent t-test. $n\geq3$ for all samples. B) XXX Diffusivity calculations of hydration water in MD simulations C) Representative conformations of HP301P (Left) and HP301L (Right) used in umbrella sampling. Grey volume indicates probe volume from which water was expelled. ODNP also probes a similar ~8Å volume surrounding the tethered MTSL spin label. k_σ is proportional to the translational dynamics of the water of the loosely associated hydration shell. D) The dewetting free energy per water molecule of P301 and 301L residues, N_w and N_0 are the instantaneous and equilibrium number of water molecule in the probe volume, respectively, and $\beta = 1/k_BT$, where k_B is the Boltzmann constant and T is the temperature. E) the top panel represent the tetrahedral angle distribution of the hydration water; the bottom panel indicate the distributions relative to the bulk water. F) The angular Shannon-entropy of bulk water, and the angular Shannon-entropy of the hydration water of P301, 301L, HP301P and HP301L as a function of temperature. G) 10 amino acids were substituted at the 301 site of the 19 amino acid HP301P peptide. The peptides vary in hydrophobicity from -0.74 (hydrophilic) to 1.8 (hydrophobic) as determined by the Tanford-Nozaki hydrophobicity scale. H) ThT fluorescence after 15 hrs of aggregation of each peptide vs. the hydrophobicity of the residue (reported in A). Linear regression shows a line of best fit; error bars indicate confidence interval generated through bootstrapping. N=3 for all peptides.

5.3.7 Surface hydrophobicity drives aggregation

Surface hydrophobicity is a complex and multidimensional parameter affected by local chemistries of the residues and the conformational landscape of the surrounding residues, among many other variables[153,155]. To test if there is a direct correlation between the hydrophobicity of the 301 site and the aggregation propensity of the HP peptide, a library of peptides of the same sequence of residues 295-313 was synthesized with 11 different mutations to the 301 site (Figure 5.4A). A ThT aggregation assay was conducted by incubating peptides with heparin at 37˚C as described previously. The maximum ThT fluorescence after 18 hours of aggregation was recorded and compared to the Tanford and Nozaki hydrophobicity scale[156]. A linear fit was performed, and confidence intervals were estimated through bootstrapping of

the data (Figure 4H). The result shows a significant correlation (p=0.0001) between the hydrophobicity of the residue estimated by the Tanford-Nozaki scale and the ThT maxima, indicating that the hydrophobicity of the 301 residue is key in determining the aggregation propensity of these peptides. Although the trend of hydrophobicity increasing b-sheet formation was strong, both the WT and P301L peptides are outliers in the fit, with significantly lower and higher fluorescence intensity, respectively.

5.3.8 Disrupted hydration shell of hydrophobic pocket favors dehydration

Given that the hydrophobicity closely correlates with the aggregation propensity of the peptides, and the hydration dynamics are slower around the P301L residue, we asked if the perturbation of the hydration layer could explain the changes in aggregation, and if the effect of P301L on the local water structure could be the primary driver for aggregation in the HP301L peptide. Specifically, the question is whether the slower, more structured, water near the 300 site of the HP301L peptide constitutes a low entropy water reservoir relative to the bulk water that renders stronger entropic driving forces for HP301L peptide assembly, or promotes protein-protein association through enhanced hydrophobic interaction.

To answer this, we used indirect umbrella sampling (INDUS) method[157,158] to compute the free energetics of dewetting the protein surface at the 301 residue (see the simulation details in SI). The dewetting free energy measures the local water-protein interactions. The hydration shell (or the probe volume) of the residue is defined by a union of spherical sub-volumes with radius R_v that are attached to the center of every heavy atom (carbon, oxygen, nitrogen, and sulfur) of the 301 residue. We chose R_v = 0.55 nm to include only the first layer of hydration water in the probe volume. After performing INDUS simulations at different windows of overlapping number of water molecules within the probe volume, we used Un-binned Weighted Histogram Analysis Method (UWHAM)[159,160] to calculate the unbiased free energy

of dewetting the hydration shell around residue 301. Due to the intrinsically disordered nature of IDPs, multiple conformations must be considered to fully describe the hydration state. To properly predict the conformational dependent free energies, the dewetting free energy of residue 301 was obtained independently for the six most probable conformations of HP301P and HP301L peptides determined by REMD simulations (see SI). The average dewetting free energy of residue 301 in HP301P and HP301L peptides was calculated by a weighted average of this subset of conformations according to the relative abundance in the REMD simulations. As illustrated in Figure 4D, the 301L site exhibits a lower dewetting free energy per water molecule, indicating that the local attractive interactions between the water and residue are weaker than the corresponding P301. Therefore, dehydration of the HP301L peptide would be more favorable than HP301P, hence rendering HP301L more hydrophobic compared to HP301P around site 301.

Depending on the topography and the chemistry of a particular residue the water molecules in the hydration shell can form ordered structures, which is expected to result in reduced entropy and diffusivity of these water molecules. Particularly, due to the lack of direct hydrogen bonding, the hydration water of a hydrophobic moiety tends to manifest ordered pseudo-clathrate motifs that are composed of multiple secondary structures (such as pentagonal and hexagonal water rings)[161]. However, computational studies of the HP301P/L peptides need to substantiate these ideas. The so-called triplet or 3-body angle distribution[162,163] -- the distribution of angles formed between any hydration water and two of its first-shell neighbors – was used as an indicator of structural behavior of the hydration water network. As shown in Figure 5.4E, the triplet distribution features a peak near the tetrahedral angle, which indicates the propensity of water to form tetrahedral networks such as the pentagonal and hexagonal

rings. Hydration water near both residues P301 and P301L shows an increase in tetrahedrality, evidenced by an increase in the distribution near the tetrahedral peak relative to bulk water. Interestingly, hydration water near residue P301L shows a greater increase in tetrahedrality relative to P301, suggesting that the hydration water of 301L forms a more tetrahedrally ordered water network. To further quantify this concept, the angular Shannon-entropy was measured for hydration water around residue 301 in HP301L and HP301P peptides at different temperatures (Figure 5.4F). An increase in temperature results in a gain of entropy in all hydration water; however, the P301L residue maintains a smaller hydration water entropy than P301 across a wide range of temperatures. To a lesser extent a similar trend is also observed for the entropy of the entire hydration water of the HP301L and HP301P peptides. Thus, one can conclude that the entropic driving force of HP301L aggregation is larger than HP301P aggregation due to the larger entropy difference with bulk water. With increasing the temperature, the hydration water of site 301L and the total hydration water of HP301L peptide more slowly approach the bulk water entropy value compared to the P301 and HP301P peptides. Therefore, even at higher temperatures, the entropic driving force of HP301L assembly remains more favorable than the entropic driving force of HP301P assembly.

5.4 Discussion

Our study set out to elucidate the mechanistic basis of the P301L/S mutation in Tau that is, by far, the most commonly used mutation to mimic AD pathology in vitro, in cell and in mouse models. We find that a minimal, 19-residue, fragment spanning residues 295-313 that includes the P301 mutation site, the PHF6 motif and is known to form a hairpin in the context of pathological Tau fibrils in CBD, PSP, GGT and GPT showcases the dramatic effect of the P301L mutation, leading to enhanced aggregation, fibril quantity, fibril stability, and seeding

competency. In short, we observe that enhanced local hydrophobicity at the 301 site gives rise to low solvation entropy that hence offers an increased entropic driving force for fibrillization upon dehydration at site P301L. Interestingly, such effect of water structuring around site 301, enhanced aggregation propensity and fibril quality is only observed in the full HP301L peptide, not in any of the truncated HPΔ(295-299) or HPΔ(304-313) fragments, even though each fragment contains the P301L mutation. This means the structural context of the 301 site is critical to its effect on Tau aggregation and seeding.

The PHF6 motif is well known for forming dry interfaces in amyloid fibrils in steric zippers and other similar b-sheet conformations[140]. It is likely that HP301P and HP301L lie in a conformation that creates a dry interface, but structural biology is needed to confirm if the conformation is a hairpin similar to that observed in the CBD or PSP structures[42,45,124]. The dry interface, formed by amide-stacking of sidechains, has a favorable enthalpic contribution from H-bonds and hydrophobic interactions, but imparts an entropic penalty on the aggregation, due to the expulsion of water[164]. It follows that a less favorable interaction, with a higher entropy in the hydration-water protein interface would lead to a smaller entropic penalty in the formation of a fibril interface. Indeed, we find through ODNP measurements that the local hydration water dynamics are retarded locally around the P301L site when compared to the WT form. Previous studies have found that in DNA the dynamics of the hydrating water are inversely correlated with the entropy of hydration[165], and the Adam-Gibb's relation has long established the relationship between water translational motion, and configurational entropy.[166] Thus, the expulsion of slower moving hydration water is more entropically favored (larger ΔS), resulting in an free-energetically more favorable dehydration around the P301L residue. We find de-wetting simulations support a similar conclusion, with the free energy

change of HP301L dehydration being lower than the HP301P. The energetic cost to remove all waters from the region surrounding the P301L residue is lower than the corresponding WT form, meaning the de-wetting on the pathway to a dry β-sheet is more likely to occur in the P301L form.

ODNP also shows that this effect is very local to the 301 site. The water retardation effect is only observed surrounding the V300C adjacent residue to P301L. The N-terminus, and C-terminus of the peptide display the same dynamics across peptides. This explains why it has been historically difficult to measure solution state differences in mutant Tau; the protein or peptide behave very similarly, as an ensemble, regardless of mutation. DEER measurements and REMD simulations both confirm that the populations of conformations are difficult to distinguish when bulk properties like R_g or end-to-end P(r) are measured.

Because of the more favorable de-wetting stage, we posit that the P301L site is the initiator for aggregation. The site may de-wet and undergo a conformational rearrangement to place the flanking residues to PHF6 in the stabilizing position described above. This explains the high primary nucleation rate of P301L others have observed in comparison to other more hydrophilic mutations (P301S).

This work highlights the importance of the hydrophobicity, and the hydration shell as important parameters to consider when designing aggregation inhibitors or investigating new mechanisms and drivers of aggregation. Much of the Tau aggregation field has focused on anionic cofactors (heparin, arachidonic acid, nucleic acids) to initiate aggregation, but perhaps more hydrophobic cofactors should also be considered as have been observed in other amyloid forming proteins.

5.5 Future Directions

5.5.1 Mutation landscape of Tau in Tauopathies

While hydrophobic differences in the P301L mutant drive aggregation of HP301L peptides, it is not known if hydrophobic driven aggregation is a common event in other Tau mutations. There are 112 SNPs identified in the *MAPT* gene. The mutations are highly concentrated in the MTBR of Tau. Of the identified SNPs, 99 are coding mutations or deletions; 41 of which are pathogenic and 26 have unclear pathogenicity. Using the Tanford-Nozaki hydrophobicity scale[156] the hydrophobicity of the mutant residue was compared to the WT residue. The mutant was more hydrophobic than the WT in 62% of instances ($p=0.03$) (Figure 5.5). While this suggests that hydrophobic mutations may bias Tau towards an aggregation prone state, it is possible that the codons coding for each mutated codon are limited by the number of replacement options and that random chance is more likely to result in a hydrophobic mutation. More work must be done to study how universal the hydrophobic aggregation effect is throughout the length of the cell.

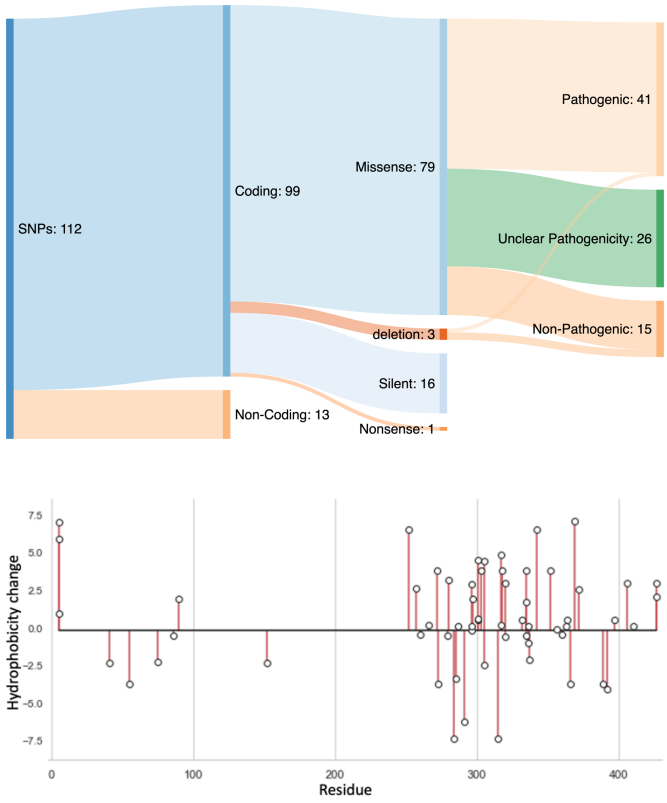


Figure 5.5: Top) The outcomes of SNPs in *MAPT* to Tau's sequence. Bottom) The change in hydrophobicity of each known pathogenic mutation to Tau. Positive values indicate a more hydrophobic mutant residue. On average, mutations cause a hydrophobic shift to Tau.

5.5.2 Post-Translation Modifications of Tau

In many Tauopathies, Tau in postmortem tissue contains a large number of post-translational modifications (PTM). Tau is hyperphosphorylated, acetylated and ubiquitinated particularly in the PRD and MTBD[167]. Many of these PTMs are predictive of pathology in AD patients[167]. Acetylation and phosphorylation both alter the charge profile of the protein, and ubiquitination adds a significant amount of steric hinderance. Many hypothesize that the PTMs, that are characteristic of each Tauopathy, influence the final structure[124]. The HP301L peptide presents a useful test peptide to begin studying the effect of mutations with 4 PTM site available. Once cryoEM structure determination is readily accessible, it should be possible to screen through all of these PTMs and compare the structural changes induced by modification.

Chapter 6: Structural determination of HP301L

6.1 Introduction

Based on negative stain TEM conducted on HP301L fibers described in Chapter 5, we believe it may be a good candidate for single particle analysis with cryo-EM. Negative stain TEM confirmed the formation of continuous, homogenous fibrils with helical periodicity that is needed for structure determination with cryoEM. At this time, there is no atomistic structure of Tau fibrils that contains a disease associated polymorphism, but there are structures of Tau from Tauopathies containing the WT form of HP301L embedded within a larger fibril core. In particular, HP301P forms distinct hairpin folds within the CBD, GGT and GPT fibril cores[123]. Resolving the atomistic structure of HP301L fibrils would give a better understanding of the effect on fibrilization by the P301L mutation by being able to compare the resulting structure with the WT form within Tauopathy fibrils.

6.1.1 Cryogenic Electron Microscopy

Electron microscopy (EM) relies on the interactions between electrons, produced by a field emission gun (FEG) in an EM beam, and the nuclei of the sample being imaged. In biological samples, most electrons pass through the sample unhindered. When an electron hits the sample, it is scattered inelastically, which contributes noise or does not make it to the detector, or elastically which induces a phase shift that provides the necessary contrast at the detector. In thin samples, the phase contrast can be very low, so EM is conducted at defocus to enhance the phase contrast. Organic molecules C, N, and O have very weak scattering, so samples must be imaged under cryogenic conditions, or by applying a heavy-atom contrast agent in negative stain electron microscopy (NS-EM).

NS-EM embeds the biological sample in a heavy-atom salt, uranyl acetate (UA) for all samples in this manuscript. In NS-EM the sample is adhered to a copper grid before going through what is essentially a buffer exchange into the UA. The copper grid is plasma cleaned to create a charged surface that attracts the particles before the remaining buffer is blotted and replaced with UA. NS-EM can retain native structures of proteins while preserving them (samples can be imaged almost indefinitely) and greatly enhance phase contrast in imaging. NS-EM is used to check sample quality and can be used for low resolution structure characterization. NS-EM is limited in resolution by the grain size of the stain (~16 Å for UA) so cryo-EM must be used if atomistic resolution is desired.

In cryo-EM, samples are vitrified on grids that consist of a uniform copper mesh with a thin film of carbon across the grid. The carbon film has uniform holes 1.2 μm to 2 μm diameter; within each hole there is a region of suspended vitreous ice. The carbon backing material contributes significant background noise, and crystalline ice has strong scattering peaks that interfere with 3D reconstructions. To form amorphous ice, a sample is applied to the grid, and plunged into liquid ethane. Liquid ethane has higher thermal conductivity than liquid nitrogen, so the initial vitrification must be done in ethane. The samples are then transferred to a liquid nitrogen bath and are stored under liquid nitrogen until imaged. Samples must be kept below the glass-transition temperature (<136 K[168]) throughout sample preparation and imaging. The cryogenic conditions also reduce beam-induced radiation damage. CryoEM is a low SNR technique because the radiation from the beam is destructive to the sample, so the number of electrons (e^-) incident to the sample must be minimized, while optimizing resolution. In practice, samples can withstand 10 to 100 $e^-/Å^2$ before major distortions begin to inhibit image quality[169,170].

The low SNR in cryoEM makes particle alignment one of the crucial steps in reconstruction. Each image is a 2D projection of the 3D object. To create 3D reconstructions, each image must be classified based on the projection angle and translational directions. Particles are aligned with a maximum likelihood algorithm with equivalent orientations can be averaged together into 2D classes. For peptides as small as HP301L, helical symmetry is crucial to structural determination. The rotational symmetry gives projections of the particles at multiple rotational angles in each particle. For example, a 700Å wide particle of an amyloid fibril will contain almost 150 subunits, each of which provides a 2D project rotated by a fixed amount (the helical twist). This characteristic of helical assemblies and the fourier projection theorem allow 3D initial models to be built from 2D classes. The initial 3D model generated from the result of 2D classification is then used as a prior in the Bayesian maximization methods used in RELION for 3D refinement. After 3D refinement is done with an initial low-resolution model, a more highly featured model should be produced. Using this new model, particles can then be refined using the model as a prior. This step allows for higher quality motion correction and CTF refinement, as it can be conduct with a reference sample, and corrected on a per-particle basis. This process of model refinement, followed by particle polishing and refinement can be iterated multiple times, and with amyloid fibrils is often more necessary than for globular proteins. In particular, particle polishing is often conducted earlier and can provide significant enhancements in resolution[171].

6.2 Sample preparation

UCSB will have a dedicated electron microscope for single particle cryo-EM by the end of 2022, but currently the capabilities to automate the image collection for single particle analysis (SPA) are not available. A proposal for microscope time was submitted, and accepted

by, the Pacific Northwest Cryo-EM Center (PNCC). Samples were sent to PNCC for the grid screening and data collection stages.

We conducted negative-stain TEM at the UCSB Microscopy and Microanalysis Facility (MMF) and observed that HP301L formed a population of fibrils with a significant portion forming fibrils a uniform helical morphology. These fibrils have a half-period of approximately 700Å and a diameter between 70 Å and 100 Å in width. This helical morphology lends itself well to single particle analysis with cryo-EM.

6.3 Results

To determine the appropriate concentrations, buffer conditions, sample preparation and vitrification parameters, cryo-EM was performed in the MMF using G2 sphera Tecnai, FEI Titan, and FEI Talos microscopes using a Gatan 626 cryo transfer holder. Acceptable conditions were determined for vitrification of fibrils to use 50 μM protein samples, in a 20 mM ammonium acetate buffer with 10 mM to 100 mM NaCl concentrations. Multiple grid application conditions were identified to produce acceptable grids conditions.

Quantifoil 2/2 grids were prepared with either one application of 5 μL, or 2 applications of 3 μL. A single blot of 2s was applied with a Vitrobot at 0 force rating, in the case of 2 applications, a manual blot was applied between the first and second sample application by touching the torn edge of a filter paper to the backside of the grid. Grids were sent to PNCC for further screening and data collection. In total, 28 grids were screened and ranked on a 200kV Arctica microscope (Figure 6.1). Grids were scored based on the overall quality of the grids, i.e. the distribution of ice thickness, and amount of contamination, the number of quality squares and holes that would be available for data collection, and the quality of samples and sample distribution within holes. Small data collections were collected on the Arctica at the

time of screening on the best grids, and the two best grids were loaded onto a Titan Krios microscope with a falcon 3 detector, and bioquantum energy filter.

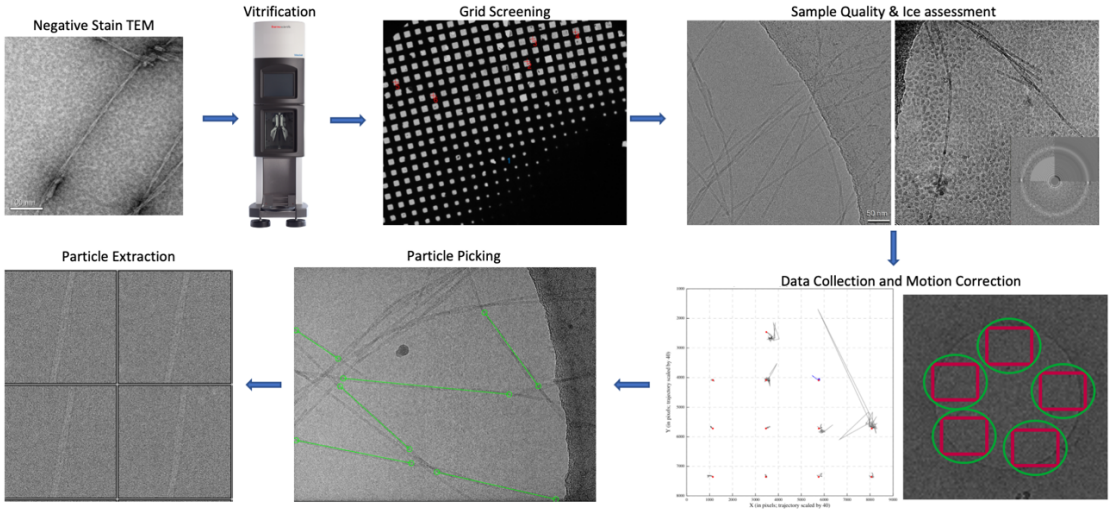


Figure 6.1: Data collection for Cryo-EM imaging. Clockwise from top left: **1)** Pure, homogenous protein is needed for high quality data collection. NS-EM is used to confirm protein homogeneity and concentration. **2)** Samples are vitrified through plunge freezing into a liquid ethane bath. **3)** Grids are screened for overall quality, number of squares with correct ice thickness for data collection. **4)** Ice quality assessment to ensure vitreous ice of appropriate thickness is present. An example of vitreous ice (left) and crystalline ice (right). (Inset) Fast Fourier Transform (FFT) of the image can be used to judge possible resolution, and diffraction spots of crystalline ice can be detected. **5)** Full data collection of best grids. For HP301L, 5 images per 2 µm hole were collected around the edges of the hole. Each hole with good quality ice was selected, and micrograph collection was automated in SerialEM. **6)** Start and end points of fibrils are picked manually. **7)** Fibrils are broken into multiple segments at defined intervals along length of fibril and extracted as individual segments. The main parameters of interest are the spacing between segments which determines how much overlap across particles is present, and the size of the particles.

2670 micrographs were collected on a Thermo-Fisher Krios at PNCC, but significant contamination and ice was observed on the grid in comparison to the screening results of the same grid. Additionally, there was a translation of the stage that caused the best regions of the grid identified in screening to be out of range of stage shift. Fibrils were observed to partition towards the edge of holes, so 5 micrographs were collected per 2 µm hole, focused around the

edge of the holes. Of the collected micrographs, 605 were determined to not have significant ice or other contamination, and some samples. Micrographs had a total dose of 50 $e^-/Å^2$, and a defocus range from -1.0 μm to -2.0 μm. Micrographs were motion corrected using the Motioncor2 implementation in relion4.0-Beta. Due to the presence of the carbon edge in many micrographs, motion correction was complicated by the relatively stationary carbon, and its associated high contrast. Motioncor2 selected the carbon edge as many of the reference points in motion correction and thus was not very effective. Particle refinement and per-particle motion correction refinement significantly enhanced resolution when used early in the refinement process as the raw images were not corrected for local ice motion[171].

Fibrils were manually picked in Relion. 768 pixel particles were extracted with between 3 and 10 asymmetrical units and downscaled to 256 pixels. 2D classification was performed with the stochastic decent method, and two distinct fibrils populations emerged: a singlet (Figure 6.2A), and doublet fibril (Figure 6.2B). The doublet fibril did not contain enough particles to proceed to 3D classification or refinement. Eight 2D classes of the singlet type were used to create an 2D initial model that was used in 3D refinement (Figure 6.2C).

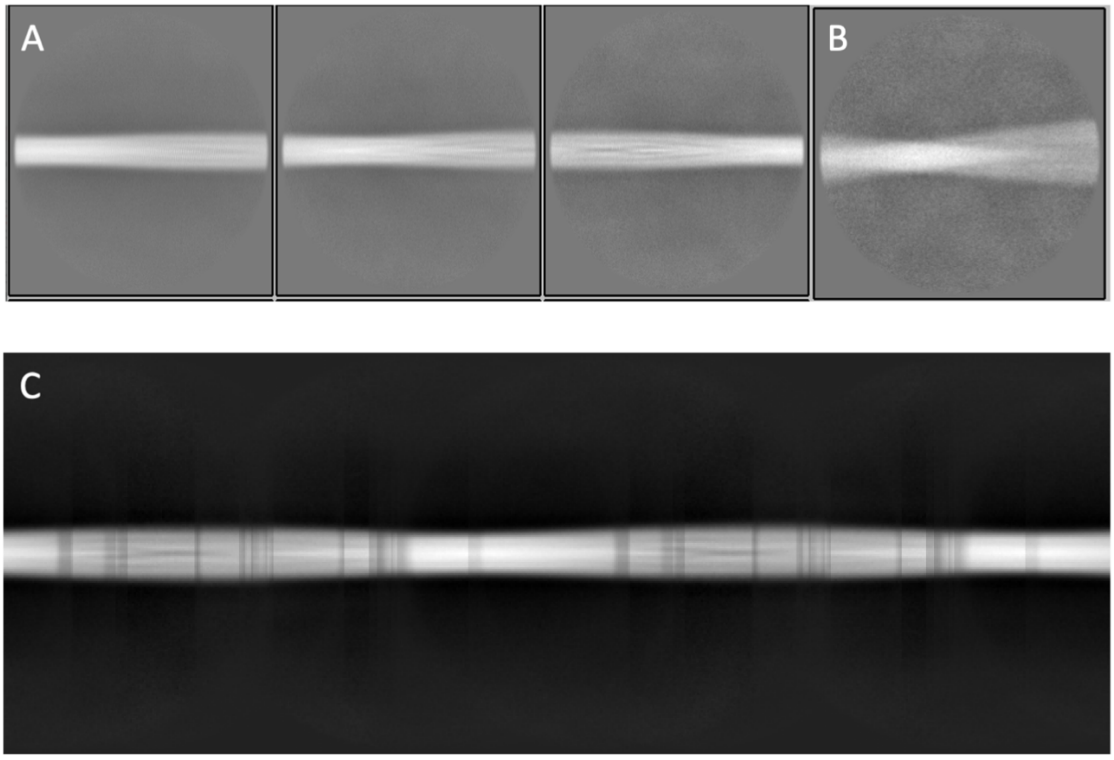


Figure 6.2: A) Representative 2D class averages of the singlet class of fibrils used for initial model building. Cross-β sheet periodicity can be observed in many of the models. B) A 2D class average C) An initial 3D model constructed from 8 2D classes. The initial model is used for initial 3D model building and refinement.

Particles went through multiple rounds of 3D refinement followed by CTF correction and particle polishing until a final resolution of 4.46 Å was resolved as determined by the gold-standard Fourier Shell Correlation (FSC) (Figure 6.3). After initial 3D refinement, a C_2 symmetry was observed, and imposed for future refinements. A pseudo-C_2 screw symmetry was also tested but gave poorer resolution after refinements, indicating that it was not the correct symmetry to impose.

The 3D electron density map shows two distinct C-shaped conformations that are approximately the size of an HP301L peptide folded in a hairpin. Atomic resolution was not

achieved, but an approximate structure was observed to be similar to the hairpin conformations observed in other Tau amyloid filaments. An unresolved density is also present in the center of the filament. It is possible that this is an additional unresolved proteinaceous density, or may be the heparin that is added to initiate the formation of the filaments.

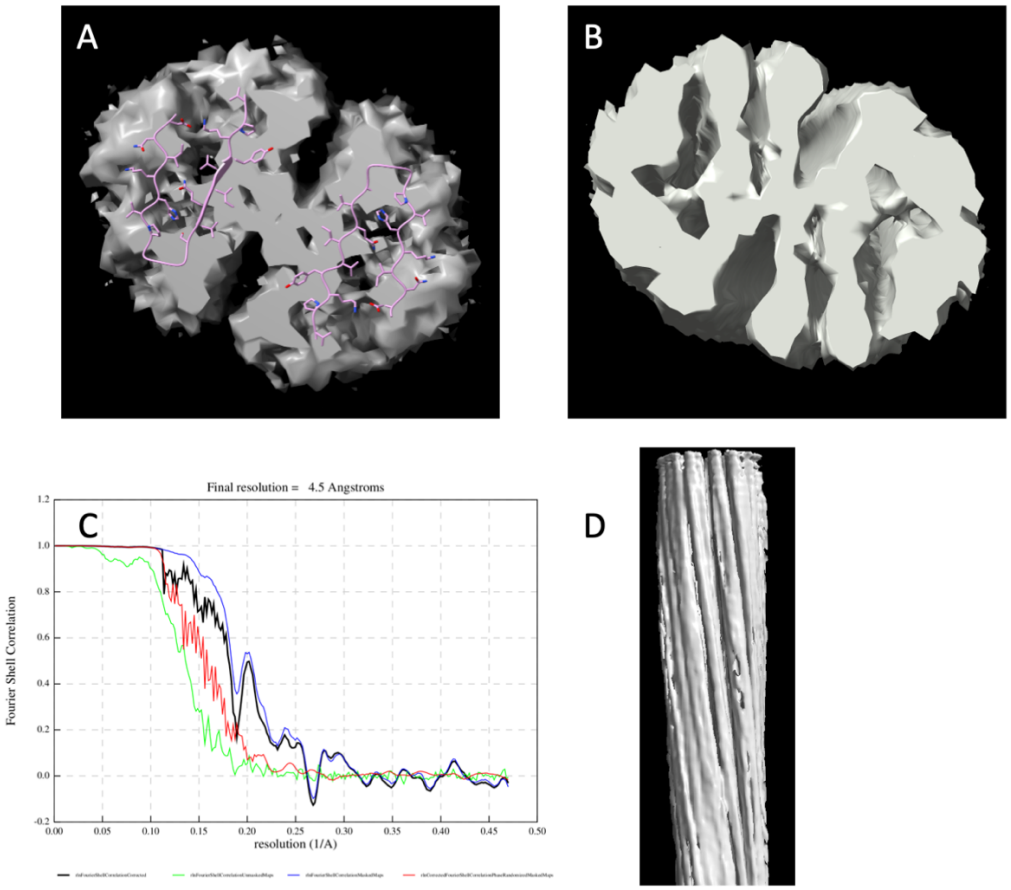


Figure 6.3: A) The final electron density map after 3D refinement is shown (grey) with a cartoon of the HP301P peptide from the CBD PDB #6TJX (pink). B) The unmasked electron density map after postprocessing. C) FSC curve of the postprocessed map. D) a view of the reconstructed density from the side. Cross-β sheet separation is observed at the top of the fibril.

6.4 Conclusion

This study laid the foundation for future cryoEM studies of amyloid fibrils at UCSB. A 4.6Å structure was produced of the HP301L peptide described in chapter 2. While atomistic

detail was not achieved, this study supports the DEER results from Chapter 5 that would indicate a hairpin-like conformation, with close end-to-end contacts. The fibril structure has a C_2 symmetry with 2 hairpin-shaped protofilaments. This is the first known structure containing the P301L mutation. P301L is the most commonly used mutation in Tau studies, and understanding the effect it has on structure is of great importance[123,172]. If P301L forms completely different structures than the equivalent wildtype protein, serious consideration must be given to its use in disease models.

Work is ongoing to produce a higher resolution structure of these fibrils. The single biggest obstacle currently is the amount of time sample iteration takes. After sample preparation, there is approximately a 5-week delay before screening can be completed. If a grid looks promising, it is submitted for data collection which takes 7 more weeks to gain time on a microscope meaning it takes approximately 3 months from sample preparation to receiving images for analysis. During this project, only a single grid had a full data collection performed and it had suffered significant degradation in quality between screening and collection. The reconstruction presented above reaches the approximated resolution limit of the images. To improve the reconstruction quality, more data must be collected. There is a lot of promise in routine use of cryoEM for amyloid fibril determination. We have now developed a pipeline for sample preparation, data collection, and processing of the HP301L samples. In future, probing the effect of other mutations and posttranslational modifications, such as phosphorylation and acetylation will provide great insight into the modification of fibril structure. There are also many other structural motifs in Tauopathies that could be explored. For example, we began to characterize fibrils formed by a portion of the AD PHF core and have identified conditions that look very promising for future cryoEM studies (See supplement).

Chapter 7: SH3 domains can bind the polyproline rich domain of Tau and regulate its ability to form Liquid-liquid phase separated condensates

7.1 Background motivation

Once the successful generation of Tau fibrils with disease material is repeatable and commonplace there will the complexity of systems studied can become more complex. Tau aggregation exists in a highly crowded, complex cellular environment. It is highly unlikely that Tau is known to interact with many proteins already. In this manuscript, Tau has been described to interact with microtubules the proteasome and the immunoproteasome, and there are multiple other proteins already identified to bind Tau. There are significant questions about how Tau is transported from its microtubule bound state into a fibrillar state and finally how prion-like forms of Tau are spread throughout the brain. Here, we detail the binding of Tau with a neuronal form of BIN1, a signal transducing adaptor protein that is involved in endocytosis of synaptic vesicles in the neuronal form. Further we identify the SH3 domain off BIN1 to bind to the proline rich domain (PRD) of Tau in liquid liquid phase separated (LLPS) condensates.

Work by Xuemei Zhang in Ken Kosik's lab identified a novel interaction between BIN1 and Tau, and found that BIN1 assisted Tau in forming LLPS condensates in cells through an interaction with the PRD. Because the PRD is involved in the microtubule (MT) binding interface of Tau, the interaction with BIN1 was found to disrupt Tau-MT complex, and instead sequester Tau with BIN1 in LLPS condensates.

We assisted with this work by producing recombinant Tau PRD, the BIN1 protein and an SH3 domain of BIN1 known to bind proline-rich sequences. We found the PRD of Tau to readily phase separate in the presence of anionic cofactors, but not to be prone to amyloid formation. CW EPR was then used to show binding of Tau PRD with the SH3 domain, and confocal microscopy showed colocalization of Tau and BIN1 or SH3 in the condensates. In contrast to the cellular results, neither SH3, nor BIN1, were found to enhance the condensate formation.

7.2 Abstract

Tau functions in dynamic exchange with microtubules, and under pathological conditions transitions to solid fibrillar aggregates. Recent *in vitro* data has revealed an additional phase state of Tau in which the protein undergoes liquid-liquid phase separation (LLPS); however, observations of this phase transition in living cells are limited. Here we have examined the domain properties of Tau using the optogenetic protein Cry2 to show that the polyproline rich domain (PRD) drives LLPS under the control of Tau phosphorylation. The condensates occupy a localized volume in which the constituents reach a higher concentration than in the surrounding solute phase. After cycling through multiple illumination periods, the rate of the disassembly gradually slowed and eventually plateaued. We discovered that Tau in the form of a condensate facilitates binding to microtubules. When dissociated from the microtubule, Tau has the potential to condense with other proteins through its PRD. The expression of EB1, a regulator of plus-end microtubule dynamic instability, resulted in polyproline rich domain Tau/EB1 colocalized biocondensates. Through an interaction with SH3 domain of BIN1, encoded by an Alzheimer risk factor gene, the Tau PRD formed condensates in cells and *in vitro*. Condensates can deliver a high local concentration of Tau,

increase the dwell time on MTs and achieve a length scale that can extend the relatively small size of the Tau protein to the macroscopic lengths of microtubules. Multivalent and intrinsically disordered domains can sense global cellular conditions and adjust the phase state of molecules needed to perform specific functions such as those related to control over MTs.

7.3 Introduction

Tau protein undergoes a liquid to solid phase transition in numerous neurodegenerative conditions, most prominently Alzheimer's disease, frontotemporal dementia and chronic traumatic encephalopathy as well as more rare conditions such as Pick's disease, progressive supranuclear palsy, corticobasal degeneration, Neiman-Pick type 3, sub-acute sclerosing encephalitis among others. The diverse conditions associated with Tau pathologies known as Tauopathies, and their pathobiological mechanisms that lead to aberrant Tau are poorly understood despite a decades long effort since the identification of Tau in these. In the past few years, the existence of liquid liquid phase separated states called condensates have been observed among several proteins that accumulate in neurodegeneration including FUS, TDP-43, HNRNPA1, and DDX4 as well as Tau. An important feature of condensates is the very high concentration of protein within a small volume compared to the solute phase, while the protein constituents continue to exchange freely. The thermodynamics of the phase transition of biopolymers from solution to condensed state by liquid liquid phase separation (LLPS) has been studied extensively using polymer physics models and tools, but the correlates of these events in living cells are mostly unknown. LLPS *in vitro* can be well-described as simple or complex coacervates; however, condensates in living cells involve multiple RNA, lipid and protein components whose binding affinities and concentration are dynamically regulated and result in an ever-changing *in vivo* phase diagram with manyfold more phases and intrinsic

variables that control the transitions among them. While *in vitro* studies can prove the sufficiency of molecules to undergo LLPS they strip the system of its inherent complexity. A complimentary technique for grasping the function of phase transitions, and how they are regulated *in vivo* is to make acute changes in the valency of a core scaffolding protein by fusing it to a photo-sensitive oligomerization domain, thereby gaining the ability to abruptly alter the location in phase space within the cytoplasm.

One way to achieve this light-mediated change in protein valency is to use the optogenetic protein Cry2, the photolyase homology region (PHR) of a cryptochrome from *Arabidopsis thaliana*, that undergoes light-sensitive self-association upon blue light exposure. Cry2WT-mCherry (Cry2 wild type) by itself will not undergo LLPS in heterologous cells, but when fused to the IDR (intrinsically disordered region) of FUS, HNRNPA1 or DDX4 these proteins can drive droplet formation. LLPS of these proteins likely involves additional cellular factors that lower the free energy for LLPS that the blue light activated Cry2 mimics. Tau is an intrinsically disordered protein (IDP), binds RNA with multivalent interactions and, like the IDR domains of many RNA binding proteins, Tau can undergo phase separation *in vitro*. Here we demonstrate that Tau undergoes LLPS in living cells, and the polyproline rich domain (PRD) of Tau controls condensate formation at discrete cytological loci in a manner that can solve length scale problems required to match single proteins to macromolecular complexes.

7.3.1 Cellular Results: The proline rich domain mediates Tau condensation in cells

Tau condensation facilitates microtubule (MT) binding

We cloned a full-length Tau 2N4R construct fused to Cry2WT- mCherry (CWT 1–441) and, in the absence of blue-light activation, we observed the expected MT pattern and some diffuse signal in the cytoplasm. With Cry2 activation by blue light, the signal from the diffuse pool was markedly reduced, and the MT signal increased (Fig. 1, A–C). The increased Tau binding to MT was quantified by computing the pixel coefficient of variation (CV) across the cell, with the lower value indicating a more uniform distribution and a higher value indicating discrete lo- calization (Fig. S1 A). This result suggested that Tau condensate enhanced its binding to MTs. To support this observation, we expressed Tau 1–441-EGFP (Fig. S1 B) in HeLa cells and treated cells with 1,6-hexanediol (1,6-HD) which interferes with weak hydrophobic protein–protein or protein–RNA interactions that can promote the formation of dynamic, liquid-like assemblies (Kroschwald et al., 2017). 3min after 1,6-HD treatment, the EGFP signal appeared diffuse in the cytoplasm (Fig. S1 C). The control, 2,5-HD, had no effect even after 20 min (Fig. S1 D). MTs were unaffected by these treatments as indicated by 1,6-HD treatment of cells labeled with mCherry-α-tubulin (Fig. S1, E and F). Similar results were observed with Tau 256–441-EGFP (Fig. S1, E–G).

The Tau PRD mediates Tau condensation in cells

With the finding that Tau phase separation as a condensate can enhance its binding to MTs, we sought to determine which do- mains ofTau drive condensation-assisted MT binding and began by analyzing them individually. The Tau domains consist of the N-terminal domain, the PRD subdivided as P1 and P2 with ∼25% prolines, the MT-binding domain (MTBD) with four repeat re- gions, and the C-terminal domain from amino acid 369 to 441, which includes a short P3 region flanking the MTBD (Fig. 1 C; Gustke et al., 1994; Mukrasch et al., 2005).

Although the MTBD can undergo LLPS in vitro (Ambadipudi et al., 2017; Zhang et al., 2017), when this domain was fused to Cry2WT-mCherry (designated CWT 244–375) and transfected into SH SY5Y cells, blue light did not induce the cytoplasmic condensates characteristic ofLLPS. In contrast, the PRD domain fused to Cry2WT-mCherry (designated CWT 151–254) expressed in SH SY5Y underwent condensation within 5 s of blue-light activation (Fig. 2, A and B). Cry2WT-mCherry alone expressed in SH SY5Y cells at a com- parable concentration did not cluster following blue-light acti- vation (Fig. 2 B; Lee et al., 2014; Taslimi et al., 2014). The extent of light-activated condensation of PRD was quantified by com- puting the pixel CV, which evaluates the overall distribution across the cell (Fig. 2 B). To probe the bulk exchange dynamics of light-induced CWT 151–254 condensates, we performed FRAP by bleaching the mCherry signal (Fig. 2, C and D;and Video 1). CWT 151–254 can exhibit up to 55 ± 15% (n = 3) fluorescence recovery. Droplets formed by blue-light activation of CWT 151–254 disassembled within a few minutes in the absence of blue light, then re- assembled when exposed again to blue light (Fig. 2, E and F), indicating that droplet formation of this construct is initially reversible. Local activation ofCWT 151–254 by directing the light to a 3 × 3-μM square demonstrated that phase separation can be spatially localized (Fig. 2 G and Video 2 for local field activation). As expected for light-activated phase separation, Tau condensation likely involves weak multivalent interactions that can be strengthened and aided in their detection by Cry2 (Bracha et al., 2019). Taken together, these data show that the PRD can undergo light-induced condensation independent of the MTBD. We sought to determine the reversibility of Tau PRD condensates over longer time periods by cycling through multiple illumination periods. A sequence ofblue light exposure followed by a recovery phase with three activation cycles was performed (Fig. 3 A; Shin et al., 2017). After the first cycle, the

condensates were mostly dissolved, but residual condensates increased progressively after each cycle. The result was quantified by counting the condensates and normalizing the result (see Ma terials and methods). The percentage of condensates that remained was 0%, 42%, and 73% after the first, second, and third cycle, respectively (Fig. 3 B). The kinetics of cluster assembly and disassembly were also analyzed by pixel CV (Fig. 3 C). As expected, the values increased during assembly and decreased during disassembly. Through the recovery, the rate of the disassembly gradually slowed and eventually reached a plateau. Upon each cycle, the recovery plateaued progressively earlier and further from the initial state. These findings resemble an "aging" process in which the liquid state transitions to a non-exchangeable gel or a solid phase (Shin et al., 2017). The PRD is the most heavily phosphorylated region in Tau, and its phosphorylation is controlled by numerous kinases in diverse signaling pathways (Martin et al., 2013). Phosphorylation in this region will modulate the charge distribution and affect the weak electrostatic interactions that mediate condensation. Placing phosphomimetic mutations within the PRD completely abolished blue light–induced phase separation of CWT (Table 1). Mono-, di-, or triple-phosphomimetic mutations at T231E or S235E, T231E/S235E or T212E/S214E/T217E, or S202D/T205D/S208D all abolished light-induced phase separation. In contrast, LLPS of CWT 151–254 with alanine mutations at the same phosphorylation sites was not abolished. Upon light activation, CWT 151–254 S202A/T205A/S208A and CWT 151–254 S202A/T205A/S208A/T212A/S214A still phase separated.

PDR condensates promote Tau association with MTs

Having characterized the isolated domain behavior of the PRD, we next sought to determine its relationship to the MT and its behavior in the context of the neighboring MTBD. Light-induced PRD condensates (CWT 151–254) were frequently observed to align along MTs (Fig. 2 A and Fig. 4 A). The MTBD on its own (CWT 244–375) and transfected into SH SY5Y cells did not bind MTsnearly as well as full-length Tau(Fig. 4, B–E). In contrast, CWT 151–375 (Fig. 4, F and G), which contains both the PRD and the MTBD, bound to MTs similarly to full-length Tau (CWT 1-441) before activation and increased its binding further after light activation, also similar to full-length Tau (Fig. 4, B and C). Inclusion of the N-terminal domain with the PRD (CWT 1–254) had an inhibitory effect over Tau condensation. Unlike the PRD domain alone, this construct showed only minimal association with MTs both before and after light activation (Fig. S2, A and B). This result suggested that the N-terminal 1–150 region negatively regulates PRD phase separation, likely owing to its many negatively charged residues, and is consistent with the negative reg- ulation ofthe N-terminal region on PRD binding to soluble tubulin and slowing polymerization (McKibben and Rhoades, 2019)and previous work showing that removal of the N-terminal region increased the affinity of Tau for MTs (Gustke et al., 1994). As additional MTBDs were sequentially added to the CWT 1–254 construct, MT binding increased, and the diffuse labeling decreased (Fig. 1 C; Fig. S2, C–M;Table S1;and Materials and methods). A construct from the N terminus that included all of the MT-binding repeats (CWT 1–375) further increased association with MTs after blue light, but was still not maximized until we used a full-length construct that included the P3 (376–400) domain (Fig. S2, I and J). The P3 domain flanks the MTBD and also undergoes proline-directed phosphorylation (Mukrasch et al., 2005). These data demonstrate that condensation of

the PRD alone can drive association with MTs and that association is inhibited by the N-terminal domain but enhanced by the MTBD as well as the C-terminal domain.

In Silico Tau PRD phase separation

The spontaneous formation of liquid droplets from an aqueous solution of protein chains requires a specific set of conditions related to the chemical environment of the cell, as well as the nature of the intermolecular interactions according to our previous study of Tau LLPS *in vitro*[98,173]. Although the complete set of these conditions for phase separation is not yet known, one requirement is a sufficient number of charged amino acids spatially arranged to favor multivalent complexation of polyelectrolytes (modulated by Coulombic interactions) to drive assembly into microscopic droplets[174,175]. To understand how charge sequence controls the phase window for droplet formation, it is useful to employ a simple model and perform numerical simulations. To test if the proline-rich domain (Tau 151-254) has a charge pattern similar to the MTBD (Tau 255-441) that might lead to charge-driven phase separation[98,173] we computationally modeled Tau 151-254 as a bead-spring polymer with a charge sequence determined by the primary amino acid sequence of Tau at pH 7 in presence of RNA (Figure 7.1C). In terms of electrostatics, Tau 255-441 and Tau 151-254 have a similar propensity to form droplets, which is much greater than the N-terminal region or C-terminal region.

7.4 Results

7.4.1 *in vitro* Tau PRD phase separation

We next compared the biophysical properties of the PRD to the MT binding domain and the extreme C-terminus (Tau 255-441) *in vitro*. Tau 151-254 protein with a 6xHIS tag at the N-terminus was purified. As demonstrated previously, phase separation can occur readily

when utilizing RNA or heparin as a cofactor[98,176]. Both Tau 151-254 and Tau 255-441 similarly phase separated with the addition of the negatively charged cofactors, polyU or heparin (Figure 3A-C). The protein was mixed with cofactor in approximately a 1:1 charge ratio, and turbidity due to phase separation was evaluated by measuring the absorbance at 500 nm, and fibril formation measured by Thioflavin T (ThT) fluorescence. Directly after addition of polyU to Tau 151-254, droplets formed and were capable of fusing and deforming as they wet the glass surface. Tau 255-441 with heparin not only formed droplets, which fuse and deform as they wet the glass surface, but also induced significant ThT fluorescence, and by proxy, beta-sheet fibrillation[173]. In contrast to Tau 255-441, Tau 151-254 did not show any ThT fluorescence (Figure 7.1B). TEM imaging confirmed that no fibrils were formed by the Tau 151-254.

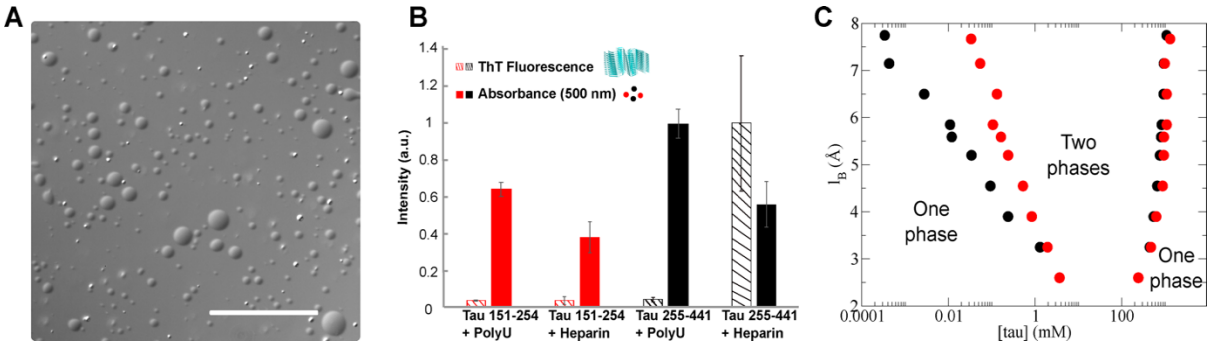


Figure 7.1 Similar to Tau 255–441, Tau 151–254 can form droplets in vitro. **(A)** Representative bright-field images of Tau 151–254/polyU droplets. Scale bar is 50 µm. **(B)** Direct comparison of Tau 151–254 (red) and Tau 255–441(black) shows phase separation occurs for both protein segments with addition of PolyU or heparin. In contrast to the Tau 255–441 heparin complex, Tau 151–254 does not show the capacity to form fibrils with either PolyU or heparin. Turbidity was evaluated at 500 nm (filled bar); fibrilization was evaluated by ThT fluorescence (striped bar). Protein concentration was 50 µM, PolyU concentration (where applicable) was 125 µg/ml, and heparin concentration was 12.5 µM. All data were normalized by the largest measured value or absorbance or fluorescence. n = 3; error bar in SD. **(C)** Predicted Tau 151–254/RNA binodal phase coexistence points modeled using FTS. Calculations performed for Tau 151–254 (red dots) are compared with our previously reported (Lin et al., 2019) phase diagram for Tau 255–441 (black dots), showing that both sequences have a stable two-phase region under similar electrostatic environments and solvent conditions. The conditions for stable binodal phase coexistence depend on the strength of the electrostatic interaction, characterized in our model by the l_B as the input

parameter. The region defined by the coexistence points is the region at which the system forms a thermodynamically stable droplet phase in coexistence with a solution phase. Areas corresponding to one-phase or two-phase are shown in the diagram.

7.4.2 BIN1 SH3 domain binds Tau PRD-polyU during phase separation *in vitro*

We analyzed the influence of the BIN1 SH3 domain upon Tau PRD phase separation *in vitro* with polyU RNA (Tau PRD-polyU). SH3 tagged with alexa-488 was incubated with Tau PRD tagged with alexa-405 in a 1:2 molar ratio. Phase separation was not observed; however, the addition of polyU RNA, which can easily phase separate with Tau PRD, now created a three-component phase separated system. The SH3 domains concentrated within Tau PRD-polyU droplets (Figure 7.2A). Fluorescence microscopy confirmed that SH3 was concentrated in the droplet phase and colocalized with the Tau PRD-dense phase. FRAP showed that the SH3 domain is mobile, and diffused between the droplet phase and the solute phase (Figure 7.2B).

Evidence for the BIN1 SH3 domain association with Tau PRD was obtained from continuous wave electron paramagnetic resonance (CW EPR) using a spin-labeled variant of SH3 (SH3-SL), see materials and methods, Figure 7.2C. CW EPR showed a distinct lineshape change towards an immobilized labeled state when BIN1 SH3 bound Tau PRD, even in the absence of droplets due to the omission of polyU RNA. Fitting of the lineshape showed 71% of the SH3-SL exhibited a bound component when incubated with BIN1 SH3:Tau PRD in a 1:2 molar ratio. Upon addition of polyU, condensation occurred, but with little change in the bound state of BIN1 SH3 (Figure 7.2D). Thus, while the system went through a phase separation, the CW EPR signature of the immobilized BIN1 SH3 domain remained unchanged. Indeed, the majority of BIN1 SH3 was found in the protein-dense droplets bound to TauPRD.

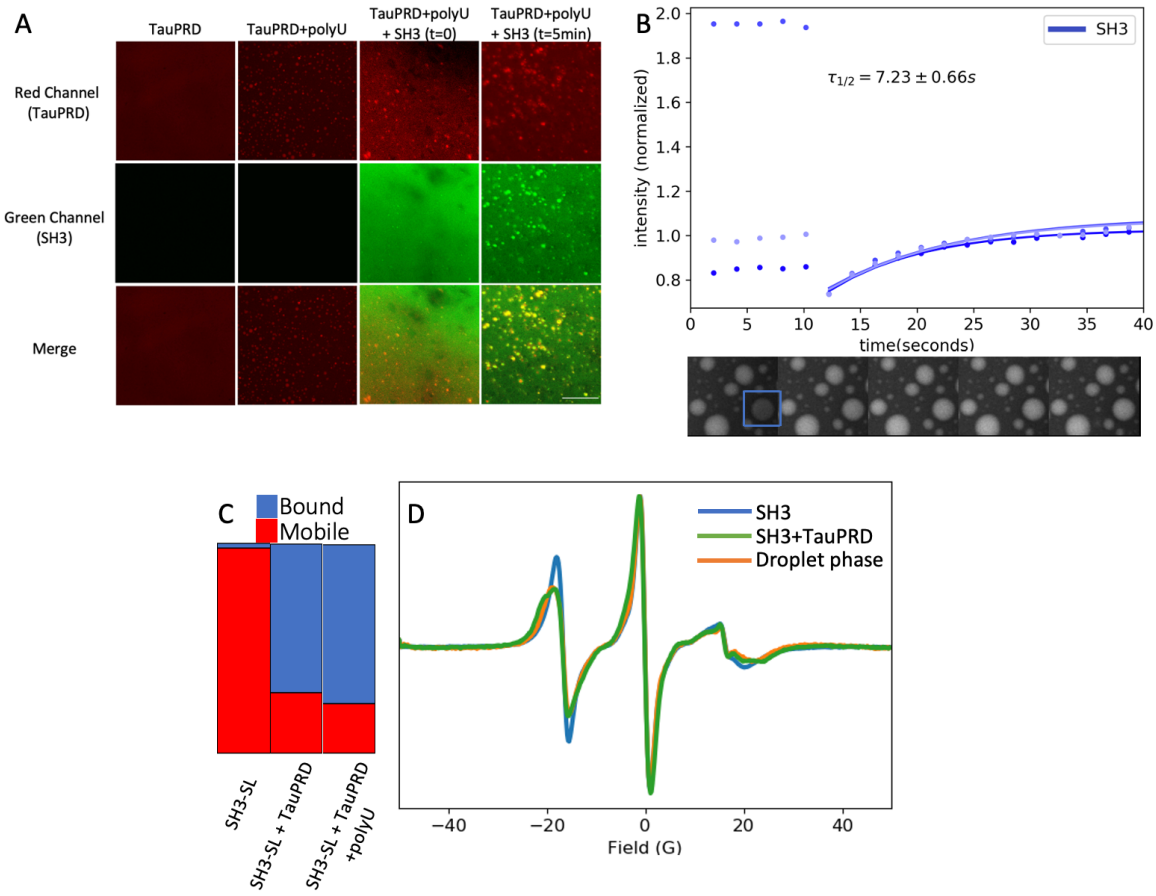


Figure 7.2: A) TauPRD-405 (column 1) was mixed with polyU to form a coacervate phase (column 2). SH3-488 was added at a 0.5 molar ratio of TauPRD, and the system was tracked for 5 minutes. When first added, SH3-488 is observed to be diffuse in the solution (column 3). After 5 minutes, the SH3-488 is observed to be highly colocalized with TauPRD-405 in the phase-separated droplets. **C)** CW EPR data was fit to a unbound (mobile) component, and a bound slow component for each sample shown in panel B. The percent of SH3-SL that is mobile is shown in red, and the percent of SH3-SL bound to TauPRD is shown in blue. Approximately 70% of SH3 is bound to Tau, regardless of the presence of a phase-separated state. **D)** SH3-SL (blue), shows a distinctly altered derivative CW EPR lineshape in the presence of TauPRD. A broadened first peak indicates a slower tumbling rate indicating SH3 binding to Tau. Little difference is observed between TauPRD and SH3 (green) and the TauPRD-polyU coacervate system when SH3 is present (yellow), suggesting SH3 remains bound to TauPRD in the coacervate phase.

A 3-component phase-separated system was observed containing TauPRD, BIN1 SH3 and polyU RNA, the BIN1 SH3 protein did not enhance the phase separation and was instead found to inhibit the formation of TauPRD-polyU condensates in a concentration dependent manner (Figure 7.3). At the 1:2 BIN1 SH3:TauPRD ratio used in the above experiments, the system remained phase separated (Figure 7A), but as the BIN1 SH3 concentration was increased, a concentration-dependent decrease in the number of condensates was observed (Figure 7.3), and as the BIN1 SH3 concentration reached equimolar with the Tau PRD, phase separation was completely inhibited (Figure 7.3). These tests were repeated with full-length neuronal BIN1, and similar concentration dependent behavior was observed to decrease the amount of condensates in the system as BIN1 increased. The two-phase system was abolished at an equimolar ratio of BIN1 and Tau PRD. At greater than equimolar concentrations of BIN1, solid precipitates formed (Figure 7.3). Taken together, these results suggest that discrete molecular binding of the Tau PRD to the BIN1 SH3 domain is not perturbed by condensation; however, the amount of the SH3 domain can determine whether the SH3 resides in the droplet, possibly as a client, or disrupts the large multivalent complexes involved in condensation.

7.5 Discussion

The discovery that Tau is capable of LLPS *in vitro*[98,177–181] raised the question of whether similar phase states of Tau are present in living cells. The use of expression constructs and optogenetic protein clustering can reveal potential protein domain functions within the environment of a living cell and is an initial step toward understanding the staggering complexity of the *in vivo* setting. Although both the full-length Tau and MTBD are capable of phase separation *in vitro*[182,183]**,** when expressed in cells, these constructs did not form LLPS droplets under any of a variety of conditions. Therefore, we amplified the weak multivalent

interactions of Tau in cells by fusing it to a light-activatable self-associating Cry2(WT) motif while ascertaining that Cry2(WT) alone did not phase separate under the same conditions. The fusion of "sticky" intrinsically disordered domains (IDR) to the photolyase homology region (PHR) of Arabidopsis thaliana Cry2[184] mimicks the LLPS of multivalent RNA binding proteins[185]. As shown previously for the IDR of FUS fused to Cry2WT as well as HNRNPA1 or the N-terminal IDR of DDX4[185], an IDR-Cry2 fusion protein can recapitulate the modular domain architecture of many phase separated proteins and confer tunable light-dependence to its multivalent interactions[185]. This approach identified the unique function of the Tau PRD (151-254) in driving condensation under a variety of experimental conditions. Focused blue light at specific subcellular locations can locally induce CWT 151-254 phase separation (Figure 1G), which undergoes dynamic exchange with the surrounding solution as assessed by fluorescence recovery after photobleaching (FRAP) experiments (Figure 1C-D). The multivalent interactions of these condensates were controlled by the charge of common phosphorylation sites, suggesting Tau condensates can be tuned through the complex phosphorylation of Tau at >30 sites[186] mediated by kinases that participate in multiple signaling pathways.

Light-induced condensation of Tau increased its association with MTs. Tau interactions with single tubulin alpha beta dimers have been identified by cryo-electron microscopy[29], but how Tau achieves the length scales necessary to interact with higher order tubulin structures is unknown. Orders of magnitude separate the nanometer dimensions of single Tau molecules from the micrometer dimensions of MT segments. Droplet size may scale more closely to MT segments. Furthermore, if Tau were uniformly distributed in the solute phase of the cell cytosol, it is unlikely that Tau could maintain the density of MT occupancy

over the MT surface given its rapid exchange between bound and free forms and a dissociation constant in the μM range. The condensed state of Tau in droplets increases the local concentration at specific MT loci to skew the on-off kinetics toward binding and the viscosity of droplets in proximity to MTs will retard cytoplasmic diffusion far from the MT surface when is not bound. Analogously to LLPS of actin regulatory protein association with the membrane[187], Tau condensates may increase MT dwell time. Finally, droplets with their ability to regulate interfacial tension can wet the solid MT surface. Interestingly, the PRD further optimizes MT binding due its ability to bind MTs[188] along with the MTBD (R1-R4). In support of this model is an *in vitro* study[189] that shows Tau initially bound diffusely along the entire microtubule (MT) lattice, but over time denser regions with the ability to FRAP (fluorescence recovery after photobleaching), fuse and dissolve on addition of 1,6-hexanediol (1,6-HD) appear gated by the nucleotide state of the MT lattice.

In addition to its association with MTs, the Tau PRD can operate in condensate formation with other proteins in the cytoplasm. Tau interacts with the MT end binding, EB1[190] in a complex with other +TIP proteins. Upon blue light activation, CWT 151-254 undergoes LLPS with EB1 with which it can collaborate in the control of MT dynamics. Through its SH3 binding domain, Tau has multiple potential interactors. One of them is BIN1[191], encoded by a gene that confers the second highest risk for AD after ApoE4[192]. The function of the BIN1-Tau complex is unknown but is thought to have a role in Tau propagation based on the finding that BIN1 is a negative regulator of endocytosis in neurons and the loss of BIN1 increases the flux to endosomes leading to internalized Tau aggregates[193]. MTs provide a competitive localization site with regard to condensation with other proteins such as BIN1, which as a condensate, can present to Tau multiple SH3 domains. Similarly, with the actin-associated

proteins nephrine-NCK-N-WASP, proline rich motifs and SH3 domains can form liquid-like droplets *in vitro* dependent on the number of SH3 domains present.

While expression studies such as this can reveal potential interactions, *in vivo* contingencies, such as stress, are sensed by multivalent molecules with adhesive domains or intrinsically disordered regions capable of condensing to serve a localized function, such as Tau does to regulate MTs. These heterotypic multivalent interactions may create membraneless bodies consisting of a scaffold and freely exchanging clients[194]. Contextualized in terms of phase states, much of Tau biological behavior will likely emerge.

7.6 Conclusion

The cellular milieu is an extremely complex environment that favors interactions and the formation of macromolecular complexes. LLPS condensates have attracted a lot of interest as a possible reversible way to create high local concentrations of proteins, and to sequester the protein out of the rest of the cytosol. Understanding Tau's interactions with other proteins in cells will be crucial to understand the complex mechanisms that drive neurodegeneration. Tau was known to phase separate *in vitro*, but there was little understanding of how cellular LLPS of Tau occurs, and what the local environment is like inside the condensates. Our collaborators identified BIN1 as a binding partner of Tau *in vivo* that colocalized in condensates.

In vitro work demonstrated Tau PRD to form LLPS condensates with an anionic cofactor of RNA or heparin. Previously, 2N4R Tau had been demonstrated to phase separate, and the MTBD of Tau had also been demonstrated to phase separate[195], but the PRD had not been implicated in the phase separation ability of Tau. Tau PRD was found to form LLPS,

independently of the MTBD indicating that the PRD may play important roles in Tau's binding and trafficking throughout the cell.

Chapter 8: Conclusion and Outlook

Disordered proteins and aggregation are fascinating and complex fields of research. Often, instead of clarifying a hypothesis, new evidence complicates our view and creates more questions than answers. Through the course of this work, we, and others, answered many questions about the mechanisms of Tau aggregation, but the solutions to prevent aggregation seem nearly as distant as they always have. This work described the effect of perturbations to Tau in multiple ways. First, we investigated how charged cofactors influence the structure of Tau filaments. Tau aggregated with heparin was shown to not reproduce the same structures as in Tauopathies, and to produce heterogenous populations of fibrils with multiple structures. Next, heparin and other cofactors were found to not be catalytic subunits to Tau aggregation. Instead, the cofactors are integral to the fibrils, and generate fibrils in stoichiometric quantities with Tau. The cofactors cannot be removed without also dissolving the fibrils of Tau.

We then turned our attention to truncations and mutations to Tau and their effects on aggregation. Because cleavages remove the conformational degrees of freedom of soluble Tau, we hypothesized that there would be a higher likelihood of aggregation into disease relevant Tau peptides. Two different length scales of Tau were studied to show that shorter forms of Tau are more aggregation prone and seeding competent. Products of proteolytic cleavages observed in AD patients were produced, and the resulting fragments were found to aggregate readily with heparin and to seed aggregation of full-length forms of Tau. The ability to recruit larger protein to the fibrils suggests that the proteolytic fragments may be a toxic seed that helps in the prion-like behavior of Tau.

Finally, a shorter fragment was studied with an additional hereditary P301L. This short peptide was more aggregation-prone than the equivalent WT form. We found differing

hydration shells can explain the difference in aggregation. The more hydrophobic leucine residue creates a lower entropy hydration shell that favors the dehydration required for amyloid aggregation. We propose a paradigm for amyloid aggregation in which hydrophobic residues drive an entropically favorable aggregation of Tau into amyloids. Additionally, cryoEM was used to show that the HP301L peptide exists in a hairpin shape similar to that observed in other Tauopathies in the fibrils formed *in vitro.*

It has a been a fascinating time to study Tau aggregation, and my hope is that these results are built upon in the coming years. Recent publications have replicated AD PHF structures with recombinant Tau[122], achieving one of the overarching goals of this work. As the knowledge base expands, replication of these results by other groups and the formation of other Tauopathy structures will be crucial. Once disease structures can be repeatedly produced *in vitro* there are many directions that should be studied. The failure of many fibril-targeting therapeutics may mean two things. First, treatments may not be targeting the correct fibril structures. If this is the case, the production of recombinant fibrils should make great strides to producing more effective therapeutics. Second, Tau fibrils may be a symptom of other dysregulation that has already progressed to a detrimental stage, and not be the correct target for therapeutics. Tau pathology is associated with dysregulation of lipid processing pathways[113,196,197], and in light of the results presented here implicating hydrophobicity as a driver of aggregation, investigations of Tau's interactions with lipids, and lipid processing pathways may provide insights into the initiation of Tau toxicity. The goal throughout this work is that our fundamental biophysical results will eventually be translated into relief for all of the patients suffering from Tauopathies.

Appendix A: Methods and Materials

Common Methods

ThT

Thioflavin T (ThT) assays for β-sheet content were conducted, unless otherwise noted with 50μM protein content unless otherwise noted, and 20mM ThT in a 384-well Corning plate. A BioTek synergy2 plate reader was used with temperature control set at 37˚C. Excitation used a 440nm filter, and emission was detected at 485nm.

TEM

Transmission Electron Microscopy (TEM) imaging was conducted with a 200kV FEI Tecnai G2 Sphera Microscope. Samples were blotted onto 200-mesh Formvar copper grids, and negative stained with uranyl acetate.

CW EPR

CW EPR measurements were carried out at room temperature with a Bruker EMX X-band spectrometer operating at 9.8 GHz (EMX; Bruker Biospin, Billerica, MA) and a dielectric cavity (ER 4123D; Bruker Biospin, Billerica, MA). A sample of 4.0 μL volume was loaded into a quartz capillary tube with 0.6 mm internal diameter (CV6084; VitroCom) and sealed at one end with critoseal, and then placed in the dielectric cavity for measurements. CW EPR spectra were acquired by using 6 mW of microwave power, 0.5 gauss modulation amplitude, 200 gauss sweep width, and 10 scans of signal averaging.

Spin counting

Spin-labeling efficiency was measured through quantitative CW EPR. A standard ladder of of 4-HydroxyTEMPO was measured at 50, 100, 200, 500 μM concentration and CW EPR spectra were recorded. The maximum of the baseline-corrected, integrated spectra was

recorded, and fit to a linear regression vs. concentration. All samples were measured under the same conditions, and the maximum of the baseline-corrected, integral spectra was converted to an effective spin label concentration. Spin labelling efficiency was calculated by SL_{eff}=[SL]/[Protein].

Multi-Component CW EPR spectra Fitting

Fitting of CW EPR was performed using MultiComponent, a software package from Christian Altenbach *The program is written in LabVIEW (National Instruments) and can be freely downloaded from the following site: http://www.chemistry.ucla.edu/directory/hubbell-wayne-l* . Fitting was done using similar procedures as described previously[198]. Briefly, the mobile component was determined by fitting the A, G and R tensors to the solution peptides. The A and G tensors were determined first, and then R was allowed to vary. Next a second component was fit to fibrils formed with XXX% spin-dilution and a single labelled site at V300C. These fibrils should display minimal exchange-coupling but exist in the slow-motion regime due to the size of the fibril. Fibrils were previously determined have an axially symmetric diffusion tensor ($\alpha_D = 0^\circ$, $\beta_D = 36^\circ$, $\gamma_D = 0^\circ$) and an ordering parameter of 20. Only the rotational correlation time (R) was fit to this component. Finally, a 3rd component was added to fit doubly-labeled peptides in fibrils. Component 3 used the same rotational correlation time (R) as the second component, as all spin labels in the fibrils should be in similar rotational correlation regimes. A Heisenberg spin-exchange component (ω_{ss}=140MHz) previously empirically determined for Tau fibrils was added to component 3. Solution state peptides were fit to only the mobile (component 1) and immobile (component 2) components. When fit with a 3rd, spin-exchanged, component the covariance of component 2 and component 3 was too large to obtain trustworthy fits.

RotamerConvolveMD simulations

The expected P(r) distribution for the structured hairpin formed in the CBD fibril core (PDB #6TJX) was predicted using the RotamerConvolveMD package[60,61]. An MTSL library from 2015 at 298K was used to simulate all possible conformations of the labels once tether to the defined structure extracted from #6TJX. To form the protein backbone all residues except 296-314 were deleted from #6TJX. Three chains (A, E, C) were retained, and the labels were attached to the termini of the middle chain (A) to account for clashes with adjacent chains.

Chapter 3 Methods

Text and Figures reprinted and adapted from Fichou, Y. *et. al*. (2018). Cofactors are essential constituents of stable and seeding-active Tau fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(52), 13234–13239. https://doi.org/10.1073/pnas.1810058115

Heparin spin labeling

Heparin sodium salt was spin-labeled with 4-amino-TEMPO through derivatization of the carboxyl group of the uronic acid residues of heparin. 0.5 mmol heparin was mixed with 0.9 mmol 4-amino-TEMPO, 0.1 mmol ethylenediamine, and 0.5 mmol N-hydroxysuccinimide on ice. 0.7 mmol EDC (1-ethyl-3-(3- dimethylaminopropyl)carbodiimide) was added and stirred on ice for 30 min. The solution was stirred for 7 h at room temperature, frozen at -20 ˚C, then lyophilized. Lyophilized samples were dissolved in 45 mL ethanol, centrifuged, and the supernatant was decanted. The pellet was washed with ethanol twice more and dried via vacuum, then dissolved in H2O and stored in freezer. Heparin sodium salt has a reported average molecular weight of 12-15 kDa. The resulting heparin was determined to contain an average of 1.2±0.2 spin labels per heparin molecule via cw-EPR double integrals.

DEER collection and analysis

Double Electron Electron Resonance (DEER). Tau187G272CS285CC291SC322S was expressed and spin labeled as Tau-SL2 which contains two spin labels in order to probe distances between residues 272 and 285. Tau187C291SC322S (Tau-cysless) was expressed as cysteine-less to avoid disulfide bonding. The protein stocks were concentrated, and buffer exchanged against D2O-based buffer (100 mM NaCl, 20 mM ammonium acetate, 1.5 mM CaCl2) using Amicon centrifugal concentrators (10 kDa cutoff). A 1:10 molar ratio Tau-SL2:Tau-cysless sample of 57 μM Tau-SL2 and 570 μM Tau-cysless was incubated with 157 μM heparin at room temperature for 24 h to prepare heparin fibrils (denoted as "before digestion" in Fig. 3B). Heparin fibrils were then incubated with 120 U of heparinase at 30 °C overnight to prepare digested heparin fibrils (denoted as "0.4× Heparinase" in Fig. 4B). 35 μL samples were mixed with 15 μL D8-glycerol (30% volume) before transferring to a quartz tube (2 mm i.d.) and frozen using liquid nitrogen. Seeded-fibrils sample (Fig. 4B) was prepared as follow: A 1:10 molar ratio Tau-SL2:Tau-cysless sample of 15 μM Tau-SL2 and 150 μM Tau-cysless was incubated at RT with 825 ug/ml polyU and 1% (protein mass) mouse brainextracted Tau fibrils. After 12h, the sample was pelleted and the fibrils solubilized in 35 μL buffer and 15 μL D8-glycerol before transferring to a quartz tube (2 mm i.d.) and freezing using liquid nitrogen. Four-pulse DEER experiments were carried out at 85 K using the Q-band Bruker E580 Elexsys pulse EPR spectrometer operating at ~34 GHz and equipped with a 300 W TWT amplifier. The following DEER pulse sequence was used: $\pi obs/2 - \tau 1 - \pi obs - (t-\pi pump) - (\tau 2-t) - \pi obs - \tau 2 -$ echo. Rectangular observe pulses were used with lengths set to $\pi obs/2$ = 10-12 ns and πobs = 20-24 ns. A chirp $\pi pump$ pulse was applied with a length of 20-24 ns and a frequency width of 133 MHz. The observe frequency was 150 MHz higher than the center of the pump frequency range. $\tau 1$ was 180 ns and $\tau 2$ was set between 1.8 ms and 2.4

ms. The DEER experiment was accumulated for ~12 h. The DEER signal was first baseline-corrected and followed by reconstruction of the distance distribution. For baseline correction, natural log of the DEER signal was taken and then linear baseline was obtained using the linear polynomial function. The DEER background signal is an exponential decay and it's logarithm provides a linear region from which the baseline can be easily obtained. A new Picard-Selected Segment-Optimized Singular Value Decomposition (or PICASSO) method (6) was applied to obtain distance distributions from the DEER signal, which is an ill-posed problem. The method finds the optimal distribution value at each distance or distance range by determining the different singular value cut-offs associated with the optimal solution at each distance using the Picard condition (6). The Picard condition informs when the solution diverges, allowing the selection of singular value cut-off before the divergence point. The method ensures optimal convergence at all distance ranges, preventing a premature or unstable solution as well as any spurious peaks. It also avoids negative values in distributions as they originate from premature and unstable solution. A singular value cut-off of 0.5-0.7 was used throughout for all datasets. Fig. S17 overlaps the experimental DEER signal with the signal calculated from the obtained distance distributions, for the 4 data sets presented in Fig. 3B and 4B. To calculate uncertainty in the distance distribution, an error analysis is conducted. The minimum and maximum values are calculated for singular value contributions where distance distributions are converged to the accurate solution and remain converged. The minimum and maximum values provide the uncertainty in the optimal distribution.

Chapter 4 Methods

Tau fragment expression and purification

The 8 fragments defined by combinations of the N-terminal and C-terminal cleavage sites were produced by recombinant expression in E.coli. cDNA for each Tau fragment was inserted into addgene plasmid #29663 with HiFi assembly. The resulting vector encodes a fusion protein consisting of a hexahistidine tag followed by eGFP, followed by a tobacco etch virus (TEV) enzymatic cleavage site attached to the desired fragment of Tau. The plasmid was transformed into DH5a, and BL21(DE3) cells. Each fragment was named with a number followed by a letter, with increasing number (1-4) corresponding to a shorter N-terminus, and a/b corresponding to the N368-K369 and Y394-K395 cleavages, respectively. *E. coli* BL21 (DE3) cells were transfected with constructed DNA variants and stored as frozen glycerol stock at −80°C. Cells from glycerol stock were grown in 10 mL luria broth (LB, Sigma Aldrich, L3022) overnight and then used to inoculate 1 L of fresh LB. Growth of cells were performed at 37°C, 200 rpm with addition of 10 µg/mL kanamycin (Fisher Scientific, BP906) until optical density at λ = 600nm reached 0.6-0.8. Expression was induced by incubation with 1 mM isopropyl-ß-D-thiogalactoside (Fisher Bioreagents, BP175510) for 3 hr. Cells were harvested with centrifugation at 5000 g for 30 min. Harvested cells were resuspended in lysis buffer (Tris-HCl, pH = 7.4, 100 mM NaCl, 0.5 mM DTT, 0.1 mM EDTA) added with 1 Pierce protease inhibitor tablet (Thermo Scientific, A32965), 1 mM PMSF, 2 mg/mL lysozyme, 20 µg/mL DNase (Sigma, DN25) and 10 mM $MgCl_2$(10 mM), and incubated on ice for 30 min. Samples were incubated at 30˚C for 20 minutes, then centrifuged at 10,000 rpm for 10 min to remove cell debris. 1 mM PMSF was added again and the resulting supernatant was incubated

for at least 4 h with Ni-NTA resins pre-equilibrated in buffer A (20 mM sodium phosphate, pH = 7.0, 500 mM NaCl, 10 mM imidazole, 100 μM EDTA).

The resin was loaded to a column and washed with 20 mL of buffer A, 25 mL buffer B (20 mM sodium phosphate, pH = 7.0, 1 M NaCl, 20 mM imidazole, 0.5 mM DTT, 100 μM EDTA). Purified protein was eluted with 15 mL of buffer C (20 mM sodium phosphate, pH = 7.0, 0.5 mM DTT, 100 mM NaCl, 300 mM imidazole), or until the resin no longer contained traces of GFP visible. The protein was concentrated to a volume of 2.5mL and was buffer exchanged into TEV buffer (50mM Tris buffer, pH7, 100mM NaCl, 50mM $CaCl_2$) to remove the imidazole, and prepare the protein for TEV cleavage. 35 μL of TEV (~1mg/mL stock) prepared in house was added to the solution and incubated at 4˚C overnight with gentle rotation.

To separate each fragment from GFP and TEV a cation exchange (GE CMFF sepharose column) separation followed by size exclusion chromatography using a Biorad superdex70 column. Cation exchange was conducted at pH 8 with a gradient of 0-1 M NaCl elution profile. All peaks showing a UV-Vis signal were collected and were concentrated to 250 μL. The SEC column was preequilibrated with 2 column volumes of working buffer, and the protein sample was loaded onto the sample loop. 0.5mL fractions were collected and run on an SDS page gel stained with Coomassie blue to determine the fractions containing Tau. All fractions containing Tau were pooled and concentrated to approximately 200 μM as determined by the UV-Vis absorption at 274 nm.

MALDI-tof

The purity, and identity of Tau fragments was confirmed with time-of-flight mass spectrometry. Protein samples were dialyzed overnight into pure H_2O. Samples were mixed with sinapinic acid, and the spectra was recorded with a Bruker Microflex LRF instrument.

All samples within 10 Da of the predicted protein weight, with no other clear species present were used for further studies.

Recombinant Tau Fibrilization

Fibrils of recombinant Tau fragments used in ThT and TEM assays were generated with 50 μM total protein concentration. 12.5 μM Heparin (Galen laboratory supplies, HEP001) was used for a 4:1 molar ratio of Tau:heparin. Samples were incubated at 37˚C in 384-well Corning plates. Fibrils used in HEK293T seeding assays were formed at 120 μM protein concentration, with 30 μM Heparin.

In Vitro Seeding and Thioflavin T experiments

Seeds for use in 2nd generation seeding studies were generated following the procedure described above at 50 μM protein concentration. The sample was vigorously pipetted to evenly disperse fibrils and was added at a 2.5 mol% (1.25 μM protein concentration) to a 50 μM sample of new protein. Aggregation was followed with ThioflavinT assays. Thioflavin T (ThT) assays for Beta-sheet content were conducted with 50 μM protein content, and 20 mM ThT in a 384-well Corning plate. A BioTek synergy2 plate reader was used with temperature control set at 37˚C.

Statistical analysis of ThT and seeding data

All reported ThT values are the increase in ThT from its minimum value near t=0. Any other normalization was conducted on data after this normalization. All analyses were conducted with a Mann-Whitney-Wilcoxon two-sided test with Bonferroni correction. Smaller p values were denoted with more stars in figures. *: 1.00e-02 < p <= 5.00e-02, **: 1.00e-03 < p <= 1.00e-02, ***: 1.00e-04 < p <= 1.00e-03, ****: p <= 1.00e-04

Fragments with the ‘a’ C-terminus showed significantly less Tht than fragments with the b C-terminus ($p=0.0005$). Different N-terminus construct a,b,c,d were found to be not significantly different ($p>0.1$) from each other, and have little effect on the total Tht observed.

Chapter 5 Methods

HP301 Sample Preparation

HP301P and HP301L monomeric/oligomeric solution mixtures were prepared similarly as TEM samples. Peptide samples were purchased in powder form and solvated and aliquoted in 20 mM Ammonium Acetate (NH_4Ac) buffer with a final concentration of 2 mg/mL. The peptide samples were stored in microcentrifuge tubes at 200 μM concentration in the freezer with -20/-80 ℃. Upon the experiment, 200 μM peptide solution were added into a certain volume of 20 mM NH_4Ac with heparin to make a final concentration of 25/50 μM of peptide and 4:1 concentration ratio of peptide to heparin. The peptide mixtures were then incubated at 37 ℃ and shaken at 650 rpm for 23 hr.

AFM Imaging

An aliquot of 4-5uL was drop-casted on a freshly cleaved mica surface (TedPella, Redding, CA) after the incubation period. The drop-casted solution was left on mica surface for about 1 min and then the excess liquid was blotted off and dried in a desiccator overnightThe topography and phase images were acquired by MFP-3D Atomic Force Microscopy (Asylum Research, Goleta, CA) in Tapping Mode Amplitude Modulation using high resolution silicon nitrate tip (NanoAndMore, CA), with a radius <10 nm and a cantilever with spring constant about 2 N/m and a resonant frequency around 70 kHz.

Fibril stability assay

Guanidinium Hydrochloride denaturation assays were conducted under the same conditions as ThT assays. Reported values were detected after signal reached equilibration at least 5 hrs after initial incubation. Fibrils were diluted 2x from initial concentration into the denaturation buffer to reach the reported GdnHCl concentration.

Spin-labeling

Peptides were functionalized with an a nitroxide free-radical, MTSL (S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate). MTSL was incubated at 10x molar concentration with peptides in 4M GdnHCl overnight at 4˚C. Labelled protein was passaged through two GE PD G-10 desalting columns and exchanged into 20 mM ammonium acetate buffer pH 7.0, with no salt present to prevent preliminary aggregation. Peptides used to measure DEER were labelled by mixing the DEER double mutant with 20x concentration of WT protein. MTSL was added in 10x molar excess and the solution was incubated at 4 ˚C for 1hr. Next, heparin was added in a 4:1 molar ratio (Tau:heparin) and the solution was incubated at 37 ˚C for 48 hrs. To remove unbound spin-labels, fibrils were concentrated using a 50 kDa cutoff filter, then diluted with buffer and concentrated again. This was repeated until at least 100x dilution of the original buffer was achieved.

ODNP

ODNP samples were prepared with 3.5 μL in a quartz capillary tube with 0.6 mm internal diameter (CV6084; VitroCom) and sealed at one end with Critoseal. The other end was sealed with beeswax. All measurements took place under 18.0±0.2 ˚C with constant

convective cooling to prevent sample heating. An NMR probe built in-house was used, and the sample was irradiated with microwave at the central electron hyperfine transition. NMR enhancement as a function of mw power was determined with a series of 20 mw power up to 6 W. A $T_1(p)$ enhancement curve was performed with a series of 5 inversion recovery experiments as previously described. $T_{1,0}$ was determined in separate experiments using unlabeled peptides to be ~2.0 s for all peptides used.

Data was processed using the hydrationGUI by Tom Casey, a package for processing ODNP data using DNPLab (https://thcasey3.github.io/hanlab/hydrationGUI.html, http://dnplab.net/). Most experiments used the automatic process function, but were manually inspected after processing for aberrations in the integration windowing, or phase correction. If an error was detected (such as a failure to correctly center the NMR peak, windowing and phase correction was performed manually, and a constant 10ppm window was used.

Peptide production

HP301 peptides were produced by Genscript to >95% purity with no additional modifications or capping beyond addition of spin labelled cysteines at the sites noted. Protein concentration was calculated from the 274nm absorbance recorded on a Shimadzu UV spectrometer. An extinction coefficient of 1.499 $cm^{-1}mM^{-1}$ was used for HP301P, HP301L, HP301LΔ(295-299), and HP301LΔ(295-299). Peptides were hydrated in 20mM ammonium acetate buffer pH7.4 to a concentration of 2 mg/mL and were immediately aliquoted and stored at -80 ˚C until use.

Chapter 7 Methods

In vitro Turbidimetry and brightfield microscopy

Turbidity of samples at room temperature measured by optical density at a 500 nm wavelength (OD_{500}), using a BioTek synergy2 plate reader. The amount of coacervates in a sample were approximated to be proportional to its maximum OD_{500}. Brightfield microscopy images of condensates were taken with a Leica BX51 or IX51 compound microscope.

In vitro Tau PRD and polyU condensate formation

Unless otherwise stated, condensates of Tau PRD with PolyU RNA were formed with 250 μM Tau PRD and 0.15 mg/mL polyU. This corresponds to a positive:negative charge ratio of 3:1. Ambient temperature and 20 mM HEPES, 40 mM NaCl pH7 buffer were used.

Recombinant protein labelling

The native cysteine in the SH3 domain was functionalized with an Alexa-488 (Thermo Fisher Scientific) fluorophore or a nitroxide free-radical, MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate). Tau PRD was labelled with Alexa-405 NHS Ester (Thermo Fisher Scientific). Protein was incubated with 5 mM DTT overnight at 4˚C, the DTT was removed through buffer exchange in a GE PD G-25 desalting column, and a 10x concentration of the labelling species was added. The sample was rotated gently overnight at 4˚C and unbound label was removed with a desalting column.

Fluorescence recovery after photobleaching for recombinant proteins

In vitro, FRAP of SH3 was conducted on a Leica Sp8 laser scanning confocal microscope at room temperature were bleached for 30 seconds at the excitation wavelength of the associated

fluorophore at 100% power. The entirety of a single droplet was bleached, and was allowed to recover for 30 seconds. Recovery data was fit to a three parameter, exponential function $I = a * e^{b*t} + c$ to establish a half-life recovery time. The reported $\tau_{1/2}$ was calculated from an average of the best-fit exponential coefficient (b) of three FRAP trials.

CW EPR measurement and data fitting

CW EPR measurements were carried out at room temperature with a Bruker EMX X-band spectrometer operating at 9.8 GHz (EMX; Bruker Biospin, Billerica, MA) and a dielectric cavity (ER 4123D; Bruker Biospin, Billerica, MA). All samples were measured with 125 μM SH3-SL. SH3-SL was estimated to have 90% spin label efficiency. A sample of 4.0 μL volume was loaded into a quartz capillary tube with 0.6 mm internal diameter (CV6084; VitroCom) and sealed at one end with critoseal, and then placed in the dielectric cavity for measurements. CW EPR spectra were acquired by using 6 mW of microwave power, 0.5 gauss modulation amplitude, 100 gauss sweep width, and 10 scans of signal averaging.

Data was analyzed using the program MultiComponent by Christian Altenbach. The program is written in LabVIEW (National Instruments) and can be freely downloaded from the following site: http://www.chemistry.ucla.edu/directory/hubbell-wayne-l. Values of the A and G tensors were established through fitting of a pure sample of Spin-labeled SH3 (SH3-SL), these values were fixed for all further components of fitting. The rotational correlation time tensor (R) was then fit to a single spectral component of the pure SH3-SL sample to establish the unbound component used for all fitting. The bound component spectra was then generated by fitting a sample of SH3-SL and Tau PRD with a two-component fit. The first component being the unbound component above, and the second component fit the R value of the bound component. The best fit was determined as a combination of the bound and unbound

components. Both components were then used to fit the soluble SH3-SL sample, and the sample containing SH3, Tau PRD and polyU in a coacervate phase. The ratio of unbound component and bound component was used as a measure of the extent of SH3 bound to Tau PRD.

Appendix B: Supplementary Material

Chapter 2 Supplementary Figures and discussion

Estimating the distance between C291 and C322

Fitzpatrick *et. Al*[38] reported the presence of additional electron density along the R3 region. They suggest that, in the case of 4R Tau, it originates from R2 forming an anti-parallel β–sheet between residues 290-305 and the R3 region of the paired helical filaments. Electron density maps (shown in Figure S2.3) depict the C-terminal end of the β–sheet to be L315. C322 is measured to be 3.4 nm away from C291 based on this result.

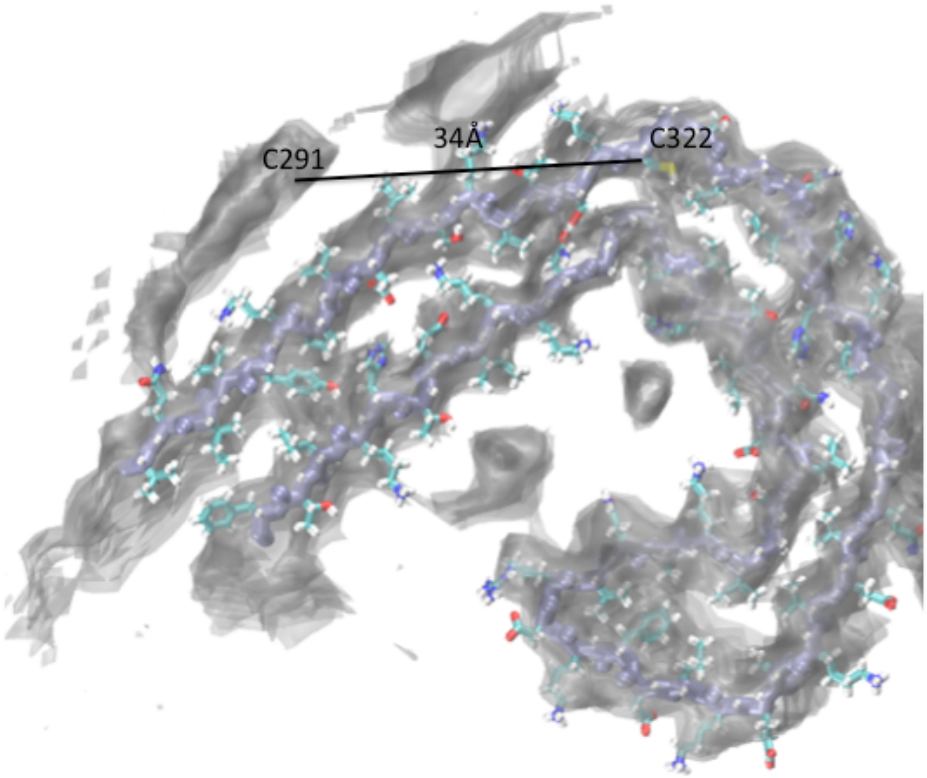


Figure S2.4: Estimation of the distance between C291 and C322 in 4R, PHF Tau is found to be 34Å.

Simulations of solution-state ensemble

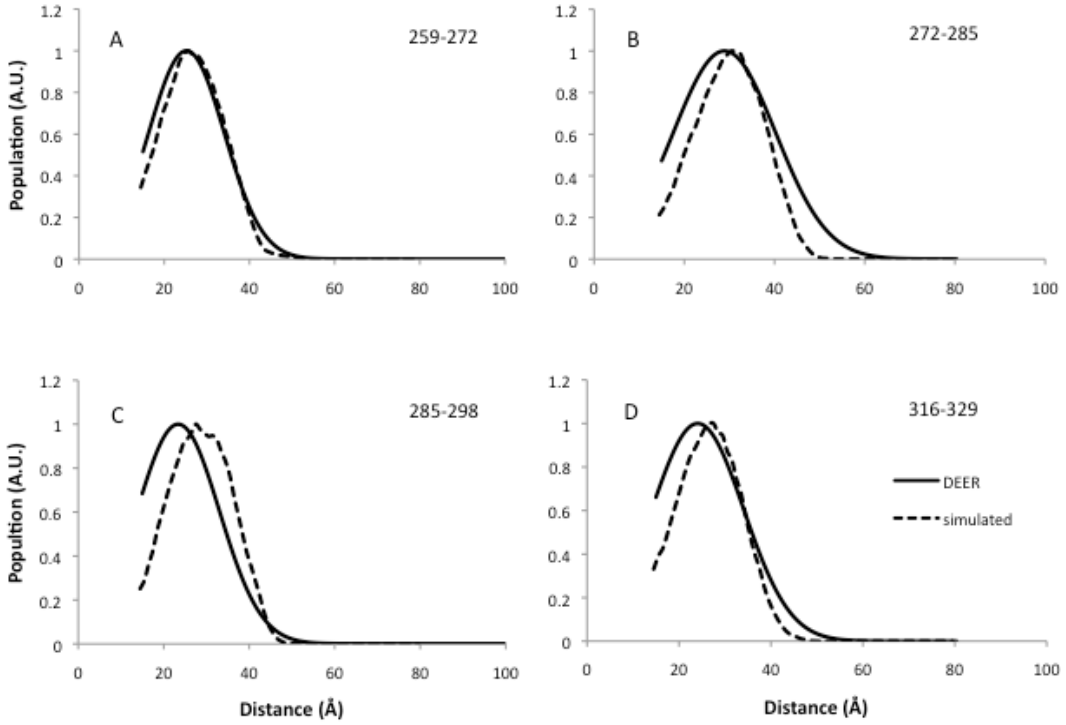


Figure S2.5: RotamerConvolveMD simulations of the expected distance distribution of MSTL spin labels in the ensemble of Tau conformers (dashed line) shows good agreement with DEER measurements (solid line). Residue pair numbers are annotated in the upper right corner of each graph.

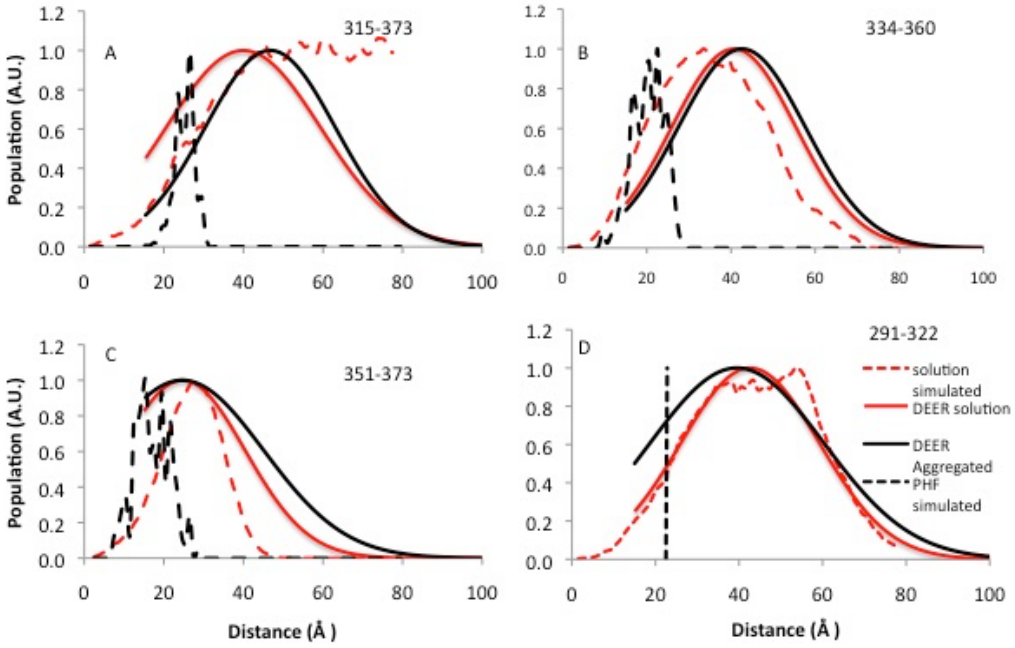


Figure S2.6: RotamerConvolveMD simulations of the expected distance distribution of MSTL spin labels in the ensemble of Tau conformers (dashed red) and DEER measurements of unaggregated Tau (solid red). A significant discrepancy is visible for sites 315-373 as the calculated large distances fall above the accessible DEER range. Simulated (dashed black) and measured (solid black) distance distributions for AD filaments (PHF and SF) and heparin filaments, respectively, are plotted as well. Residue pair numbers are annotated in the upper right corner of each graph.

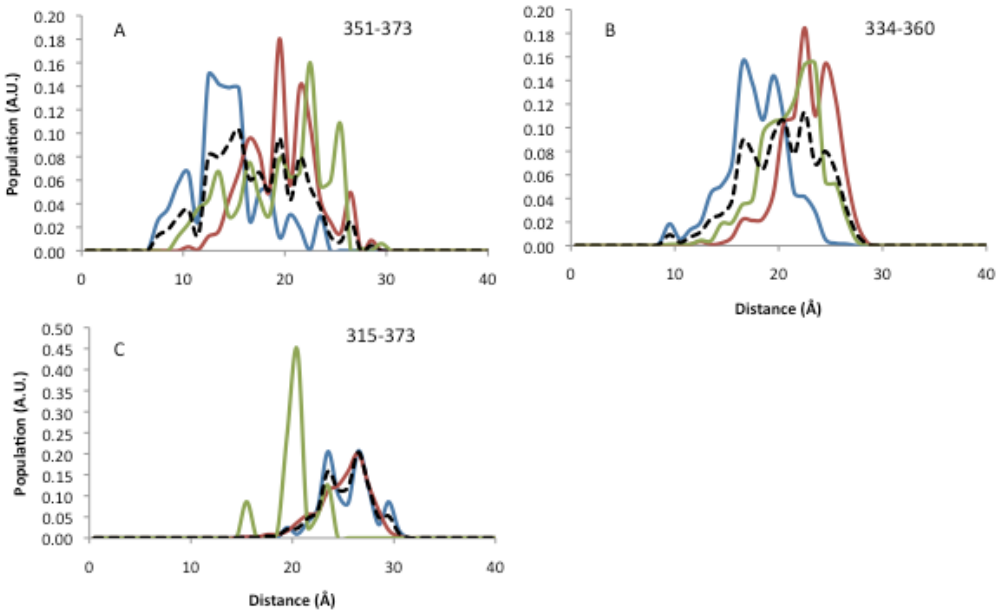


Figure S2.7: DEER distance distribution **s**imulated for PHF (blue), SF (red), pronase treated PHF (green), and a 1:1 mixture of SF and PHF (dashed black). A mixture of PHF and SF provides a broader distribution than any single conformation, and for comparison to DEER measurements throughout the text.

Chapter 3 Supplementary Information

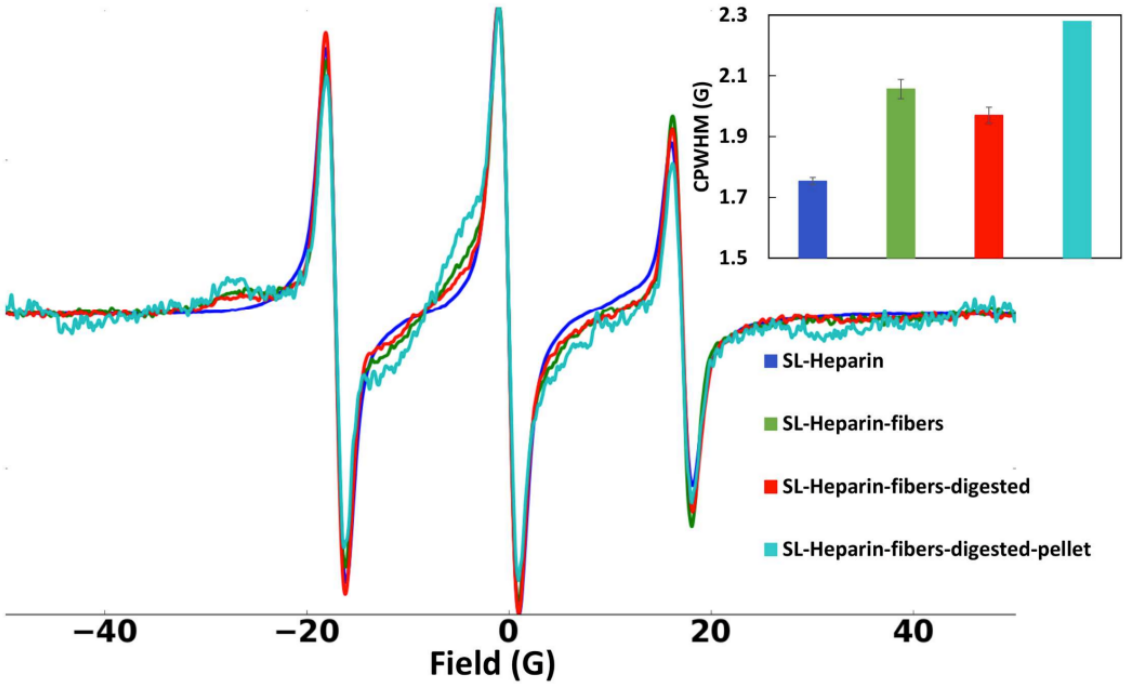


Figure S3.1 cw-EPR lineshape analysis of heparin-SL. Changes in cw-EPR lineshape show the evolution of heparin mobility with respect to aggregation and digestion. (inset) The central peak width half-maximum (CPWHM) of digested and undigested fibrils, before and after filtration shows the relative mobility of heparin. Soluble heparin is the most mobile species (dark blue), and upon fibrillation, the heparin becomes less mobile as indicated by a broader peak (green). Digestion increase the average mobility indicated by a narrower peak (red), but not to the same point than free heparin, suggesting that a portion of bound heparin is released upon digestion. When this digested sample is pelleted to select only fibril (blue), the peak become even broader showing that the remaining fibrils after digestion contain bound heparin.

Chapter 5 Supplementary information

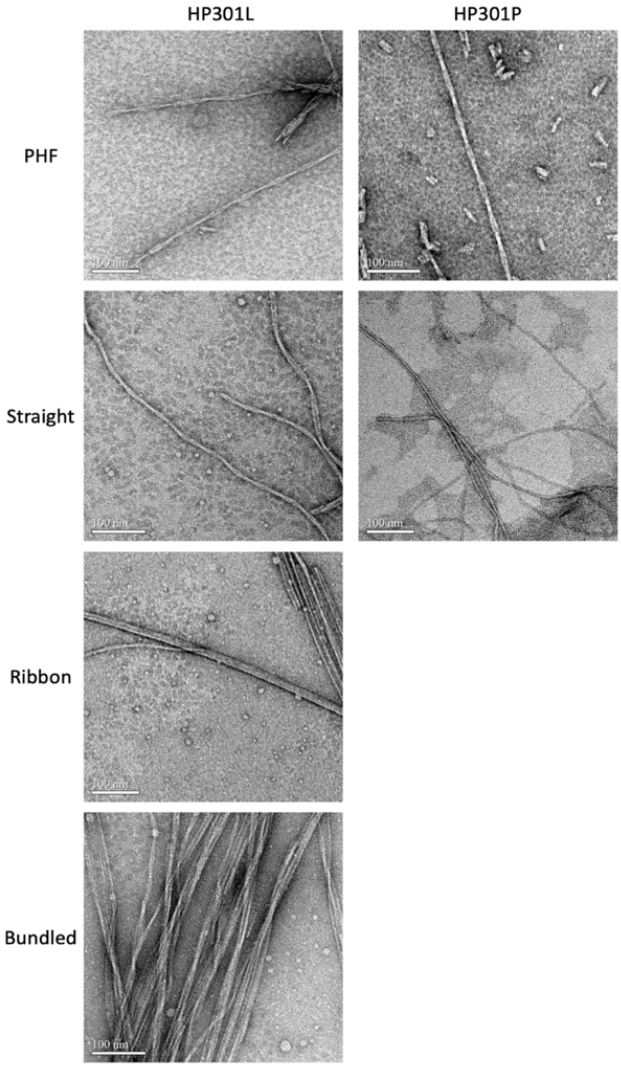


Figure S4.1: Example morphologies observed in HP301P and HP301L fibrils populations

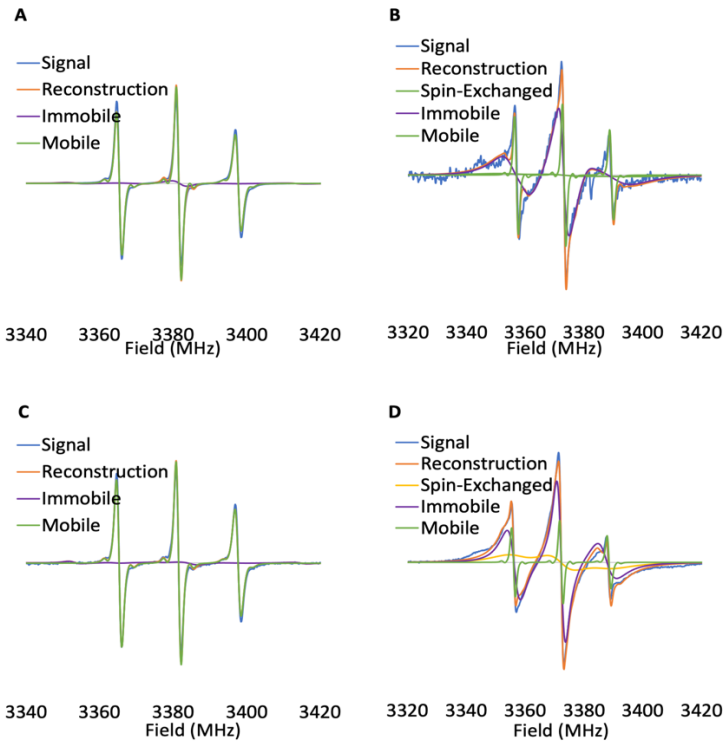


Figure S4.2: EPR spectra used for fitting of spin components. A) HP301P B) HP301P fibers C) HP301L D) HP301L fibers.

References

1. Uversky, V. N. Biophysical Methods to Investigate Intrinsically Disordered Proteins: Avoiding an “Elephant and Blind Men” Situation. in *Advances in Experimental Medicine and Biology* vol. 870 215–260 (2015).
2. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
3. Zarin, T. *et al.* Proteome-wide signatures of function in highly diverged intrinsically disordered regions. *Elife* **8**, 1–27 (2019).
4. Ferreon, A. C. M., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390–394 (2013).
5. Uversky, V. N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Curr. Opin. Struct. Biol.* **44**, 18–30 (2017).
6. Kim, J. Y., Meng, F., Yoo, J. & Chung, H. S. Diffusion-limited association of disordered protein by non-native electrostatic interactions. *Nat. Commun.* **9**, 1–10 (2018).
7. Keul, N. D. *et al.* The entropic force generated by intrinsically disordered segments tunes protein function. *Nature* **563**, 584–588 (2018).
8. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases : Introducing the D 2 Concept. **2**, (2008).
9. Dobson, C. M., Swoboda, B. E. P., Joniau, M. & Weissman, C. The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 133–145 (2001).
10. De Simone, A. *et al.* Intrinsic disorder modulates protein self-assembly and aggregation. doi:10.1073/pnas.1118048109/-/DCSupplemental.
11. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
12. Chebaro, Y., Ballard, A. J., Chakraborty, D. & Wales, D. J. Intrinsically disordered energy landscapes. *Sci. Rep.* **5**, 1–12 (2015).
13. Tolnay, M. & Probst, a. REVIEW: tau protein pathology in Alzheimer’s disease and related disorders. *Neuropathol. Appl. Neurobiol.* **25**, 171–187 (1999).
14. Goedert, M. Alzheimer’s and Parkinson’s diseases: The prion concept in relation to assembled Aβ, tau, and α-synuclein. *Science (80-.).* **349**, (2015).
15. Association, A. *Alzheimer’s Disease Facts and Figures. Alzheimers Dement*. (2022).
16. Wong, W. Economic burden of Alzheimer disease and managed care considerations. *Am. J. Manag. Care* **26**, S177–S183 (2020).
17. Nelson, R. & Eisenberg, D. Recent atomic models of amyloid fibril structure. *Curr. Opin. Struct. Biol.* **16**, 260–265 (2006).
18. Knowles, T. P. *et al.* Role of Intermolecular Forces in Protein Nanofibrils. **318**, 1900–1904 (2007).
19. Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L. C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 315–320 (2005).
20. Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K. & Caporini, M. A. Atomic structure and hierarchical assembly of a cross- β amyloid fi bril. **5590**,

(2013).
21. Jahn, T. R. *et al.* The Common Architecture of Cross-β Amyloid. *J. Mol. Biol.* **395**, 717–727 (2010).
22. Makin, O. S. & Serpell, L. C. Structures for amyloid fibrils. *FEBS J.* **272**, 5950–5961 (2005).
23. Knowles, T. P. J. *et al.* Kinetics and thermodynamics of amyloid formation from direct measurements of fluctuations in fibril mass. (2007).
24. Tycko, R. Physical and structural basis for polymorphism in amyloid fibrils. **23**, 1528–1539 (2014).
25. Holmes, B. B. *et al.* Proteopathic tau seeding predicts tauopathy in vivo. *Proc. Natl. Acad. Sci.* **111**, (2014).
26. Frost, B. & Diamond, M. I. Prion-like mechanisms in neurodegenerative diseases. *Nat. Rev. Neurosci.* **11**, 155–159 (2010).
27. Kaufman, S. K. *et al.* Tau Prion Strains Dictate Patterns of Cell Pathology, Progression Rate, and Regional Vulnerability In Vivo. *Neuron* **92**, 796–812 (2016).
28. Wang, Y. & Mandelkow, E. Tau in physiology and pathology. *Nat. Rev. Neurosci.* **17**, 5–21 (2016).
29. Kellogg, E. H. *et al.* Near-atomic model of microtubule-tau interactions. *Science (80-.).* **360**, 1242–1246 (2018).
30. Mandelkow, E. & Mandelkow, E. Biochemistry and Cell Biology of Tau Protein in Neurofibrillary Degeneration. *Cold Spring Harb Perspect Med* **2**, (2012).
31. Stelzmann, R. A., Norman Schnitzlein, H. & Reed Murtagh, F. An english translation of alzheimer's 1907 paper, ?Uber eine eigenartige erkankung der hirnrinde? *Clin. Anat.* **8**, 429–431 (1995).
32. Small, D. H. & Cappai, R. Alois Alzheimer and Alzheimer's disease: a centennial perspective. *J. Neurochem.* **99**, 708–710 (2006).
33. Crowther, R. A. Straight and paired helical filaments in Alzheimer disease have a common structural unit. **88**, 2288–2292 (1991).
34. Crowther, R. A. & Wischik, C. M. Image reconstruction of the Alzheimer paired helical filament. *EMBO J.* **4**, 3661–3665 (1985).
35. Wischik, C. M., Crowther, R. A., Stewart, M. & Roth, M. Subunit Structure of Paired Helical Filaments in Alzheimer ' s Disease. 1905–1912 (1985).
36. Kosik, K. S., Joachim, C. L. & Selkoe, D. J. Microtubule-associated protein tau (τ) is a major antigenic component of paired helical filaments in Alzheimer disease. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 4044–4048 (1986).
37. Neve, R. L., Harris, P., Kosik, K. S., Kurnit, D. M. & Donlon, T. A. Identification of cDNA clones for the human microtubule-associated protein tau and chromosomal localization of the genes for tau and microtubule-associated protein 2. *Mol. Brain Res.* **1**, 271–280 (1986).
38. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer ' s disease. *Nat. Publ. Gr.* **547**, 185–190 (2017).
39. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer ' s disease. *Nat. Publ. Gr.* (2017) doi:10.1038/nature23002.
40. Irwin, D. J. Tauopathies as clinicopathological entities. *Park. Relat. Disord.* **22**, S29–S33 (2016).
41. Schöll, M. *et al.* PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron*

89, 971–982 (2016).
42. Shi, Y. *et al.* Structure-based Classification of Tauopathies. *bioRxiv* 1–27 (2021).
43. Rauch, J. N. *et al.* LRP1 is a master regulator of tau uptake and spread. *Nature* **580**, 381–385 (2020).
44. Al-Hilaly, Y. K. *et al.* Tau (297-391) forms filaments that structurally mimic the core of paired helical filaments in Alzheimer's disease brain. *FEBS Lett.* **594**, 944–950 (2020).
45. Zhang, W. *et al.* Novel tau filament fold in corticobasal degeneration. *Nature* **580**, 283–287 (2020).
46. Falcon, B. *et al.* encephalopathy encloses hydrophobic molecules. *Nature* doi:10.1038/s41586-019-1026-5.
47. Friedhoff, P., Schneider, A., Mandelkow, E. M. & Mandelkow, E. Rapid assembly of Alzheimer-like paired helical filaments from microtubule-associated protein tau monitored by fluorescence in solution. *Biochemistry* **37**, 10223–10230 (1998).
48. Fichou, Y. *et al.* The elusive tau molecular structures: Can we translate the recent breakthroughs into new targets for intervention? *Acta Neuropathol. Commun.* **7**, 1–17 (2019).
49. Weil, J. A. & Bolton, J. R. *Electron Paramagnetic Resonance*. (Wiley-Interscience, 2007).
50. Bordignon, E. EPR Spectroscopy of Nitroxide Spin Probes. **6**, 235–254 (2017).
51. Stoll, S. & Britt, R. D. General and efficient simulation of pulse EPR spectra. *Phys. Chem. Chem. Phys.* **11**, 6614–6625 (2009).
52. Mitchelle S. de Vera, I., Blackburn, M., Luis, G. & Fanucci, G. *Pulsed EPR Distance Measurements in Soluble Proteins by Site- directed Spin-labeling (SDSL)*. *Curr Protoc Protein Sci.* vol. 74 (2015).
53. Goedert, M., Crowther, R. A. & Spillantini, M. G. Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature* **383**, (1996).
54. Sanders, D. W. *et al.* Distinct Tau Prion Strains Propagate in Cells and Mice and Define Different Tauopathies. *Neuron* **82**, 1271–1288 (2014).
55. Kaufman, S. K., Thomas, T. L., Tredici, K. Del, Braak, H. & Diamond, M. I. Characterization of tau prion seeding activity and strains from formaldehyde- fixed tissue. 1–12 (2017) doi:10.1186/s40478-017-0442-8.
56. Falcon, B. *et al.* Conformation Determines the Seeding Potencies of Native and Recombinant Tau Aggregates. *J. Biol. Chem.* **290**, 1049–1065 (2015).
57. Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **63**, 419–446 (2012).
58. Fichou, Y., Eschmann, N. A., Keller, T. J. & Han, S. Conformation-based assay of tau protein aggregation. in 89–112 (2017). doi:10.1016/bs.mcb.2017.06.008.
59. Schwalbe, M. *et al.* Predictive Atomic Resolution Descriptions of Intrinsically Disordered hTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. *Structure* **22**, 238–249 (2014).
60. Stelzl, L. S., Fowler, P. W., Sansom, M. S. P. & Beckstein, O. Flexible Gates Generate Occluded Intermediates in the Transport Cycle of LacY. *J. Mol. Biol.* **426**, 735–751 (2014).
61. Polyhach, Y., Bordignon, E. & Jeschke, G. Rotamer libraries of spin labelled cysteines

for protein studies w. 2356–2366 (2011) doi:10.1039/c0cp01865a.
62. Khan, S. *et al.* The Solution Structure of Heparan Sulfate Differs from That of Heparin. *J. Biol. Chem.* **288**, 27737–27751 (2013).
63. Khorramian, B. A. & Stivala, S. S. Small-angle x-ray scattering of high- and low-affinity heparin. *Arch. Biochem. Biophys.* **247**, 384–392 (1986).
64. Eschmann, N. A. *et al.* Signature of an aggregation-prone conformation of tau. *Nat. Publ. Gr.* 1–10 (2017) doi:10.1038/srep44739.
65. Bibow, S. *et al.* The Dynamic Structure of Filamentous Tau. *Angew. Chemie Int. Ed.* **50**, 11520–11524 (2011).
66. Sillen, A. *et al.* Regions of Tau Implicated in the Paired Helical Fragment Core as Defined by NMR. *ChemBioChem* **6**, 1849–1856 (2005).
67. Pavlova, A. *et al.* Protein structural and surface water rearrangement constitute major events in the earliest aggregation stages of tau. *Proc Natl Acad Sci U S A* **113**, E127-36 (2016).
68. Shammas, S. L. *et al.* A mechanistic model of tau amyloid aggregation based on direct observation of oligomers. *Nat. Commun.* **6**, 1–10 (2015).
69. Ramachandran, G. & Udgaonkar, J. B. Understanding the Kinetic Roles of the Inducer Heparin and of Rod-like Protofibrils during Amyloid Fibril Formation by Tau Protein. *J. Biol. Chem.* **286**, 38948–38959 (2011).
70. Novak, P., Prcina, M. & Kontsekova, E. Tauons and Prions : Infamous Cousins ? **26**, 413–430 (2011).
71. Zhang, W. *et al.* Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. 1–33 (2018) doi:10.1101/468892.
72. Greenwald, J. & Riek, R. Biology of Amyloid: Structure, Function, and Regulation. *Structure* **18**, 1244–1260 (2010).
73. Jahn, T. R. & Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **272**, 5962–5970 (2005).
74. Almeida, M. R. & Saraiva, M. J. Clearance of extracellular misfolded proteins in systemic amyloidosis: Experience with transthyretin. *FEBS Lett.* **586**, 2891–2896 (2012).
75. Mandelkow, E.-M. & Mandelkow, E. Biochemistry and Cell Biology of Tau Protein in Neurofibrillary Degeneration. *Cold Spring Harb. Perspect. Med.* **2**, a006247–a006247 (2012).
76. Goedert, M. *et al.* Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature* **383**, 550–553 (1996).
77. 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.* **399**, 344–349 (1996).
78. Wilson, D. M. & Binder, L. I. Free fatty acids stimulate the polymerization of tau and amyloid beta peptides. In vitro evidence for a common effector of pathogenesis in Alzheimer's disease. *Am. J. Pathol.* **150**, 2181–95 (1997).
79. Dinkel, P. D., Holden, M. R., Matin, N. & Margittai, M. RNA Binds to Tau Fibrils and Sustains Template-Assisted Growth. *Biochemistry* **54**, 4731–4740 (2015).
80. Meyer, V., Dinkel, P. D., Rickman Hager, E. & Margittai, M. Amplification of Tau Fibrils from Minute Quantities of Seeds. *Biochemistry* **53**, 5804–5809 (2014).
81. Carlson, S. W. *et al.* A Complex Mechanism for Inducer Mediated Tau

Polymerization. *Biochemistry* **46**, 8838–8849 (2007).
82. Sibille, N. *et al.* Structural Impact of Heparin Binding to Full-Length Tau As Studied by NMR Spectroscopy. *Biochemistry* **45**, 12560–12572 (2006).
83. von Bergen, M. *et al.* The Core of Tau-Paired Helical Filaments Studied by Scanning Transmission Electron Microscopy and Limited Proteolysis. *Biochemistry* **45**, 6446–6457 (2006).
84. 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 Unmodi fi ed Recombinant Protein. (2013).
85. Woerman, A. L. *et al.* Tau prions from Alzheimer's disease and chronic traumatic encephalopathy patients propagate in cultured cells. *Proc. Natl. Acad. Sci.* **113**, E8187–E8196 (2016).
86. Mirra, S. S. *et al.* Tau Pathology in a Family with Dementia and a P301L Mutation in Tau. *J. Neuropathol. Exp. Neurol.* **58**, 335–345 (1999).
87. Mirbaha, H. *et al.* Inert and seed-competent tau monomers elucidate the structural origins of aggregation. **12**, 145 (2017).
88. Eisenberg, D. S. & Boyer, D. R. A structure-based model for the electrostatic interaction of the N-terminus ofprotein tau with the fibril core of Alzheimer's Disease filaments. **1**, 1–8 (2018).
89. Fichou, Y., Vigers, M., Goring, A. K., Eschmann, N. A. & Han, S. Heparin-induced tau filaments are structurally heterogeneous and differ from Alzheimer's disease filaments. *ChemComm* **54**, 8–10 (2018).
90. 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 Unmodi fi ed Recombinant Protein. (2013) doi:10.1021/bi400866w.
91. Fichou, Y., Vigers, M., Goring, A. K., Eschmann, N. A. & Han, S. Heparin-induced tau filaments are structurally heterogeneous and differ from Alzheimer's disease filaments. *Chem. Commun.* **54**, 4573–4576 (2018).
92. Guo, J. L. *et al.* Unique pathological tau conformers from Alzheimer's brains transmit tau pathology in nontransgenic mice. *J. Exp. Med.* **213**, 2635–2654 (2016).
93. Xia, D., Gutmann, J. M. & Götz, J. Mobility and subcellular localization of endogenous, gene-edited Tau differs from that of over-expressed human wild-type and P301L mutant Tau. *Sci. Rep.* **6**, 29074 (2016).
94. Sultan, A. *et al.* Nuclear Tau, a Key Player in Neuronal DNA Protection. *J. Biol. Chem.* **286**, 4566–4575 (2011).
95. Yamada, K. Extracellular Tau and Its Potential Role in the Propagation of Tau Pathology. *Front. Neurosci.* **11**, (2017).
96. Maïza, A. *et al.* The role of heparan sulfates in protein aggregation and their potential impact on neurodegeneration. *FEBS Lett.* **592**, 3806–3818 (2018).
97. Rauch, J. N. *et al.* Tau Internalization is Regulated by 6-O Sulfation on Heparan Sulfate Proteoglycans (HSPGs). *Sci. Rep.* **8**, 6382 (2018).
98. Zhang, X. *et al.* RNA stores tau reversibly in complex coacervates. *PLOS Biol.* **15**, e2002183 (2017).
99. Ginsberg, S. D. *et al.* RNA sequestration to pathological lesions of neurodegenerative diseases. *Acta Neuropathol.* **96**, 487–494 (1998).
100. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's

disease. *Nature* **547**, 185–190 (2017).
101. Falcon, B., Zhang, W., Murzin, A. G., Murshudov, G. & Holly, J. Structures of filaments from Pick ' s disease reveal a novel tau protein fold. (2018).
102. Mirbaha, H. *et al.* Inert and seed-competent tau monomers suggest structural origins of aggregation. *Elife* 1–29 (2018).
103. Gamblin, T. C. *et al.* Caspase cleavage of tau : Linking amyloid and neurofibrillary tangles in Alzheimer ' s disease. (2003).
104. Goedert, M. & Spillantini, M. G. REVIEWS A Century of Alzheimer ' s Disease. 777–782 (2006).
105. Goedert, M., Eisenberg, D. S. & Crowther, R. A. Propagation of Tau Aggregates and Neurodegeneration. *Annu. Rev. Neurosci.* **40**, 189–210 (2017).
106. Goedert, M. & Spillantini, M. G. Pathogenesis of the tauopathies. *J. Mol. Neurosci.* **45**, 425–431 (2011).
107. Dujardin, S. *et al.* Tau molecular diversity contributes to clinical heterogeneity in Alzheimer's disease. *Nat. Med.* (2020) doi:10.1038/s41591-020-0938-9.
108. Basler, M., Kirk, C. J. & Groettrup, M. The immunoproteasome in antigen processing and other immunological functions. *Curr. Opin. Immunol.* **25**, 74–80 (2013).
109. Murata, S., Takahama, Y., Kasahara, M. & Tanaka, K. The immunoproteasome and thymoproteasome: functions, evolution and human disease. *Nat. Immunol.* **19**, 923–931 (2018).
110. Ukmar-Godec, T. *et al.* Proteasomal degradation of the intrinsically disordered protein tau at single-residue resolution. *Sci. Adv.* **6**, (2020).
111. Chen, D. *et al.* Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nat. Commun.* **10**, (2019).
112. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
113. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
114. Livingston, G. *et al.* Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413–446 (2020).
115. Orre, M. *et al.* Reactive glia show increased immunoproteasome activity in Alzheimer's disease. *Brain* **136**, 1415–1431 (2013).
116. Busche, M. A. *et al.* Tau impairs neural circuits, dominating amyloid-β effects, in Alzheimer models in vivo. *Nat. Neurosci.* **22**, 57–64 (2019).
117. Yeo, I. J. *et al.* A dual inhibitor of the proteasome catalytic subunits LMP2 and Y attenuates disease progression in mouse models of Alzheimer's disease. *Sci. Rep.* **9**, 18393 (2019).
118. Bhaskar, K. *et al.* Regulation of Tau Pathology by the Microglial Fractalkine Receptor. *Neuron* **68**, 19–31 (2010).
119. Wischik, C. M. *et al.* Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease. *Proc. Natl. Acad. Sci.* **85**, 4506–4510 (1988).
120. Al-hilaly, Y. K. *et al.* Alzheimer ' s Disease-like Paired Helical Filament Assembly from Truncated Tau Protein Is Independent of Disulfide Crosslinking. *J. Mol. Biol.* **429**, 3650–3665 (2017).
121. Al-Hilaly, Y. K. *et al.* Tau (297-391) forms filaments that structurally mimic the core

of paired helical filaments in Alzheimer's disease brain. *FEBS Lett.* **594**, 944–950 (2020).
122. Lövestam, S. *et al.* Assembly of recombinant tau into filaments identical to those of Alzheimer's disease and chronic traumatic encephalopathy. *Elife* **11**, (2022).
123. Shi, Y. *et al.* Structure-based classification of tauopathies. *Nature* (2021) doi:10.1038/s41586-021-03911-7.
124. Xiang, X. *et al.* Role of molecular polymorphism in defining tau filament structures in neurodegenerative diseases. (2021).
125. Jensen, M. R. *et al.* Quantitative Determination of the Conformational Properties of Partially Folded and Intrinsically Disordered Proteins Using NMR Dipolar Couplings. *Structure* **17**, 1169–1185 (2009).
126. Guo, J. L. & Lee, V. M. Y. Seeding of normal tau by pathological tau conformers drives pathogenesis of Alzheimer-like tangles. *J. Biol. Chem.* **286**, 15317–15331 (2011).
127. Frost, B., Jacks, R. L. & Diamond, M. I. Propagation of Tau Misfolding from the Outside to the Inside of a Cell *. **284**, 12845–12852 (2009).
128. Falcon, B. *et al.* Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* (2018).
129. Woerman, A. L. *et al.* Propagation of prions causing synucleinopathies in cultured cells. *Proc. Natl. Acad. Sci.* **112**, E4949 LP-E4958 (2015).
130. Ingram, E. M. & Spillantini, M. G. Tau gene mutations: dissecting the pathogenesis of FTDP-17. *Trends Mol. Med.* **8**, 555–562 (2002).
131. Goedert, M. Tau gene mutations and their effects. *Mov. Disord.* **20**, (2005).
132. Oddo, S. *et al.* Triple-transgenic model of Alzheimer's Disease with plaques and tangles: Intracellular Aβ and synaptic dysfunction. *Neuron* **39**, 409–421 (2003).
133. Allen, B. *et al.* Abundant Tau Filaments and Nonapoptotic Neurodegeneration in Transgenic Mice Expressing Human P301S Tau Protein. *J. Neurosci.* **22**, 9340–9351 (2002).
134. Lewis, J. *et al.* Neurofibrillary tangles, amyotrophy and progressive motor disturbance in mice expressing mutant (P301L) tau protein. *Nat. Genet.* **25**, 402–405 (2000).
135. Gratuze, M. *et al.* Impact of TREM2R47H variant on tau pathology–induced gliosis and neurodegeneration. *J. Clin. Invest.* **130**, 4954–4968 (2020).
136. Platt, B. *et al.* Abnormal Cognition, Sleep, EEG and Brain Metabolism in a Novel Knock-In Alzheimer Mouse, PLB1. *PLoS One* **6**, e27068 (2011).
137. Terwel, D. *et al.* Changed Conformation of Mutant Tau-P301L Underlies the Moribund Tauopathy, Absent in Progressive, Nonlethal Axonopathy of Tau-4R/2N Transgenic Mice. *J. Biol. Chem.* **280**, 3963–3973 (2005).
138. Kfoury, N., Holmes, B. B., Jiang, H., Holtzman, D. M. & Diamond, M. I. Trans-cellular Propagation of Tau Aggregation by Fibrillar Species. *J. Biol. Chem.* **287**, 19440–19451 (2012).
139. Yao, Q. Q., Hong, L., Wu, S. & Perrett, S. Distinct microscopic mechanisms for the accelerated aggregation of pathogenic Tau mutants revealed by kinetic analysis. *Phys. Chem. Chem. Phys.* **22**, 7241–7249 (2020).
140. Seidler, P. M. *et al.* Structure-based inhibitors of tau aggregation. **10**, (2018).
141. Kawasaki, R. & Tate, S. I. Impact of the hereditary p301l mutation on the correlated conformational dynamics of human tau protein revealed by the paramagnetic

relaxation enhancement NMR experiments. *Int. J. Mol. Sci.* **21**, (2020).
142. Mylonas, E. *et al.* Domain Conformation of Tau Protein Studied by Solution Small-Angle X-ray. *Biochemistry* **47**, 10345–10353 (2008).
143. Stöhr, J. *et al.* A 31-residue peptide induces aggregation of tau's microtubule-binding region in cells. *Nat. Chem.* **9**, 874–881 (2017).
144. Laos, V. *et al.* Catalytic Cross Talk between Key Peptide Fragments That Couple Alzheimer's Disease with Amyotrophic Lateral Sclerosis. *J. Am. Chem. Soc.* **143**, 3494–3502 (2021).
145. Arya, S. *et al.* Terminal Capping of an Amyloidogenic Tau Fragment Modulates Its Fibrillation Propensity. *J. Phys. Chem. B* **124**, 8772–8783 (2020).
146. Ganguly, P. *et al.* Tau Assembly: The Dominant Role of PHF6 (VQIVYK) in Microtubule Binding Region Repeat R3. *J. Phys. Chem. B* **119**, 4582–4593 (2015).
147. Nguyen, P. H. *et al.* Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chem. Rev.* **121**, 2545–2647 (2021).
148. Jeschke, G., Koch, A., Jonas, U. & Godt, A. Direct conversion of EPR dipolar time evolution data to distance distributions. *Journal of Magnetic Resonance* vol. 155 72–82 (2002).
149. Sherck, N. *et al.* End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* **142**, 19631–19641 (2020).
150. Banham, J. E. *et al.* Distance measurements in the borderline region of applicability of CW EPR and DEER: A model study on a homologous series of spin-labelled peptides. *J. Magn. Reson.* **191**, 202–218 (2008).
151. Altenbach, C., Oh, K. J., Trabanino, R. J., Hideg, K. & Hubbell, W. L. Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: Experimental strategies and practical limitations. *Biochemistry* **40**, 15471–15482 (2001).
152. Cheng, C.-Y., Varkey, J., Ambroso, M. R., Langen, R. & Han, S. Hydration dynamics as an intrinsic ruler for refining protein structure at lipid membrane interfaces. *Proc. Natl. Acad. Sci.* **110**, 16838–16843 (2013).
153. Barnes, R. *et al.* Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* **139**, 17890–17901 (2017).
154. Franck, J. M., Pavlova, A., Scott, J. A. & Han, S. Quantitative cw Overhauser effect dynamic nuclear polarization for the analysis of local water dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **74**, 33–56 (2013).
155. Rego, N. B. & Patel, A. J. Understanding Hydrophobic Effects: Insights from Water Density Fluctuations. *Annu. Rev. Condens. Matter Phys.* **13**, 303–324 (2022).
156. Nozaki, Y. & Tanford, C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions. *J. Biol. Chem.* **246**, 2211–2217 (1971).
157. Patel, A. J., Varilly, P., Chandler, D. & Garde, S. Quantifying Density Fluctuations in Volumes of All Shapes and Sizes Using Indirect Umbrella Sampling. *J. Stat. Phys.* **145**, 265–275 (2011).
158. Jiang, Z., Remsing, R. C., Rego, N. B. & Patel, A. J. Characterizing Solvent Density Fluctuations in Dynamical Observation Volumes. *J. Phys. Chem. B* **123**, 1650–1661 (2019).

159. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
160. Tan, Z., Gallicchio, E., Lapelosa, M. & Levy, R. M. Theory of binless multi-state free energy estimation with applications to protein-ligand binding. *J. Chem. Phys.* **136**, 144102 (2012).
161. Hudait, A., Odendahl, N., Qiu, Y., Paesani, F. & Molinero, V. Ice-Nucleating and Antifreeze Proteins Recognize Ice through a Diversity of Anchored Clathrate and Ice-like Motifs. *J. Am. Chem. Soc.* **140**, 4905–4912 (2018).
162. Monroe, J. I. & Shell, M. S. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *J. Chem. Phys.* **151**, 094501 (2019).
163. Jiao, S. *et al.* Sequence Modulates Polypeptoid Hydration Water Structure and Dynamics. *Biomacromolecules* **23**, 1745–1756 (2022).
164. Nelson, R. *et al.* Structure of the cross-β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
165. Jana, B. *et al.* Entropy of water in the hydration layer of major and minor grooves of DNA. *J. Phys. Chem. B* **110**, 19611–19618 (2006).
166. Adam, G. & Gibbs, J. H. On the Temperature Dependence of Cooperative Relaxation Properties in Glass-Forming Liquids. *J. Chem. Phys.* **43**, 139–146 (1965).
167. Wesseling, H. *et al.* Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's Disease. *Cell* **183**, 1699-1713.e13 (2020).
168. Ghormley J. A., J. Am. Chem. Soc. 79, 1862 (1957).
169. Glaeser, R. M., Nogales, E. & Chiu, W. *Single-particle Cryo-EM of Biological Macromolecules*. (IOP publishing, 2021). doi:10.1088/978-0-7503-3039-8ch1.
170. Glaeser, R. M. & Hall, R. J. Reaching the information limit in cryo-EM of biological macromolecules: Experimental aspects. *Biophys. J.* **100**, 2331–2337 (2011).
171. Lövestam, S. & Scheres, S. H. W. High-throughput cryo-EM structure determination of amyloids. (2022).
172. Chang, C.-W., Shao, E. & Mucke, L. Tau: Enabler of diverse brain disorders and target of rapidly evolving therapeutic strategies. *Science (80-.).* **371**, (2021).
173. Lin, Y. *et al.* Narrow equilibrium window for complex coacervation of tau and RNA under cellular conditions. *Elife* **8**, (2019).
174. Veis, A. A review of the early development of the thermodynamics of the complex coacervation phase separation. *Adv. Colloid Interface Sci.* **167**, 2–11 (2011).
175. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **11**, 899–904 (2015).
176. Fichou, Y. *et al.* Cofactors are essential constituents of stable and seeding-active tau fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 13234–13239 (2018).
177. Wegmann, S. *et al.* Tau protein liquid–liquid phase separation can initiate tau aggregation. *EMBO J.* **37**, (2018).
178. Hernández-Vega, A. *et al.* Local Nucleation of Microtubule Bundles through Tubulin Concentration into a Condensed Tau Phase. *Cell Rep.* **20**, 2304–2312 (2017).
179. Ambadipudi, S., Biernat, J., Riedel, D., Mandelkow, E. & Zweckstetter, M. Liquid-liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat. Commun.* **8**, (2017).

180. Tan, R. *et al.* Microtubules gate tau condensation to spatially regulate microtubule functions. *Nat. Cell Biol.* **21**, 1078–1085 (2019).
181. Siahaan, V. *et al.* Kinetically distinct phases of tau on microtubules regulate kinesin motors and severing enzymes. doi:10.1038/s41556-019-0374-6.
182. Lin, Y., Fichou, Y., Zeng, Z., Hu, N. Y. & Han, S. Electrostatically Driven Complex Coacervation and Amyloid Aggregation of Tau Are Independent Processes with Overlapping Conditions. *ACS Chem. Neurosci.* **11**, 615–627 (2020).
183. Fichou, Y. *et al.* Cofactors are essential constituents of stable and seeding-active tau fibrils. *Proc. Natl. Acad. Sci.* **115**, 13234–13239 (2018).
184. Bugaj, L. J., Choksi, A. T., Mesuda, C. K., Kane, R. S. & Schaffer, D. V. Optogenetic protein clustering and signaling activation in mammalian cells. *Nat. Methods* **10**, 249–252 (2013).
185. Shin, Y. *et al.* Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171.e14 (2017).
186. Gong, C.-X., Liu, F., Grundke-Iqbal, I. & Iqbal, K. Post-translational modifications of tau protein in Alzheimer's disease. *J. Neural Transm.* **112**, 813–838 (2005).
187. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
188. Mckibben, K. & Rhoades, E. Regulation of tau ' s proline rich region by its N-terminal domain. (2019).
189. Tan, R. *et al.* Microtubules gate tau condensation to spatially regulate microtubule functions. *Nat. Cell Biol.* **21**, 1078–1085 (2019).
190. Akhmanova, A. & Steinmetz, M. O. Microtubule +TIPs at a glance. *J. Cell Sci.* **123**, 3415–3419 (2010).
191. Lasorsa, A. *et al.* Structural Basis of Tau Interaction With BIN1 and Regulation by Tau Phosphorylation. *Front. Mol. Neurosci.* **11**, (2018).
192. Seshadri, S. Genome-wide Analysis of Genetic Loci Associated With Alzheimer Disease. *JAMA* **303**, 1832 (2010).
193. Calafate, S., Flavin, W., Verstreken, P. & Moechars, D. Loss of Bin1 Promotes the Propagation of Tau Pathology. *Cell Rep.* **17**, 931–940 (2016).
194. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
195. Zhang, X. *et al.* RNA stores tau reversibly in complex coacervates. 1–28 (2017).
196. Cornblath, E. J. *et al.* Computational modeling of tau pathology spread reveals patterns of regional vulnerability and the impact of a genetic risk factor. 1–16 (2021).
197. van der Kant, R. *et al.* Cholesterol Metabolism Is a Druggable Axis that Independently Regulates Tau and Amyloid-β in iPSC-Derived Alzheimer's Disease Neurons. *Cell Stem Cell* vol. 24 363-375.e9 (2019).
198. Pavlova, A. *et al.* Protein structural and surface water rearrangement constitute major events in the earliest aggregation stages of tau. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E127–E136 (2016).