

## Article

# Hexameric Aggregation Nucleation Core Sequences and Diversity of Pathogenic Tau Strains

Ling Wu<sup>1,2</sup>, Sidharth S. Madhavan<sup>3,4</sup> , Christopher Tan<sup>3</sup> and Bin Xu<sup>1,2,3,4,\*</sup> 

<sup>1</sup> Biomanufacturing Research Institute and Technology Enterprise (BRITE), Department of Pharmaceutical Sciences, North Carolina Central University, Durham, NC 27707, USA

<sup>2</sup> Duke/UNC Alzheimer's Disease Research Center, Durham, NC 27710, USA

<sup>3</sup> Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

<sup>4</sup> School of Neuroscience, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

\* Correspondence: bxu@ncsu.edu

**Abstract:** Tau aggregation associates with multiple neurodegenerative diseases including Alzheimer's disease and rare tauopathies such as Pick's disease, progressive supranuclear palsy, and corticobasal degeneration. The molecular and structural basis of tau aggregation and related diverse misfolded tau strains are not fully understood. To further understand tau-protein aggregation mechanisms, we performed systematic truncation mutagenesis and mapped key segments of tau proteins that contribute to tau aggregation, where it was determined that microtubule binding domains R2 and R3 play critical roles. We validated that R2- or R3-related hexameric PHF6 and PHF6\* peptide sequences are necessary sequences that render tau amyloidogenicity. We also determined that the consensus VQI peptide sequence is not sufficient for amyloidogenicity. We further proposed single- and dual-nucleation core-based strain classifications based on recent cryo-EM structures. We analyzed the structural environment of the hexameric peptide sequences in diverse tau strains in tauopathies that, in part, explains why the VQI consensus core sequence is not sufficient to induce tau aggregation. Our experimental work and complementary structural analysis highlighted the indispensable roles of the hexameric core sequences, and shed light on how the interaction environment of these core sequences contributes to diverse pathogenic tau-strains formation in various tauopathy brains.

**Keywords:** tau aggregation; tauopathies; nucleation core sequences; amyloidogenicity; tau strains diversity



**Citation:** Wu, L.; Madhavan, S.S.; Tan, C.; Xu, B. Hexameric Aggregation Nucleation Core Sequences and Diversity of Pathogenic Tau Strains. *Pathogens* **2022**, *11*, 1559. <https://doi.org/10.3390/pathogens11121559>

Academic Editor: Lawrence S. Young

Received: 27 September 2022

Accepted: 14 December 2022

Published: 19 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

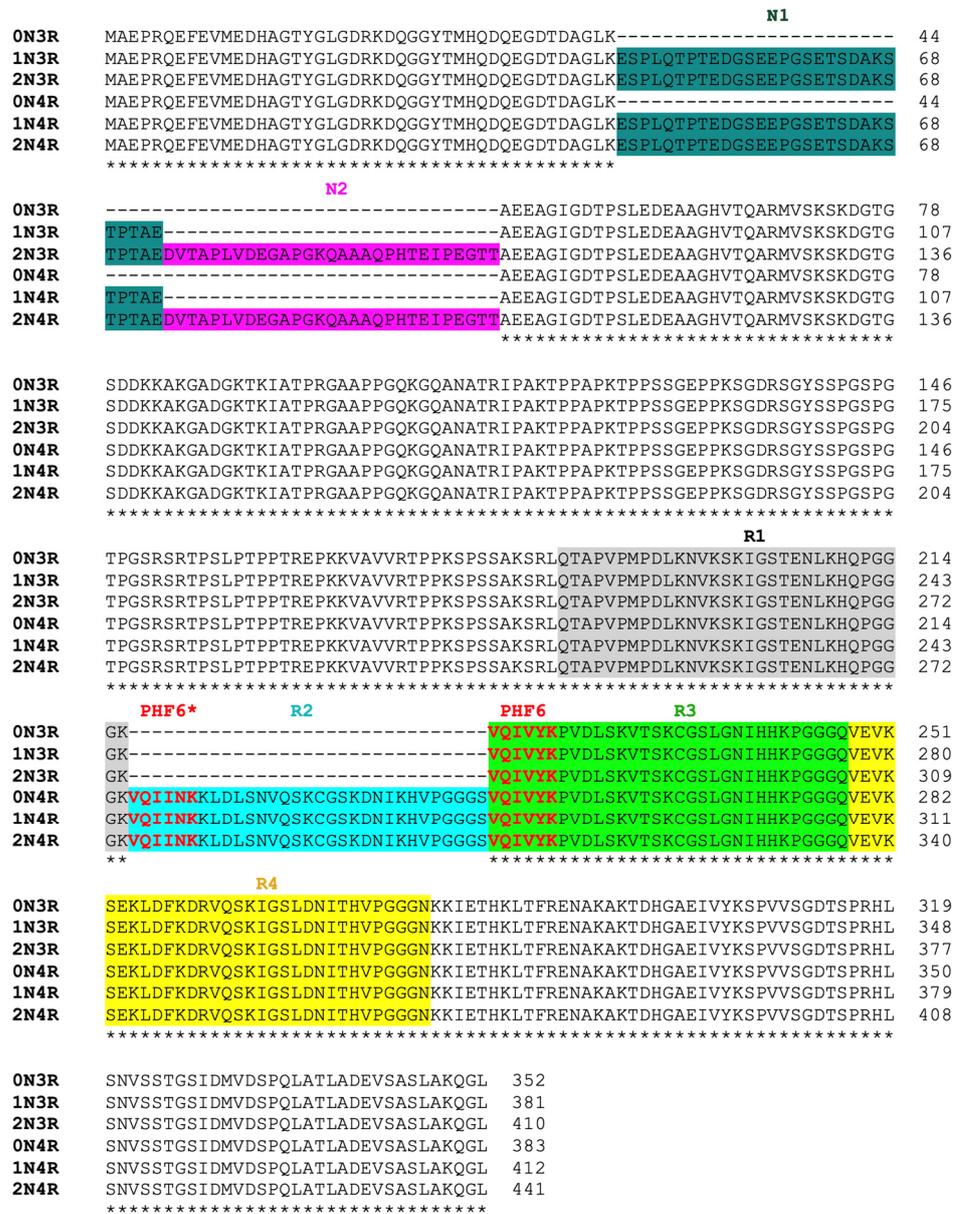


**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Extracellular A $\beta$  amyloid plaques and intraneuronal tau neurofibrillary tangles represent hallmark features of Alzheimer's disease (AD) [1,2]. These abnormal structures were generally believed to cause the synaptic dysfunction and neuronal death that lead to the memory and cognitive deficits characteristic of AD patients [3,4]. Multiple lines of evidence have suggested that pathological changes in tangles correlate better with neuronal dysfunction than A $\beta$  deposits [5,6]. Furthermore, a close relationship between tau aggregates and neuronal loss is well established in hippocampus and cerebral-cortex tissues [7]. Tau aggregates are present in multiple neurodegenerative diseases known as "tauopathies" [8,9], including AD, Pick's disease (PiD), progressive supranuclear palsy (PSP), and corticobasal degeneration (CBD). Six tau isoforms are expressed in human brains. These isoforms contain either three or four microtubule-binding repeats (3R or 4R tau) and 0–2 N-terminal inserts (0N, 1N, or 2N tau) (Figure 1) [10–12]. Based on the tau isoforms that constitute the abnormal filaments, tauopathies can be divided into three groups. In AD and chronic traumatic encephalopathy (CTE), both 3R and 4R isoforms make up the neuronal inclusions [13,14]; 3R tau is found in PiD, whereas 4R tau isoforms are present in the filaments of PSP, CBD, argyophilic grain disease (AGD), and globular glial tauopathy (GGT) [15–17]. Transmission of the AD pathology and related tauopathies is not fully

understood, but is thought to be through “prion-like seeding” mechanisms that ultimately yield intercellular spreading of toxic tau aggregates [18–25].



**Figure 1.** Human tau isoforms sequence alignment. Identical residues in the alignment are labelled with asterisks. Nucleation core sequences VQIVYK (PHF6) in all six isoforms and VQIINK (PHF6\*) in 4R-tau isoforms are highlighted in red. N-terminal insert domain N1 is shaded in dark green, and N2 in purple. Microtubule repeat segments R1–R4 are highlighted in grey, cyan, green, and yellow, respectively.

Significantly, isoform composition and the morphology of tau filaments can differ between tauopathies, suggesting the existence of distinct misfolded tau strains, molecular heterogeneity and complexity of these tauopathy diseases [26–29]. Recently, cryo-EM structures of tau filaments extracted from postmortem brains of different tauopathies patients have been solved. The diversity of the structural folds, which are in part reflected in the different morphology of tau filaments, allows us to gain insights into distinct misfolded tau strains. Furthermore, past experimental studies have recognized the important roles the microtubule-binding domain repeats (R1–R4) play in tau aggregation (Figure 1; [30]). Specifically, two nucleation core sequences in the microtubule binding region (MTBR),

<sup>275</sup>VQIINK<sup>280</sup> (PHF6\*) and <sup>306</sup>VQIVYK<sup>311</sup> (PHF6) were identified, each located at the beginning of R2 or R3 repeat region, respectively (Figure 1; [31]).

The objectives of this study are to understand sequence-specific human tau aggregation mechanisms and how key sequences contribute to the structural diversity of different tauopathies. Through systematic truncation mutagenesis, we not only mapped out R2 and R3 as key segments contributing to tau aggregation, but also validated that R2- and R3-related hexameric PHF6 and PHF6\* peptide sequences are sufficient sequences that render tau amyloidogenicity, yet the consensus VQI peptide sequence alone is not sufficient. Furthermore, by analysing a collection of tau filament structures from diverse pathogenic tau strains, we proposed single- and dual-nucleation core-based strain classifications with respect to the hexameric PHF6 and PHF6\* core sequences. Information on the structural environment of the hexameric peptide nucleation core sequences from a diverse set of pathogenic tau strains gives insights into why VQI consensus sequence alone is not sufficient for tau aggregation. Overall, our work not only highlighted the indispensable roles of hexameric core sequences, but also shed light on how their interaction environment contributes to diverse pathogenic tau strain formation, which is important for distinct tau strain propagation.

## 2. Materials and Methods

### 2.1. Plasmid Constructs

Expression vectors for his-tagged versions of all six wild-type human tau isoforms were kindly provided by the late Dr. Lester “Skip” Binder and Dr. Nicolas Kanaan of Michigan State University (via Prof. George Bloom of the University of Virginia). All the truncation constructs were first PCR-amplified of each specified segment of sequence from 2N4R tau plasmid (except for construct “C7”, where 2N3R tau plasmid was used as parent) and cloned into the same expression vector using Nde I and Xho I restriction sites. Oligo primers and related annealing temperatures for all the truncation mutagenesis are described in Supplemental Table S1. The PCR mixture preparation and thermocycling followed standard protocol. The PCR mixture consisted of 1 µL of 10X reaction buffer, 0.2 µL of 10 mM dNTP mix (Invitrogen, Waltham, MA, USA; cat. #18427013; 200 µM each final), 1 µL each of forward and reverse primers (200 nM each primer final), 0.2 µL of Pfu DNA polymerase (Agilent, Santa Clara, CA, USA; cat. #99903-844; 1 U/µL), and 10 ng of template in a total volume of 10 µL. PCR was performed with the following cycling profile: initial denaturation at 95 °C for 2 min, followed by 30 cycles of 30 sec denaturation at 95 °C, annealing temperature (different in primer pairs relevant to their T<sub>m</sub>) for 30 sec, and extension at 72 °C for 1 min. The time for the final extension step was increased to 10 min. The PCR product was cleaned up by Qiagen PCR purification kit (cat. #28104), and then digested with NdeI/XhoI (Thermo Fisher Scientific, Waltham, MA, USA; cat. #ER0585 and ER0695), the fragments were then purified using the QIA quick gel extraction kit (Qiagen, Hilden, Germany; cat. #20021). The enzyme-digested product was ligated into the vector derived from NdeI/XhoI digested tau 2N4R product. Plasmid ligation was achieved by T4 DNA ligase (New England BioLabs, Ipswich, MA, USA; cat. #M0202S) at 22 °C overnight. The recombinant plasmids were transformed into E. coli XL1Blue (Agilent; cat. #200249) for amplification. Plasmid DNA from the resulting clones was purified using the QIAprep spin miniprep kit (Qiagen; cat. #27104). All constructs were designed with a his<sub>6</sub>-tag at their carboxy-termini to facilitate protein purification and were verified by DNA sequencing.

### 2.2. Recombinant Truncated Tau Fragments Expression and Purification

Recombinant tau protein expression and purification, in general, follows protocol as described [32,33]. Plasmids encoding human tau isoforms or truncated tau fragments were transformed into BL21-DE3 E. coli cells (Thermo Fisher Scientific; cat. #C600003). Overnight starter cultures of BL21-DE3 E. coli cells transformed with recombinant tau plasmids were inoculated into multi-liter LB broth at 1:50 dilution and 100 µg/mL ampicillin. Cultures were incubated at 37 °C, and shaken until OD<sub>600</sub> reached between 0.5–0.6. Tau expression

was induced using 1 mM IPTG and the cultures continued for an additional 4 h. BL21-DE3 cells containing expressed tau were pelleted and resuspended in 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0 and 300 mM NaCl (sonication lysis buffer) at a concentration of 20 mL/L of culture preparation and sonicated at 60% power in ten 30-s intervals over 10 min. Cell lysates were centrifuged and supernatant containing the protein was applied to Ni-NTA (Qiagen; cat. #30250) column equilibrated with sonication lysis buffer. The columns were washed with 40–60 times of bed volumes of column buffer followed by washing buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8, 300 mM NaCl, and 20 mM imidazole). Recombinant tau proteins were then eluted using elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8, 300 mM NaCl, and 200 mM imidazole). Fractions were tested for protein concentration using 5 µL of protein sample mixed with 10 µL Coomassie Protein Assay reagent (Thermo Fisher Scientific; cat. #23200). Pooled fractions were concentrated to 4 mL using 10 kD or 3 kD molecular weight cut-off spin columns (MilliporeSigma, Burlington, MA, USA; cat. #UFC901096 and 900324) and filtered using 0.22 µm low-binding Durapore PVDF membrane filters (MilliporeSigma; cat. #UFC40GV00 and UFC30GV00). Truncated tau proteins were further purified by FPLC using size-exclusion Superdex75 10/300 GL column (GE Healthcare, Chicago, IL, USA; cat. #17517401) in 1X PNE buffer (25 mM PIPES, 150 mM NaCl and 1 mM EDTA at pH 7.0). Purified tau fragments were evaluated by SDS-PAGE for purity and quantified by Pierce BCA protein assays (Thermo Fisher Scientific; cat. #23225).

### 2.3. Thioflavin-T Fluorescence Aggregation Kinetic Analysis

Fluorescence-based experiments were performed using a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA, USA). All kinetic reads were taken at 37 °C in non-binding all-black clear-bottom Greiner 96-well plates covered with optically clear films and stirred for 10 s prior to each reading. ThT fluorescence was measured at 444 nm and 491 nm as excitation and emission wavelengths. Each kinetic assay consisted of final concentrations of 30 µM tau protein, 60 µg/mL heparin, and 10 µM ThT [34]. Kinetic signals were collected until fluorescence signals plateaued, typically in 40–80 h.

### 2.4. Protein Sequence Alignment

Tau protein isoform alignment was performed by MAFFT version 7, a similarity-based multiple sequence alignment (MSA) program [35]. MSA program assumes all the input sequences are homologous and descended from a common ancestor. Fully conserved regions were labelled as asterisks.

### 2.5. Structural Analysis of Tau Strains from Different Tauopathies

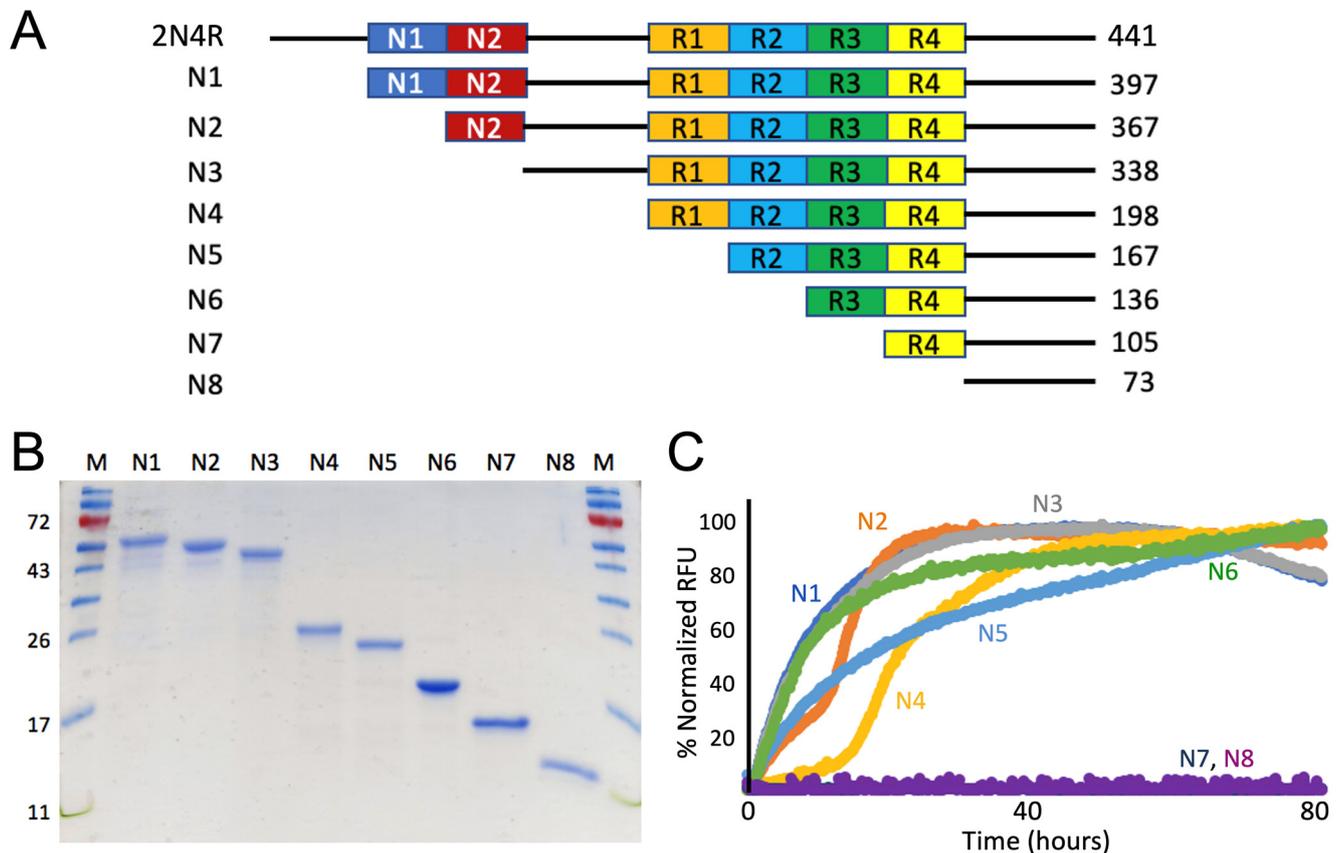
Protein structural coordinates were retrieved from PDB Protein Data Bank ([www.rcsb.org](http://www.rcsb.org), accessed on 24 September 2022). For all tau cryo-EM structures, PDB access codes are 5O3L for the Alzheimer's fold [36], 6NWP for the CTE fold [37], 6GX5 for the Pick's fold [38], 6TJO for the CBD fold [39], 7P6D for the AGD fold [17], 7P65 for the PSP fold [17], 7P66 for the GGT fold [17], and 7P6A for the GPT fold [17]. PyMol molecular visualization system (Schrodinger, Inc., New York, NY, USA) was used for all protein structure display and atomic structural analysis applications. Hydrogen bonding length was limited to 3.3 angstrom in maximum and hydrophobic interaction distance cutoff was set for 5.0 angstrom in maximum.

## 3. Results

### 3.1. Truncation Mutagenesis of N-terminal & C-terminal Tau Sequences

Using 2N4R tau full-length expression plasmid as a parent template, we performed PCR amplification reactions with appropriate oligo primer sequences (Table S1) and generated eight tau protein constructs truncated from the full-length sequence systematically deleted from the N-terminus (named as N1–N8; Figure 2A). Domain boundaries were set by distinct domains, such as N1, N2, proline-rich domain, R1, R2, R3, and R4. Similarly, we also generated five tau-protein sequences truncated from the carboxy-terminus (named as

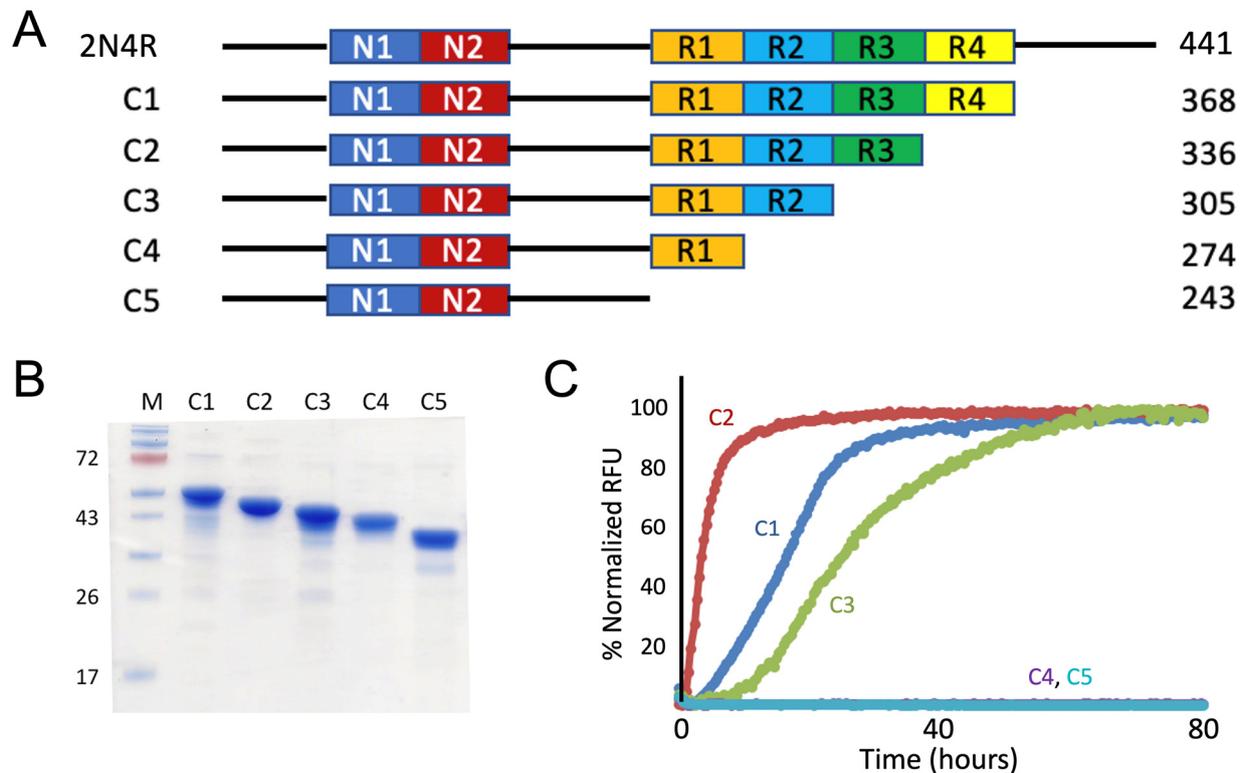
C1–C5; Figure 3A). A his<sub>6</sub>-tag was appended to either the N-terminus or C-terminus to facilitate protein purification. All the plasmids were sequence-verified to confirm there were no PCR-induced errors. Truncation mutants were transformed into expression host BL21DE3 bacteria. Ni-NTA affinity chromatography and size-exclusion chromatography were used to purify all the truncation mutants. Individual mutants were purified to homogeneity (>90% pure; Figures 2B and 3B).



**Figure 2.** Mapping the key segments in 2N4R tau for protein aggregation using N-terminal truncation mutagenesis: critical role of the R3 repeat for tau aggregation. (A) N-terminal truncation constructs are schematically illustrated. All truncation constructs were cloned using 2N4R tau (total 441 residues) as the template. A his<sub>6</sub>-tag was added to each construct in the carboxy-terminal (not shown) to facilitate protein purification. Sizes for each construct in amino-acid residues are specified. (B) SDS-PAGE of purified recombinant N-terminal truncation tau proteins corresponding constructs schematically drawn in the panel A. Molecular weight markers are shown on either side and corresponding molecular weights (in kD) are labelled to the left of the gel. (C) ThT fluorescence assays for each N-terminal truncation mutant. A total of 15–30  $\mu$ M of tau mutant proteins were incubated in 20 mM Tris pH 7.4 with 0.06 mg/mL of heparin at 37 °C for specified time periods. The results are representative of at least triplicated experiments.

Protein-aggregation properties were characterized by standard ThT fluorescence-based assays for detection and monitoring the aggregation of the resulting truncated tau proteins for an extended period of time (40–80 h) until the fluorescence intensities were well in the plateau phases (Figures 2C and 3C; [33,34]). For the N-terminal truncation mutants, it is evident that mutants N7 and N8 lost amyloidogenic properties while N1–N6 were competent in forming aggregation, as shown in Figure 2C. Based on the sequence difference boundary between N6 and N7, our ThT fluorescence results indicated the critical role of R3 repeat for aggregation. Similarly, for the C-terminal truncation mutants, mutants C4 and C5 lost amyloidogenicity (Figure 3C). Based on the sequence difference boundary between C3 and C4, ThT fluorescence data from the C-terminal truncation mutant set indicated the

critical role of R2 repeat for aggregation. Overall, our systematic truncation mutagenesis work pointed to the critical roles R2 and R3 play in tau aggregation.



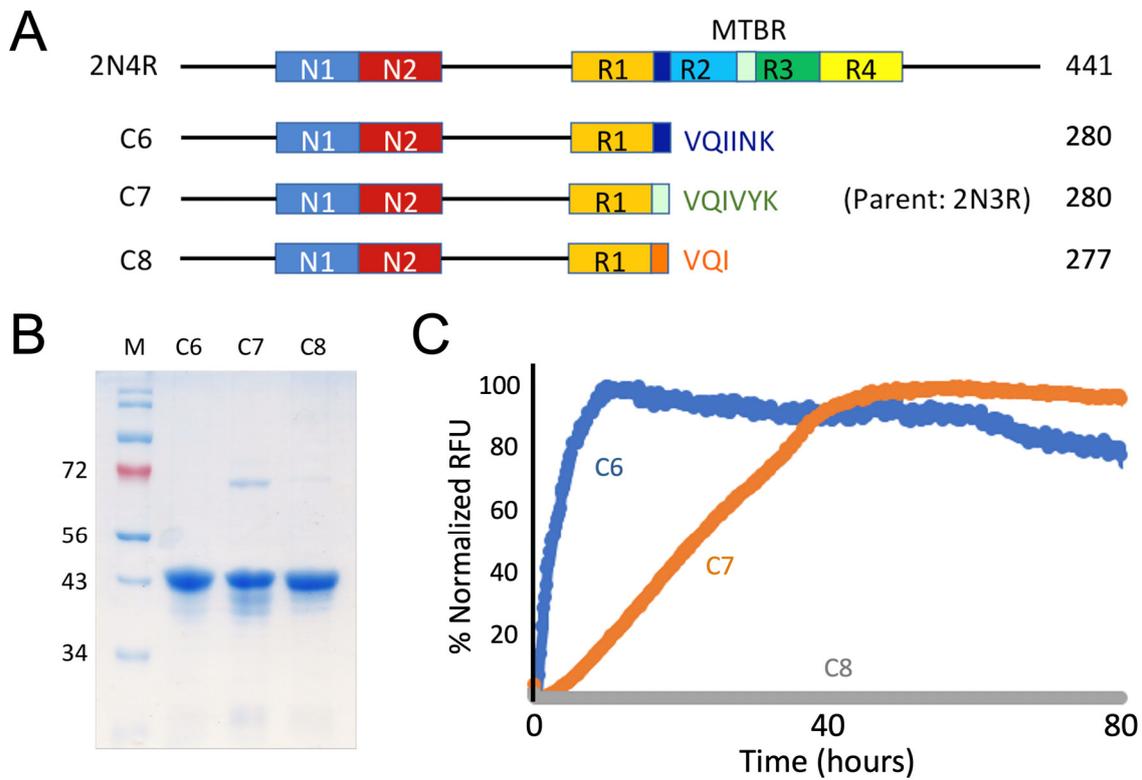
**Figure 3.** Mapping the key segments in 2N4R tau for protein aggregation using C-terminal truncation mutagenesis: critical role of the R2 repeat in tau aggregation. (A) Each C-terminal construct is schematically illustrated. All constructs were cloned using 2N4R tau (total 441 residues) as the template. A his<sub>6</sub>-tag was added to each construct in the carboxy terminal (not shown) to facilitate protein purification. Sizes for each construct in amino-acid residues are specified. (B) SDS-PAGE of purified recombinant C-terminal truncation tau proteins corresponding constructs schematically drawn in the panel A. Molecular weight markers (in kD) are labelled to the left of the gel. (C) ThT fluorescence assays for each truncation mutant. A total of 15–30  $\mu$ M of tau mutant proteins were incubated in 20 mM Tris pH 7.4 with 0.06 mg/mL of heparin at 37 °C for specified time periods. The results are representative of at least triplicated experiments.

### 3.2. Recombinant Constructs with PHF6\*, PHF6, and VQI Only Sequences

Based on our systematic truncation mutagenesis results and identification of R2 and R3 segments as two key segments in tau aggregation, and recognizing hexameric VQIINK (PHF6\*) and VQIVYK (PHF6) as nucleation core sequences from past studies [31], we engineered three additional tau mutants C6–C8. All three constructs used C4, a non-amyloidogenic construct, as a carrier (Figure 3A): each was extended by VQIINK (C6), VQIVYK (C7), or the consensus sequence of PHF6 and PHF6\*, VQI (C8). These constructs allowed us to test the hypothesis if the entire length of the hexameric sequences for both nucleation cores are necessary or they can be further shortened to consensus sequence VQI.

Similarly to the truncation mutants described in 3.1., recombinant constructs C6–C8 were cloned using PCR amplification steps and appropriate oligos. Both C6 and C8 used 2N4R as parent template since the VQIINK sequence, in the case of C6 (or VQI sequence in the case of C8), from the R2 segment directly joins R1. As, in the 2N3R tau parent construct, VQIVYK sequence from R3 segment directly joins R1, the C7 construct was engineered using 2N3R as the parent template (Figure 4A). Recombinant C6, C7, and C8 were purified

to homogeneity (>90% purity) after Ni-NTA affinity chromatography and size-exclusion chromatography (Figure 4B).



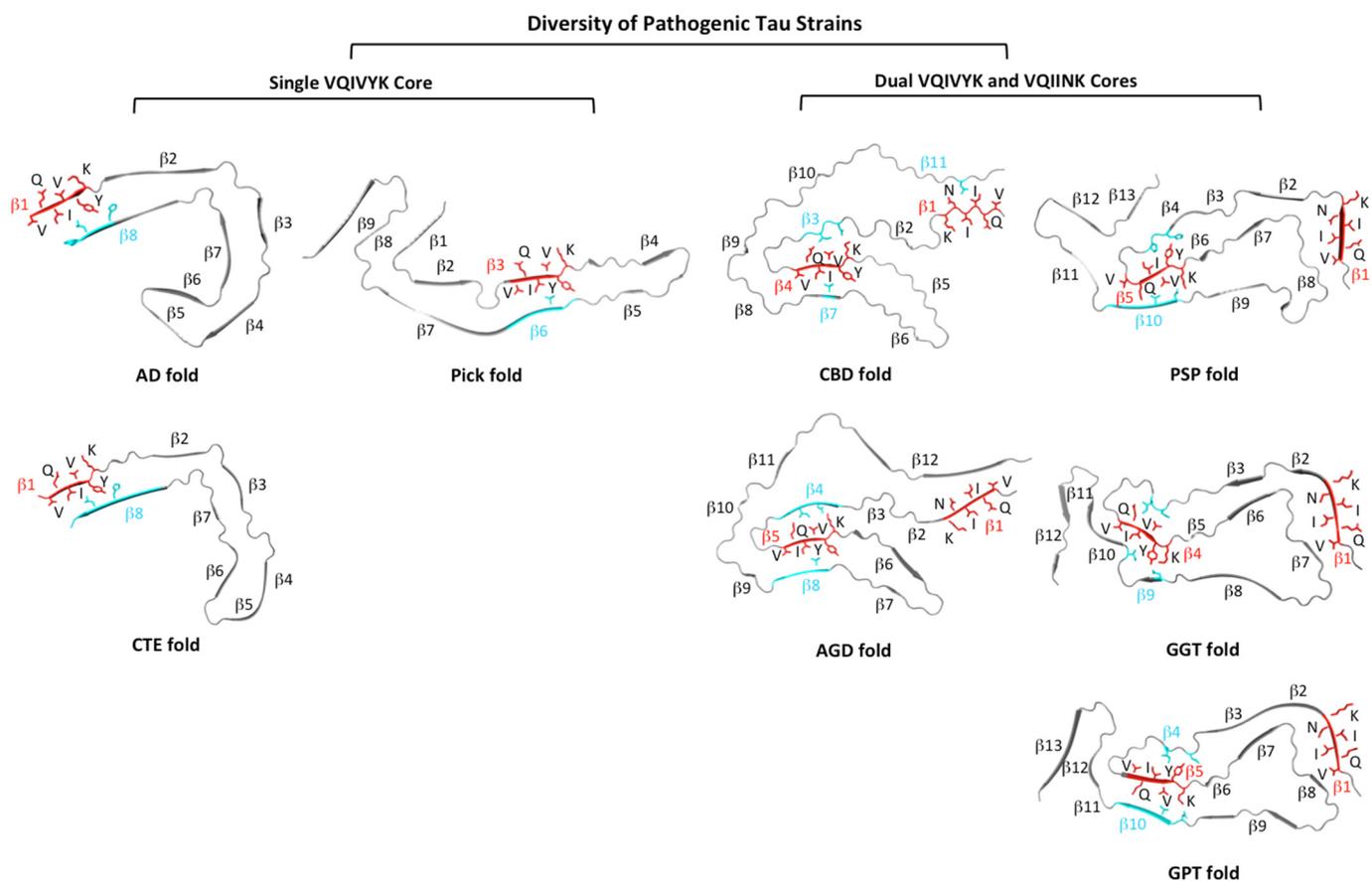
**Figure 4.** Mapping the key segments in 2N4R tau for aggregation using truncation mutagenesis: hexameric sequences PHF6 or PHF6\* each is sufficient to be amyloidogenic but consensus VQI sequence is not. (A) Each construct is schematically illustrated. MTBR stands for microtubule binding region. Except for C7 construct, which was cloned using 2N3R tau as the template, all other constructs were cloned using 2N4R tau as the template. Hexameric core sequences VQIINK, VQIVYK, and their consensus sequence VQI are shown color-coded in blue, light green and orange, respectively. A his<sub>6</sub>-tag was added to each construct in the carboxy terminal (not shown) to facilitate protein purification. Amino-acid residue numbers for each construct are specified. (B) SDS-PAGE of purified recombinant C-terminal truncation mutants corresponding to constructs schematically drawn in the panel A. Molecular weight markers (in kD) are labelled to the left of the gel. (C) ThT fluorescence assays for each truncation mutant. A total of 15–30 μM of tau mutant proteins were incubated in 20 mM Tris pH 7.4 with 0.06 mg/mL of heparin at 37 °C for specified time periods. The results are representative of at least triplicated experiments.

Protein aggregation was characterized by ThT fluorescence assays. As expected, C6 and C7 constructs demonstrated competency in forming tau amyloid (Figure 4C). However, the C8 construct (with VQI consensus sequence only) was not capable of forming tau amyloid. Comparison data of C6–C8 constructs showed that VQIINK or VQIVYK are alone sufficient to be amyloidogenic but shortened trimeric VQI sequence is not.

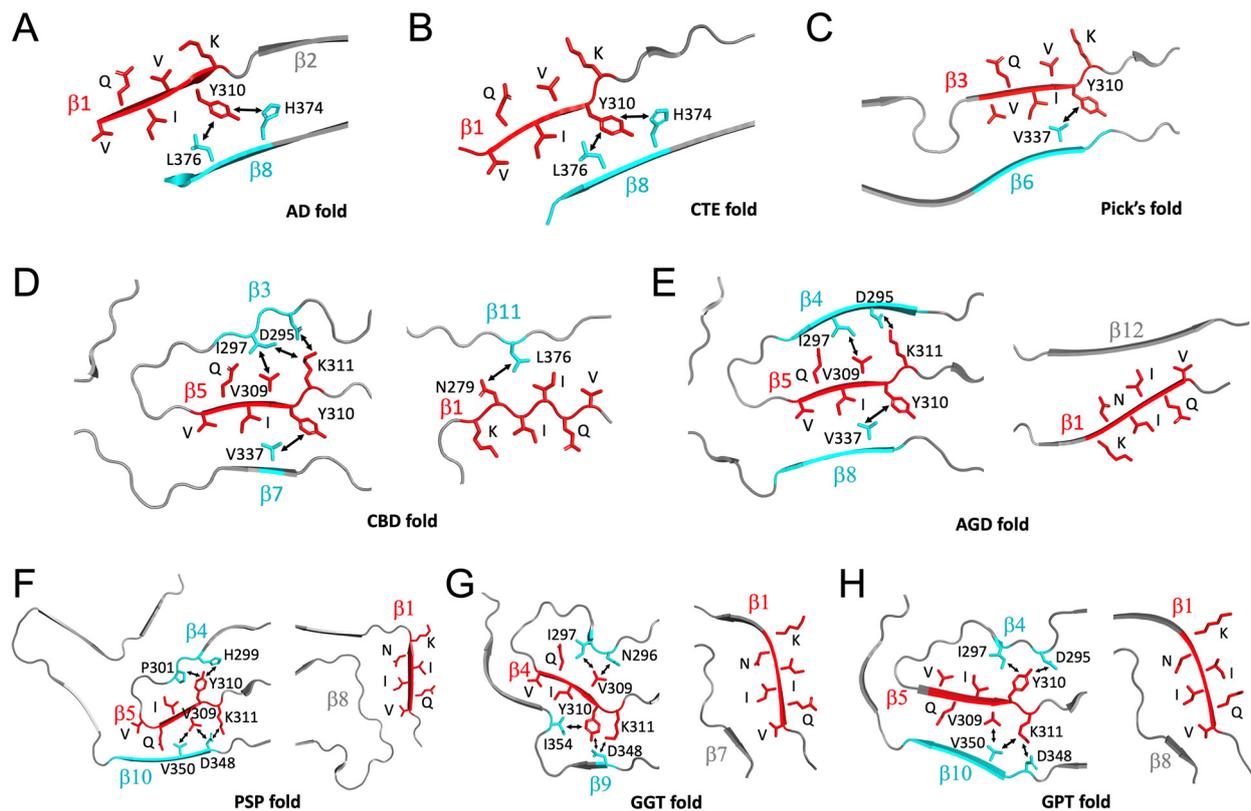
### 3.3. Hexameric Core Sequences in Diverse Pathogenic Tau Strains

Cryo-EM structures of tau filaments extracted from diverse tauopathy brains with definitive neuropathological diagnosis have been solved recently [17,36–39]. The diversity of tau-protein structural folds from different tauopathies in part reflects the diversity of pathogenic tau strains. Tauopathies are often classified based on tau-isoform composition in the respective filaments. Based on the presence of the PHF6 and PHF6\* nucleation core sequences, we propose tau structural folds in two classes (Figure 5): (1) Single VQIVYK Core. Three tauopathies, AD, CTE, and PiD each have a single VQIVYK core. (2) Dual VQIVYK

and VQIINK Cores. CBD AGD, PSP, GGT, and GPT have both VQIVYK and VQIINK cores. In each category, they may be further sub-classified based on structural folds. The CTE structural fold is quite similar to the “C”-shaped AD fold, but is significantly different from the elongated/extended Pick’s fold. In the dual nucleation core, CBD and AGD share similar structural folds; PSP fold, GGT fold, and GPT fold, however, demonstrated a different structural fold. Regarding two-nucleation-core hexameric sequences, VQIVYK is present in all tauopathy structures so far determined; however, VQIINK is only presented in tau filaments structures with the presence of dual-hexamer core sequences (CBD, AGD, PSP, GGT, and GPT). Comparing nucleation sequences in various tau structural folds and each core sequence’s interacting environment, we speculated that VQIVYK likely serves as the primary nucleation core because of its tight packing with other interacting elements. The VQIINK sequence likely serves as a secondary nucleation core since it has much fewer interactions with neighbouring strands or residues. The N-terminals of CBD, AGD, PSP, GGT, and GPT folds, where the VQIINK sequence is located, only loosely pack against the structural fold with few interactions with their neighbouring residues: very few intramolecular bonds (< 5.0 angstrom in distance) were found involving INK residues in the VQIINK core (Figures 5 and 6 and Table S2).



**Figure 5.** Molecular structure illustration of the diversity of pathogenic tau strains. Cryo-EM structures of tau protein folds in diverse tauopathies classified into single VQIVYK core strains (AD, CTE, and Pick folds) and dual VQIVYK and VQIINK core strains (CBD, AGD, PSP, GGT, and GPT folds). PDB access codes are 5O3L for AD fold, 6NWP for CTE fold, 6GX5 for Pick fold, 6TJO for CBD fold, 7P6D for AGD fold, 7P65 for PSP fold, 7P66 for GGT fold, and 7P6A for GPT fold. Hexameric core sequences are shown in red and their interaction amino acids/strands are shown in cyan. All  $\beta$  strands are labelled.



**Figure 6.** Atomic interactions involving residues **VYK** of **VQIVYK** (PHF6) or **INK** of **VQIINK** (PHF6\*) hexameric core sequences in diverse tau strains. (A–C) Atomic interactions with **VYK** of PHF6 core sequence. (D–H) Atomic interactions with **VYK** of PHF6 sequence (left panels) or with **INK** of PHF6\* sequence (right panels). Residue interactions are highlighted by arrows. Detailed inter-atomic interaction and distances are listed in Table S2. Hexameric core sequences are shown in red and their interacting amino acids/strands are shown in cyan. Relevant amino acids and  $\beta$  strands are labelled.

### 3.4. Why Consensus VQI Sequence Is not Sufficient for Tau Aggregation?

We further investigated the atomic basis of why the consensus sequence **VQI** is not sufficient to induce tau aggregation while either **VQIVYK** or **VQIINK** sequence is sufficient. We analysed the atomic interactions changes (loss of interactions) due to the absence of “**VYK**” in **VQIVYK** core sequence or “**INK**” in **VQIINK** core sequence in all available structural folds of tau filaments (Figure 6 and Table S2). In the single-nucleation-core **AD** fold, **CTE** fold, and **Pick’s** fold (Figure 6A–C), intra-molecular interactions between Tyr310 of **VQIVYK** with spatial neighbours Leu376 and His374 (in the **AD** fold and **CTE** fold), or between Tyr310 with spatial neighbour Val337 (in the **Pick’s** fold) will be lost in the absence of “**VYK**” residues. In the structural folds with dual nucleation cores, intra-molecular bonds between Val309 and Ile297, Tyr310 and Val337, Lys311 and Ile297, Lys311 and Asp295 related to core **VQIVYK**, and Asn279 and Leu376 related to core **VQIINK**, will be lost in the absence of “**VYK**” or “**INK**” in **CBD** fold (Figure 6D); similarly, bonds between Val309 and Ile297, Tyr310 and Val337, and Lys311 and Asp295 relevant to core **VQIVYK** will be lost in **AGD** fold (Figure 6E). In **PSP**, **GGT**, and **GPT** structural folds with dual cores, intra-molecular bonds between Val309 and Val350, Val309 and Asp348, Tyr310 and Pro301, Tyr310 and His299, Lys311 and Asp348 in **PSP** fold (Figure 6F), bonds between Val309 and Asn296, Val309 and Ile297, Tyr310 and Asp348, Tyr310 and Ile354, Lys311 and Asp348 in **GGT** fold (Figure 6G), Val309 and Val350, Tyr310 and Ile297, Tyr310 and Asp295, Lys311 and Val350, and Lys311 and Asp348 in **GPT** fold (Figure 6H), will be lost. Such loss of multiple intra-molecular bonds may significantly destabilize each structural fold to a

degree that the remaining VQI-induced intra-molecular bonds are not strong enough to form a nucleation core for amyloid growth.

The nature of interactions between “VYK” from VQIVYK and its neighbouring residues is primarily hydrophobic interactions supplemented with a few hydrogen bonds (Table S2). There are very few interactions between “INK” from VQIINK and its neighbours in the dual hexameric core tau structures (Figure 6). It is interesting to observe that VQIVYK nucleation core has more intra-molecular bonds with its neighbouring residues than those for VQIINK in the dual-core structural folds (Figure 6D–H), which is further supported by the evidence that VQIVYK is the only core present in the tau structural folds with a single nucleation core (Figure 6A–C). From the structural point of view, we propose VQIVYK may be the primary drive force for tau aggregation and VQIINK may play secondary roles in tau aggregation.

#### 4. Discussion

We determined the critical roles of R2 and R3 in tau aggregation and further validated that R2- or R3-related hexameric PHF6 and PHF6\* peptide sequences are necessary sequences that render tau amyloidogenicity. These results may have significant values for future translational applications. For example, we can use R2, R3, R2–R3, or PHF6, PHF6\* peptides as surrogates for tau-aggregation mechanistic studies or inhibitor screening [40]. Significantly, R3 or R2–R3 each represents a key segment in determining 3R-tau or 4R-tau isoform aggregation; therefore, these two peptides may be used for 3R- or 4R-tau isoform-specific aggregation mechanistic studies or inhibitor screening.

While significant past work had been performed to elucidate major roles played by R1–R4 repeating segments in tau aggregation, the functions of various domains and segments of tau proteins are not fully understood. For examples, past studies identified the critical roles of VQIVYK and VQIINK hexameric nucleation core sequences [31]. In this study, we provided experimental data to demonstrate that entire hexameric sequences are necessary and VQI alone is not sufficient for tau amyloidogenicity. We further complemented experimental data with structure-analysis-based explanations using multiple cryo-EM atomic structures of tau strains extracted from different postmortem tauopathy brains. Significant questions remain regarding the roles of other segments of the tau protein, for example, what are the roles of N-terminal N1 and N2 insert domains in tau aggregation? Our recent study showed significant kinetic differences in tau aggregation from pair-wise comparison of tau isoforms, 0N3R/0N4R, 1N3R/1N4R, and 2N3R/2N4R, where the differences in the  $t_{1/2}$  fold of changes may be assigned to the presence or absence of the N1 and N2 segments of the isoforms evaluated [33,41]. Novel roles of N1 and N2 remain to be elucidated and this is an interesting future direction.

It appears that the VQIVYK (PHF6) core sequence may be the primary nucleation core for tau aggregation and VQIINK plays a secondary role. This is supported by the evidence that there are more intra-molecular interactions between PHF6 and its environment versus those for PHF6\* (Figure 6D–H). However, inhibitors based on the structure of the VQIVYK segment only partially inhibit full-length tau aggregation and are ineffective at inhibiting seeding by full-length fibrils [42]. Designed VQIINK inhibitors, however, can more potently block seeding by full-length tau while VQIVYK-based inhibitors cannot. Therefore, it is still an open question as to which core sequence plays the dominant roles in driving tau aggregation and seeding. In physiological and pathological states, other factors may also play important roles in driving tau aggregation, as discussed below.

Multiple cell-based and in-vivo mice-based tau-seeding experiments provided strong evidence that tau protein has essential characteristics of a prion [19,20,27]. Prion-like propagation of tau aggregates may, therefore, underlie the disease pathogenesis and progression of neurodegenerative tauopathies. Recent groundbreaking cryo-EM structures of tau filaments from patients with AD, PiD, CBD, and other primary tauopathies further provided atomic evidence of different molecular conformers for a list of distinct neurodegenerative tauopathies [17,36–39]. Utilizing concepts and techniques such as RT-QuIC originally used

in prion disease research, we and others demonstrated “prion-like” tau-isoform seeding activities in diseased AD and related tauopathy brains [33,43–45]. Tau proteins from *in vivo* sources, including tauopathy brains, are heavily “decorated” by post-translational modifications (PTMs). Such modifications not only include well-recognized hyperphosphorylation, but also acetylation, ubiquitination, and methylation [46–48]. Therefore, it will be interesting and important to understand the roles of site-specific PTMs in tau aggregation, as these modifications exist *in vivo* and may play significant roles in tauopathy pathogenesis [47,49]. What are the roles of the six human tau isoforms in tau aggregation: do they play any role in disease progression in the context of rapid and slow progressive AD patients given that biochemical data suggests that they have different aggregation kinetics? How can we utilize the diversity of pathogenic tau strains and their atomic structures to design novel diagnostic tests for translation? New technological advances such as advanced mass spectrometry and ultrasensitive misfolded protein detection will certainly facilitate addressing many such important questions.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/pathogens11121559/s1>, Table S1: Oligo Primer Sequence Information for Truncation Mutagenesis. PCR annealing temperatures are calculated by IDT OligoAnalyzer Tool based on primer pair sequences and experimental reaction conditions; Table S2: Atomic interactions involving residues VYK of VQIVYK (PHF6) or INK of VQIINK (PHF6\*) hexameric core sequences in diverse tau strains. Corresponding intra-molecular interactions are highlighted by arrows in Figure 6. Inter-atomic distances are measured using Pymol Molecular Display Software (Version 2.5.4—Updated 17 August 2022).

**Author Contributions:** L.W. performed experiments, acquired and analysed the data, and prepared the manuscript. S.S.M. and C.T. performed experiments and acquired and analysed the data. B.X. conceived of and designed the experiments, analysed the data, and prepared the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported in part by grants from the Alzheimer’s Association, Biomarkers Across Neurodegenerative Diseases (BAND) Program (BAND-19-614848) (BX), Awards No. 16-1 and No. 18-4 from the Commonwealth of Virginia’s Alzheimer’s and Related Diseases Research Award Fund (BX and LW), National Institute of Health grants R03 AG061531 (BX) and R01 AG067607 (BX).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets used and/or analysed during the current study are available from the corresponding authors upon reasonable request.

**Acknowledgments:** We thank Virginia Tech Center for Drug Discovery and BRITE Research Institute of North Carolina Central University for instrument access.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Abbreviations

AD: Alzheimer’s disease; AGD: Argyophilic grain disease; CBD: Corticobasal degeneration; CNS: central nervous system; Cryo-EM: cryo-electron microscopy; CTE: chronic traumatic encephalopathy; GGT: globular glial tauopathy; GPT: GGT/PSP hybrid tauopathy; MSA: multiple sequence alignment; MT: microtubule; NFT: neurofibrillary tangle; PBS: phosphate-buffered saline; PHF: paired helical filament; PiD: Pick’s disease; PSP: progressive supranuclear palsy; PTM: post-translational modification; RFU: relative fluorescence unit; ThT: thioflavin T.

## References

1. Braak, H.; Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **1991**, *82*, 239–259. [[CrossRef](#)] [[PubMed](#)]
2. Selkoe, D.J. Alzheimer's disease: Genes, proteins, and therapy. *Physiol. Rev.* **2001**, *81*, 741–766. [[CrossRef](#)]
3. Bloom, G.S. Amyloid- $\beta$  and tau: The trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol.* **2014**, *71*, 505–508. [[CrossRef](#)] [[PubMed](#)]
4. Spire-Jones, T.L.; Hyman, B.T. The intersection of amyloid beta and tau at synapses in Alzheimer's disease. *Neuron* **2014**, *82*, 756–771. [[CrossRef](#)] [[PubMed](#)]
5. Wilcock, G.K.; Esiri, M.M. Plaques, tangles and dementia. A quantitative study. *J. Neurol. Sci.* **1982**, *56*, 343–356. [[CrossRef](#)]
6. Nelson, P.T.; Alafuzoff, I.; Bigio, E.H.; Bouras, C.; Braak, H.; Cairns, N.J.; Castellani, R.J.; Crain, B.J.; Davies, P.; Tredici, K.D.; et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: A review of the literature. *J. Neuropathol. Exp. Neurol.* **2012**, *71*, 362–381. [[CrossRef](#)]
7. Goedert, M.; Spillantini, M.G. Propagation of Tau aggregates. *Mol. Brain* **2017**, *10*, 18. [[CrossRef](#)]
8. Arai, T.; Ikeda, K.; Akiyama, H.; Shikamoto, Y.; Tsuchiya, K.; Yagishita, S.; Beach, T.; Rogers, J.; Schwab, C.; McGeer, P.L. Distinct isoforms of tau aggregated in neurons and glial cells in brains of patients with Pick's disease, corticobasal degeneration and progressive supranuclear palsy. *Acta Neuropathol.* **2001**, *101*, 167–173. [[CrossRef](#)]
9. Lee, V.M.; Goedert, M.; Trojanowski, J.Q. Neurodegenerative tauopathies. *Annu. Rev. Neurosci.* **2001**, *24*, 1121–1159. [[CrossRef](#)]
10. Goedert, M.; Spillantini, M.G.; Jakes, R.; Rutherford, D.; Crowther, R.A. Multiple isoforms of human microtubule-associated protein tau: Sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron* **1989**, *3*, 519–526. [[CrossRef](#)]
11. Goedert, M.; Spillantini, M.G.; Cairns, N.J.; Crowther, R.A. Tau proteins of Alzheimer paired helical filaments: Abnormal phosphorylation of all six brain isoforms. *Neuron* **1992**, *8*, 159–168. [[CrossRef](#)] [[PubMed](#)]
12. Espinoza, M.; de Silva, R.; Dickson, D.W.; Davies, P. Differential incorporation of tau isoforms in Alzheimer's disease. *J. Alzheimers Dis.* **2008**, *14*, 1–16. [[CrossRef](#)] [[PubMed](#)]
13. Togo, T.; Akiyama, H.; Iseki, E.; Uchikado, H.; Kondo, H.; Ikeda, K.; Tsuchiya, K.; de Silva, R.; Lees, A.; Kosaka, K. Immunohistochemical study of tau accumulation in early stages of Alzheimer-type neurofibrillary lesions. *Acta Neuropathol.* **2004**, *107*, 504–508. [[CrossRef](#)] [[PubMed](#)]
14. Cherry, J.D.; Kim, S.H.; Stein, T.D.; Pothast, M.J.; Nicks, R.; Meng, G.; Huber, B.R.; Mez, J.; Alosco, M.L.; Tripodis, Y.; et al. Evolution of neuronal and glial tau isoforms in chronic traumatic encephalopathy. *Brain Pathol.* **2020**, *30*, 913–925. [[CrossRef](#)]
15. Arendt, T.; Stieler, J.T.; Holzer, M. Tau and tauopathies. *Brain Res. Bull.* **2016**, *126*, 238–292. [[CrossRef](#)]
16. Chung, D.C.; Roemer, S.; Petrucelli, L.; Dickson, D.W. Cellular and pathological heterogeneity of primary tauopathies. *Mol. Neurodegener.* **2021**, *16*, 57. [[CrossRef](#)]
17. Shi, Y.; Zhang, W.; Yang, Y.; Murzin, A.G.; Falcon, B.; Kotecha, A.; van Beers, M.; Tarutani, A.; Kametani, F.; Garringer, H.J.; et al. Structure-based classification of tauopathies. *Nature* **2021**, *598*, 359–363. [[CrossRef](#)]
18. Prusiner, S.B. Biology and genetics of prions causing neurodegeneration. *Annu. Rev. Genet.* **2013**, *47*, 601–623. [[CrossRef](#)]
19. Guo, J.L.; Lee, V.M. Seeding of normal Tau by pathological Tau conformers drives pathogenesis of Alzheimer-like tangles. *J. Biol. Chem.* **2011**, *286*, 15317–15331. [[CrossRef](#)]
20. Guo, J.L.; Narasimhan, S.; Changolkar, L.; He, Z.; Stieber, A.; Zhang, B.; Gathagan, R.J.; Iba, M.; McBride, J.D.; Trojanowski, J.Q.; et al. Unique pathological tau conformers from Alzheimer's brains transmit tau pathology in nontransgenic mice. *J. Exp. Med.* **2016**, *213*, 2635–2654. [[CrossRef](#)]
21. Swanson, E.; Breckenridge, L.; McMahon, L.; Som, S.; McConnell, L.; Bloom, G.S. Extracellular Tau Oligomers Induce Invasion of Endogenous Tau into the Somatodendritic Compartment and Axonal Transport Dysfunction. *J. Alzheimers Dis.* **2017**, *58*, 803–820. [[CrossRef](#)] [[PubMed](#)]
22. Gibbons, G.S.; Lee, V.M.Y.; Trojanowski, J.Q. Mechanisms of Cell-to-Cell Transmission of Pathological Tau: A Review. *JAMA Neurol.* **2019**, *76*, 101–108. [[CrossRef](#)] [[PubMed](#)]
23. Vaquer-Alicea, J.; Diamond, M.I. Propagation of Protein Aggregation in Neurodegenerative Diseases. *Annu. Rev. Biochem.* **2019**, *88*, 785–810. [[CrossRef](#)]
24. Wang, Z.; Manca, M.; Foutz, A.; Martinez, M.C.; Raymond, G.J.; Race, B.; Orru, C.D.; Yuan, J.; Shen, P.; Li, B.; et al. Early Preclinical Detection of Prions in the Skin of Prion-Infected Animals. *Nat. Commun.* **2019**, *10*, 247. [[CrossRef](#)]
25. Goedert, M. Tau proteinopathies and the prion concept. *Prog. Mol. Biol. Transl. Sci.* **2020**, *175*, 239–259. [[PubMed](#)]
26. Kaufman, S.K.; Sanders, D.W.; Thomas, T.L.; Ruchinkas, A.J.; Vaquer-Alicea, J.; Sharma, A.M.; Miller, T.M.; Diamond, M.I. Tau Prion Strains Dictate Patterns of Cell Pathology, Progression Rate, and Regional Vulnerability in vivo. *Neuron* **2016**, *92*, 796–812. [[CrossRef](#)] [[PubMed](#)]
27. Dujardin, S.; Bégard, S.; Caillierez, R.; Lachaud, C.; Carrier, S.; Lieger, S.; Gonzalez, J.A.; Deramecourt, V.; Déglon, N.; Muraige, C.A.; et al. Different tau species lead to heterogeneous tau pathology propagation and misfolding. *Acta Neuropathol. Commun.* **2018**, *6*, 132. [[CrossRef](#)]
28. Vaquer-Alicea, J.; Diamond, M.I.; Joachimiak, L.A. Tau strains shape disease. *Acta Neuropathol.* **2021**, *142*, 57–71. [[CrossRef](#)]
29. Xu, H.; O'Reilly, M.; Gibbons, G.S.; Changolkar, L.; McBride, J.D.; Riddle, D.M.; Zhang, B.; Stieber, A.; Nirschl, J.; Kim, S.J.; et al. In vitro amplification of pathogenic tau conserves disease-specific bioactive characteristics. *Acta Neuropathol.* **2021**, *141*, 193–215. [[CrossRef](#)]

30. von Bergen, M.; Barghorn, S.; Müller, S.A.; Pickhardt, M.; Biernat, J.; Mandelkow, E.M.; Davies, P.; Aebi, U.; Mandelkow, E. The core of tau-paired helical filaments studied by scanning transmission electron microscopy and limited proteolysis. *Biochemistry* **2006**, *45*, 6446–6457. [[CrossRef](#)]
31. von Bergen, M.; Friedhoff, P.; Biernat, J.; Heberle, J.; Mandelkow, E.M.; Mandelkow, E. Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5129–5134. [[CrossRef](#)] [[PubMed](#)]
32. Barghorn, S.; Biernat, J.; Mandelkow, E. Purification of recombinant tau protein and preparation of Alzheimer-paired helical filaments in vitro. *Methods Mol. Biol.* **2005**, *299*, 35–51. [[PubMed](#)]
33. Wu, L.; Wang, Z.; Lad, S.; Gilyazova, N.; Dougharty, D.T.; Marcus, M.; Henderson, F.; Ray, W.K.; Siedlak, S.; Li, J.; et al. Selective Detection of Misfolded Tau From Postmortem Alzheimer’s Disease Brains. *Front. Aging Neurosci.* **2022**, *14*, 945875. [[CrossRef](#)]
34. Wu, L.; Velander, P.; Brown, A.M.; Wang, Y.; Liu, D.; Bevan, D.R.; Zhang, S.; Xu, B. Rosmarinic Acid Potently Detoxifies Amylin Amyloid and Ameliorates Diabetic Pathology in a Transgenic Rat Model of Type 2 Diabetes. *ACS Pharmacol. Transl. Sci.* **2021**, *4*, 1322–1337. [[CrossRef](#)] [[PubMed](#)]
35. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
36. Fitzpatrick, A.W.P.; Falcon, B.; He, S.; Murzin, A.G.; Murshudov, G.; Garringer, H.J.; Crowther, R.A.; Ghetti, B.; Goedert, M.; Scheres, S.H. Cryo-EM structures of tau filaments from Alzheimer’s disease. *Nature* **2017**, *547*, 185–190. [[CrossRef](#)] [[PubMed](#)]
37. Falcon, B.; Zivanov, J.; Zhang, W.; Murzin, A.G.; Garringer, H.J.; Vidal, R.; Crowther, R.A.; Newell, K.L.; Ghetti, B.; Goedert, M.; et al. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **2019**, *568*, 420–423. [[CrossRef](#)] [[PubMed](#)]
38. Falcon, B.; Zhang, W.; Murzin, A.G.; Murshudov, G.; Garringer, H.J.; Vidal, R.; Crowther, R.A.; Ghetti, B.; Scheres, S.H.; Goedert, M. Structures of filaments from Pick’s disease reveal a novel tau protein fold. *Nature* **2018**, *561*, 137–140. [[CrossRef](#)]
39. Zhang, W.; Tarutani, A.; Newell, K.L.; Murzin, A.G.; Matsubara, T.; Falcon, B.; Vidal, R.; Garringer, H.J.; Shi, Y.; Ikeuchi, T.; et al. Novel tau filament fold in corticobasal degeneration. *Nature* **2020**, *580*, 283–287. [[CrossRef](#)]
40. Velander, P.; Wu, L.; Hildreth, S.; Vogelaar, N.J.; Mukhopadhyay, B.; Zhang, S.; Helm, R.F.; Xu, B. Catechol-containing compounds are a broad class of amyloid inhibitors: Redox state is a key determinant of the inhibitory activities. *Pharmacol. Res.* **2022**, *184*, 106409. [[CrossRef](#)]
41. Shahpasand-Kroner, H.; Portillo, J.; Lantz, C.; Seidler, P.M.; Sarafian, N.; Loo, J.A.; Bitan, G. Three-repeat and four-repeat tau isoforms form different oligomers. *Protein Sci.* **2022**, *31*, 613–627. [[CrossRef](#)] [[PubMed](#)]
42. Seidler, P.M.; Boyer, D.R.; Rodriguez, J.A.; Sawaya, M.R.; Cascio, D.; Murray, K.; Gonen, T.; Eisenberg, D.S. Structure-based inhibitors of tau aggregation. *Nat. Chem.* **2018**, *10*, 170–176. [[CrossRef](#)] [[PubMed](#)]
43. Saijo, E.; Ghetti, B.; Zanusso, G.; Oblak, A.; Furman, J.L.; Diamond, M.I.; Kraus, A.; Caughey, B. Ultrasensitive and selective detection of 3-repeat tau seeding activity in Pick disease brain and cerebrospinal fluid. *Acta Neuropathol.* **2017**, *133*, 751–765. [[CrossRef](#)] [[PubMed](#)]
44. Kraus, A.; Saijo, E.; Metrick, M.A., 2nd; Newell, K.; Sigurdson, C.J.; Zanusso, G.; Ghetti, B.; Caughey, B. Seeding selectivity and ultrasensitive detection of tau aggregate conformers of Alzheimer disease. *Acta Neuropathol.* **2019**, *137*, 585–598. [[CrossRef](#)] [[PubMed](#)]
45. Metrick, M.A., 2nd; Ferreira, N.D.C.; Saijo, E.; Kraus, A.; Newell, K.; Zanusso, G.; Vendruscolo, M.; Ghetti, B.; Caughey, B. A single ultrasensitive assay for detection and discrimination of tau aggregates of Alzheimer and Pick diseases. *Acta Neuropathol. Commun.* **2020**, *8*, 22. [[CrossRef](#)]
46. Arakhamia, T.; Lee, C.E.; Carlomagno, Y.; Kumar, M.; Duong, D.M.; Wesseling, H.; Kundinger, S.R.; Wang, K.; Williams, D.; DeTure, M.; et al. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell* **2020**, *180*, 633–644. [[CrossRef](#)]
47. Wesseling, H.; Mair, W.; Kumar, M.; Schlaffner, C.N.; Tang, S.; Beerepoot, P.; Fatou, B.; Guise, A.J.; Cheng, L.; Takeda, S.; et al. Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer’s Disease. *Cell* **2020**, *183*, 1699–1713. [[CrossRef](#)]
48. Wu, L.; Gilyzova, N.; Ervin, J.F.; Wang, S.J.; Xu, B. Site-Specific Phospho-Tau Aggregation-Based Biomarker Discovery for Alzheimer’s Diagnosis and Differentiation. *ACS Chem. Neurosci.* **2022**, *13*, 3281–3290. [[CrossRef](#)]
49. Dujardin, S.; Commins, C.; Lathuiliere, A.; Beerepoot, P.; Fernandes, A.R.; Kamath, T.V.; De Los Santos, M.B.; Klickstein, N.; Corjuc, D.L.; Corjuc, B.T.; et al. Tau molecular diversity contributes to clinical heterogeneity in Alzheimer’s disease. *Nat. Med.* **2020**, *26*, 1256–1263. [[CrossRef](#)]