Terminal Capping of an Amyloidogenic Tau Fragment Modulates Its Fibrillation Propensity

Authors: Shruti Arya[a], Pritam Ganguly[a], Andrea Arsiccio[a], Sarah L Claud[b], Benjamin Trapp[c], Grace E. Schonfeld[b], Xikun Liu[a], Kristi Lazar Cantrell[b], Joan-Emma Shea[a*] and Michael T. Bowers[a*]

[a] Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, CA, United States

[b] Department of Chemistry, Westmont College, Santa Barbara, CA, United States

[c] Neon Therapeutics, Cambridge, MA, United States

*Corresponding authors

Michael T. Bowers

E-mail: bowers@chem.ucsb.edu

Joan-Emma Shea

E-mail: shea@chem.ucsb.edu

ABSTRACT

Aberrant protein folding leading to the formation of characteristic cross-β-sheet rich amyloid structures is well-known for its association with a variety of debilitating human diseases. Often, depending upon amino acid composition, only a small segment of a large protein participates in amyloid formation and is in fact capable of self-assembling into amyloid, independent of the rest of the protein. Therefore, such peptide fragments serve as useful model systems for understanding the process of amyloid formation. An important factor that has often been overlooked while using peptides to mimic full length protein, is the charge on the termini of these peptides. Here, we show the influence of terminal charges on the aggregation of an amyloidogenic peptide from microtubule associated protein Tau, implicated in Alzheimer's disease and Tauopathies. We found that modification of terminal charges by capping the peptide at one or both of the termini drastically modulates the fibrillation of the hexapeptide sequence PHF6 (paired helical filament 6) from repeat 3 of Tau, both with and without heparin. Without heparin, the PHF6 peptide capped at both termini and PHF6 capped only at the N-terminus, self-assembled to form amyloid fibrils. With heparin, all capping variants of PHF6, except for PHF6 with both termini free, formed typical amyloid fibrils. However, the rate and extent of aggregation both with and without heparin as well as the morphology of aggregates were found to be highly dependent on the terminal charges. Our molecular dynamics simulations on PHF6 capping variants corroborated our experiments and provided critical insights into the mechanism of PHF6 self-assembly. Overall, our results emphasize the importance of terminal modifications in fibrillation of small peptide fragments and provide significant insights into the aggregation of a small Tau fragment which is considered essential for Tau filament assembly.

INTRODUCTION

The incorrect folding of proteins leading to the formation of β-sheet rich oligomeric structures that subsequently convert to amyloid fibrils is implicated in the pathogenesis of a large number of debilitating neurodegenerative diseases and systemic disorders.[1–8] Even though the amyloid fibrils formed by different proteins share the common characteristic of highly ordered cross-β sheet structure, the underlying amino acid sequence is believed to have a strong influence on the fibrillation propensity of a protein.[7,9,10] In many instances, often depending upon amino acid sequence composition, only a small segment of a large protein is prone to fibrillation and hence participates in amyloid core formation and in most cases, can self-assemble to form amyloid independent of the rest of the protein.[5,11] Therefore, select small peptide fragments of large proteins can serve as useful model systems. Additionally, many peptide fragments are produced by proteolytic cleavage inside the cell under healthy as well as under diseased conditions[12–15] and therefore studying these fragments is of paramount importance from the standpoint of both function and disease.

The fragments of large proteins or peptides associated with amyloid related diseases have been extensively studied to elucidate the mechanism of amyloid formation.[16–19] There are several studies that have investigated the effect of amino acid composition, side chain interactions and solvent conditions on the aggregation behavior of amyloidogenic peptide fragments.[20–25] However, there are only a few reports that have highlighted the influence of the termini of the small peptide fragments on their aggregation propensity and fibril morphology.[18,19,26,27] The charges on the termini of smaller peptide fragments are believed to have a prominent effect on the peptide hydrophobicity, hydrogen bonding pattern, electrostatic interactions, and the propensity of the peptide towards secondary structure formation. Hence, capping can significantly influence the intricate balance between peptide-solvent and peptide-peptide interactions that underlie the major physical driving force for protein aggregation. Additionally, the free termini introduce extra charges on the peptide, which makes an uncapped peptide a poor model of the full-length protein. In this study, we focus on the effect of termini capping on the fibrillation propensity of an amyloidogenic peptide fragment derived from the microtubule associated protein Tau. The misfolding of Tau is associated with Alzheimer's disease and several other neurodegenerative diseases collectively known as "Tauopathies".[28–33] Tau is the primary component of neurofibrillary tangles (NFTs) found in the brain of Alzheimer's disease patients.[31] It is primarily expressed in

the central and peripheral nervous system.[31] The full-length Tau protein is intrinsically disordered and can be divided into an N-terminal projection domain, a proline rich region, a repeat region, and a C-terminal domain.[30,34,35] There are six isoforms of Tau in the adult human brain that differ in the number of N-terminal inserts (0-2) in the N-terminal half and either contain or lack a 31-residue long imperfect repeat (R2) which is one of the four imperfect repeats (R1-R4) present in the C-terminal half of the longest Tau isoform.[36] The inclusion of R2 results in the production of three isoforms that contain four repeats in total (4R-Tau) and exclusion of R2 gives rise to three isoforms containing three repeats in total (3R-Tau). The repeat sequences constitute a major part of the microtubule binding region in Tau and are also known to influence the propensity of Tau to form tangles, which suggest that the physiological function and pathological assembly of Tau may be mutually exclusive.[31] NFTs found in AD are a mixture of 3R and 4R isoforms.[37] Unlike other amyloid associated proteins or peptides, Tau is a highly soluble protein and has a low tendency to self-assemble in vitro under physiologically relevant pH and temperature conditions.[38] However, in the presence of polyanionic co-factors such as heparin and RNA it readily undergoes aggregation[39,40] but the exact nature of underlying interactions between the co-factors and Tau is not well understood. Also, the biological significance of these polyanionic co-factors remains unclear. The two short hexapeptide sequences VQIINK and VQIVYK that are located in R2 and R3, respectively are considered crucial for Tau filament assembly.[41,42] Additionally, a recent cryo-EM study on Tau filaments isolated from AD patients has revealed that the Tau filament core is comprised of residues 306-378 and the VQIVYK sequence from the R3 repeat forms the first β-strand of this filament core.[43] The VQIVYK sequence is commonly referred to as paired helical filament 6 (PHF6). While all repeats of Tau are considered important for microtubule binding, the repeat R3 is believed to play a critical role in inducing Tau aggregation and PHF6 is considered essential for Tau filament assembly.[44,45] In addition, it has been found that endopeptidase cleavage can lead to in-situ generation of fragments containing R3 in human brain.[46] Moreover, the R3 repeat is part of all six Tau isoforms and has been shown to self-assemble into amyloid.[17] Most of the aggregation studies on PHF6 and R3 fragments have been carried out in the presence of heparin as an anionic co-factor.[47–50] Furthermore, even though the aggregation behavior of various Tau fragments has been investigated extensively,[11,17,47–57] the importance of the charges on the termini of these fragments has either been overlooked or has been underestimated. Therefore, in this study,

we have investigated the effect of terminal charges on the self-assembly of four capping variants of Tau peptide fragment PHF6 (Scheme 1), both with and without heparin.

MATERIALS AND METHODS

Materials

Ammonium acetate, thioflavin-T (ThT) and sodium chloride were purchased from EM science (Merck), Sigma-Aldrich and Fisher Scientific, respectively. Heparin (H4784) was purchased from Sigma-Aldrich. All solutions were prepared in 10mM ammonium acetate buffer, which was prepared with HPLC grade water (Alfa Aesar). The pH of the ammonium acetate buffer was adjusted to 7.0 using a Thermo Fisher pH meter.

Methods

Peptide Synthesis and Purification

PHF6 was purchased from Genscript inc., R3 was purchased from Anaspec (Catalog number AS-65435-1), and Ac-R3-NH_2 was purchased from Aapptec. The PHF6 peptides were synthesized using standard 9-fluorenylmethoxycarbonyl (Fmoc) and HBTU/HOBt manual solid-phase synthesis. PHF6 and Ac-PHF6 were synthesized on an Fmoc-Wang Resin (Aldrich) that was preloaded with lysine. PHF6-NH_2 and Ac-PHF6-NH_2 were synthesized on a Fmoc-Rink Amide Resin (Anaspec). Ac-PHF6 and Ac-PHF6-NH_2 were acetylated with acetic anhydride prior to cleavage from the resin. The PHF6 peptides were cleaved from the resin using 94% TFA, 5% of triisopropylsilane and 1% phenol for 2 h at 22°C. Crude peptides were purified by RP-HPLC on a semi-preparative C18 column (Phenomenex) using gradients of water (0.1% v/v TFA) and acetonitrile (0.1% v/v TFA). Peptide purity was greater than 95% by analytical RP-HPLC. The molecular masses of the peptides were verified by ESI mass spectrometry. All the peptides were dissolved in HFIP and were dried under vacuum for ~24 h. The dry peptides were stored at -20°C. Prior to experiments, the dried peptides were dissolved in the desired buffer (see below).

Determination of Peptide Concentrations

Peptide concentrations were calculated using the extinction coefficient for the tyrosine residue (ε_{275} = 1,420 M^{-1} cm^{-1}). The absorbances were measured on a Thermo Electron Corporation Nicolet Evolution 300.

ThT Fluorescence Assay

The ThT fluorescence assay was performed on a BioTek Synergy 2 multimode plate reader in kinetic mode for 24 hours. The fibril assays were performed at 37ºC by incubating the peptides at 150 μM in 10 mM ammonium acetate buffer (pH 7.0) containing either 150 mM NaCl or 15 μM heparin and 50 μM ThT. The excitation and emission filters used were 440/30 nm and 485/20 nm, respectively. The PMT gain used was either 55 or 60. The buffer (10 mM ammonium acetate, pH 7.0) containing either 150 mM NaCl or 15 μM heparin and 50 μM ThT was used as baseline and the data were always corrected for the baseline. For the 2-week time point, the samples were incubated at 37ºC under the desired experimental conditions (mentioned above) and the fluorescence was recorded in the plate reader using the same excitation and emission filters. All data were then plotted and smoothened in Origin software.

Far-UV Circular Dichroism

CD data were collected on a Jasco J-810 CD spectrometer equipped with a temperature-controlled holder. The CD spectra were collected at 1 nm intervals from 260 to 195 nm with a 4-s response time and a 1 nm bandwidth. All spectra were collected in a 1 mm quartz cuvette at 37ºC. The CD spectra of freshly dissolved (t = 0) peptides were measured at a concentration of 150 μM in 10 mM ammonium acetate, pH 7.0, containing either 150 mM NaCl or 15 μM heparin. The samples were incubated at 37ºC in between time points and run again after two weeks to detect the formation of β-sheet. Ten scans were collected at each time point, and the data were averaged. A buffer baseline was subtracted from the averaged data and the curves were base-line corrected. The data were then plotted in the Origin software.

Transmission Electron Microscopy

Fibril assays were performed by incubating the peptides at 150 μM in 10 mM ammonium acetate buffer (pH 7.0) containing either 150 mM NaCl or 15 μM heparin. The samples were incubated under quiescent conditions at 37ºC for two weeks prior to EM imaging. Aliquots of fibril samples (5 μL) were applied to a glow-discharged, 400-mesh, carbon-coated support film and stained with 1% uranyl acetate. Images of fibrils at a magnification of 137,200x were recorded using a FEI Tecnai F30 STEM Microscope operated at 300 kV and equipped with a Gatan CCD digital micrograph.

Molecular Dynamics Simulations

All-atom molecular dynamics (MD) simulations were performed using the GROMACS (versions 5.1.2 and 2016.3) package.[58] Aggregation propensities of the PHF6 peptide with different capping conditions in 1 M NaCl were studied using unbiased MD simulations. For the unbiased MD simulations, large cubic boxes with ~25.4 nm linear dimension containing 25 peptides, 9860 Na+ and 9860 Cl- ions, and ~531,000 water molecules were studied. Additional Cl- ions were added to assure electroneutrality of the systems as follows: 25 for Ac-PHF6-NH2 and PHF6 (uncapped) peptides, 50 for PHF6-NH2 peptide and 0 for Ac-PHF6 peptide. The concentrations of the peptides were ~2.5 mM. The peptides were simulated using the OPLS-AA force field.[59-61] Water molecules were modeled with the rigid TIP3P force field.[62] For Na+ and Cl- ions we reparameterized a Kirkwood-Buff derived force field (Smith force field)[63] in order to reproduce experimental activity derivative of aqueous NaCl solutions. The details of the development of the NaCl force field can be found below. Periodic boundary conditions in the x, y and z directions were applied. 500 ns long NpT simulations at 300 K temperature and 1 bar pressure were performed. The temperature of the systems was controlled using the Nosé-Hoover thermostat with a time constant of 0.5 ps.[64,65] The pressure of the systems was maintained utilizing the Parrinello-Rahman barostat with a time constant of 3 ps.[66] Prior to the production runs, 5 ns long NpT simulations using the Berendsen thermostat (0.5 ps time constant) and Berendsen barostat (3 ps time constant) were performed.[67] The trajectories were generated using a leap-frog integrator with 2 fs time-steps.[68] The last 250 ns of the trajectories were used for the calculations of the inter-peptide hydrogen bonds. Coulombic interactions were calculated using the Particle Mesh Ewald method with a grid-spacing of 0.12 nm.[69] For all the non-bonded interactions, a cut-off of 1.2 nm was used. The bonds in the peptides were constrained using the LINCS algorithm.[70] The SETTLE algorithm was used to keep the water molecules rigid.[71] The hydrogen bonds were identified using a linear cut-off of 0.25 nm for the hydrogen-acceptor distance and a 30-degree cut-off for the hydrogen-donor-acceptor angle. The peptide clusters were identified using the backbone (N-C_α-C) atoms with a cut off distance of 0.70 nm. The radius of gyration and the asphericity of the clusters were also calculated based on the backbone atoms. The secondary structures of the peptides were calculated with the DSSP algorithm (using Gromacs do_dssp plugin)[72].

The Smith force field for NaCl uses the SPC/E water model with a 0.75 scaling factor for the cation-water van der Waals interactions to reproduce the derivative of the activity coefficient of

NaCl. The developers of the Smith force field had shown that the force field predicts a very low molar activity derivative (1.18) relative to experiment (1.63) when the TIP3P water is used at 4 M concentration. Since in our simulations the TIP3P water is used, we used the Smith NaCl force field with a re-parameterized scaling factor (0.65) for the cation-water van der Waals interactions. The new parameterization yields a molar activity derivative for 4 M NaCl solutions, very close (1.62) to experiment. For comparison, the original Smith NaCl force field with SPC/E water predicts the corresponding value as 1.47.[63] With the newly parametrized force field the molar activity derivative of 1 M NaCl solution was calculated to be 1.08, which is also reasonably close to the experimental value (1.01).[63] To develop the NaCl force field, we used cubic boxes with linear dimensions of ~5.8 nm. NpT simulations with the Nosé-Hoover thermostat[73, 74] and the Parrinello-Rahman barostat[66] were run for 70 ns and 20 ns respectively for 1 M and 4 M NaCl solutions. These simulations were preceded by 5 ns long NpT simulations with the Berendsen thermostat and barostat.[67] The activity derivatives were calculated from the Kirkwood-Buff integrals (KBIs) between the ions and between the ions and water. All the ions were grouped together to calculate the KBIs.[75] The last 65 ns and 15 ns of trajectories were used for the calculation of the activity derivatives in 1 M and 4 M NaCl solutions respectively. The KBIs were corrected using the method by Ganguly and van der Vegt.[76] The values of the KBIs were obtained by averaging the values between 1.0 nm and 1.4 nm. The other simulation parameters are the same as the simulations of the peptides.

RESULTS AND DISCUSSION

ThT fluorescence shows that terminal capping influences the kinetics of fibrillation

In order to compare the fibrillation propensities of capping variants of PHF6, we monitored the fibrillation kinetics in continuous mode for 24 hours using Thioflavin T (ThT), which is a well-known amyloid marker.[77] The capping variants of PHF6 were incubated with and without heparin. The peptides incubated without heparin were dissolved in 10 mM ammonium acetate (pH 7.0) containing 150 mM NaCl. The peptides incubated with heparin were dissolved in 10 mM ammonium acetate (pH 7.0) containing 15 μM heparin. *Without heparin*, Ac-PHF6-NH_2, which is acetylated at the N-terminus and amidated at the C-terminus showed an increase in ThT fluorescence within 2 hours (Figure 1A). However, no increase in ThT fluorescence was observed

for PHF6 (free N- and C-termini), Ac-PHF6 (acetylated at the N-terminus) and PHF6-NH_2 (amidated at the C-terminus) (Figure 1A). Upon 2 weeks of incubation, a small increase in ThT fluorescence was observed for Ac-PHF6 (Figure 1B). Thus, from this set of data, we can conclude that the extent of increase in ThT fluorescence is less for Ac-PHF6 when compared to Ac-PHF6-NH_2 but both are able to convert to ThT active (amyloid competent) species (Figure 1B). On the contrary, PHF6-NH_2 and PHF6 remain ThT inactive even after 2 weeks of incubation (Figure 1B). *With heparin*, Ac-PHF6-NH_2 immediately showed an increase in ThT fluorescence (Figure 1C). However, no increase in ThT fluorescence was observed for other PHF6 capping variants (Figure 1C). Upon 2 weeks of incubation, an increase in ThT fluorescence was observed for Ac-PHF6 (Figure 1D). Thus, both with and without heparin, Ac-PHF6 and Ac-PHF6-NH_2 convert to ThT active species (Figure 1D).

CD measurements show the differences in the secondary structure of PHF6 capping variants

Without heparin, PHF6, Ac-PHF6, and PHF6-NH_2 did not show any significant change in secondary structure even after 2 weeks (Figure 2A). The fully capped variant, Ac-PHF6-NH_2 showed a small change in the secondary structure initially (Figure S1A). However, after 2 weeks, a significant change in the secondary structure from predominantly random coil to a mixture of random coil, α-helix and β-sheet rich structure was observed (Figure 2A). *With heparin*, Ac-PHF6-NH_2 immediately attained β-sheet rich structure (Figure S1B). Interestingly, after 2 weeks, the CD structure changed from predominantly β-sheet to a mixture of random coil, α-helix and β-sheet, akin to the CD signature observed without heparin (Figure 2B). The singly capped variants, Ac-PHF6 and PHF6-NH_2 initially retained their disordered structure (Figure S1B). However, after 2 weeks, they converted into β-sheet, and α-helix rich structures, while retaining some disordered (random coil) structure (Figure 2B). It is interesting that the CD signature observed by us for Ac-PHF6-NH_2 with/without heparin after 2 weeks shows resemblance to the CD structure observed for paired helical filaments (PHFs) isolated from AD patients which also has high α-helical content.[78] The CD spectra observed for Ac-PHF6 and PHF6-NH_2 with heparin are different from that of Ac-PHF6-NH_2 (Figure 2B). The CD signal for PHF6-NH_2 was found to be particularly different and is indicative of more β-sheet content.

TEM reveals distinct morphological differences between the aggregates formed by PHF6 capping variants

Without heparin, both Ac-PHF6-NH_2 and Ac-PHF6 showed the formation of regular amyloid fibrils (Figure 3B, 3D). Upon careful examination of the TEM images, it appears that while Ac-PHF6 predominantly forms fibrils that have a more twisted morphology and resemble paired helical filaments (PHFs), Ac-PHF6-NH_2 forms mostly straight filaments along with some PHFs. The fibrils formed by Ac-PHF6 also appeared thicker than the Ac-PHF6-NH_2 fibrils. No fibrils or oligomeric structures were observed for PHF6 and PHF6-NH_2 (Figure 3A, 3C), which agrees with our ThT and CD data. *With heparin*, all PHF6 capping variants, except for PHF6 (with free N- and C-termini), showed the formation of regular amyloid fibrils (Figure 3E-3H). However, the fibrils were most dense in case of Ac-PHF6-NH_2. In the case of PHF6-NH_2, we observed only a few isolated fibrils (Figure 3G). Even in the presence of heparin, we observed differences in the morphologies of fibrils due to terminal capping. In fact, the fibrils formed with heparin appeared different from the fibrils formed without heparin. For example, Ac-PHF6 formed thicker fibrils *without heparin* (Figure 3B) while the fibrils formed *with heparin* were comparatively thinner and were a mixture of straight filaments and PHFs (Figure 3F). Also, *with heparin*, Ac-PHF6-NH_2 predominantly formed fibrils that resembled PHFs (Figure 3H).

Overall comparison of fibrillation propensities of PHF6 capping variants based on our experimental data

Taken together, our results suggest that under identical experimental conditions: (i) the fibrillation propensities for PHF6 fragments are highly capping dependent and follow this order: (a) *without heparin* Ac-PHF6-NH_2 > Ac-PHF6 > PHF6-NH_2 ≅ PHF6 and (b) *with heparin* Ac-PHF6-NH_2 > Ac-PHF6 > PHF6-NH_2 > PHF6 (Table 1). (ii) No fibril formation or any change in secondary structure is observed for PHF6 (free N- and C-termini), whether or not heparin is present (Table 1). (iii) In the case of PHF6-NH_2, although the ThT fluorescence does not change after 2 weeks when heparin was present, CD data showed a change in secondary structure from random coil to partial β-sheet rich structure and TEM imaging showed the formation of a few isolated fibrils. (iv) Terminal capping also had a strong influence on the final fibrillar morphologies, both *with* and *without heparin*.

Collectively, our experimental data strongly indicate that *without heparin*, the charge on the termini play a crucial role in assembly and subsequent fibrillation of the PHF6 fragment. In fact,

even in the presence of the anionic co-factor heparin, the fibril formation appears to be strongly influenced by the terminal charges. The introduction of positive charge on the N-terminus and negative charge on the C-terminus completely eliminates the amyloidogenicity of the PHF6 fragment whether or not heparin is present. When positive charge is present on the N-terminus and the C-terminus is amidated (PHF6-NH_2), the amyloidogenicity of PHF6 is only recovered slightly upon the addition of heparin. In the case of Ac-PHF6, the charged C-terminus makes it less aggregation prone than fully capped Ac-PHF6-NH_2 but the Ac-PHF6 fragment eventually forms fibrils. However, the morphology of fibrils formed by Ac-PHF6 is different from Ac-PHF6-NH_2 (Figure 3B, 3D). Thus, our data suggest that while charges on both termini make PHF6 less aggregation prone, the charge on the N-terminus particularly seems to play an important role in determining the fibrillation propensity of PHF6.

MD simulations provide critical mechanistic and structural insights into the self-assembly of PHF6 capping variants

In the past, MD simulations of Tau fragments were used to predict local structural changes[57,79,80] and to identify the aggregation pathways and the structure of oligomers.[55,57,81-89] One of the major challenges in studying aggregation of peptides through all-atomistic MD simulations is to achieve experimentally relevant peptide concentrations, which may typically be 10^3 to 10^6 times lower than the MD simulations.[90] By studying the aggregation of 25 peptides in a cubic box with linear dimension of ~25.5 nm, we have achieved a peptide concentration of 2.5 mM, which is >15 times higher than the peptide concentrations used in ThT, TEM and CD experiments. However, the relatively low concentration of the peptides in our MD simulations allow qualitative comparison of the aggregation propensities of the peptides as observed in the MD simulations and in the experiments. Our simulations of 500 ns allow us to probe the molecular mechanism of aggregation and the structures of the aggregates in their early stages.

In Figure S2A, we plot the average number of peptide aggregates versus cluster size (N). We found that the probability of forming larger aggregates (N>6) is significantly higher for Ac-PHF6-NH_2 and Ac-PHF6 compared to PHF6-NH_2 and PHF6. The total number of hydrogen bonds formed within the peptide clusters qualitatively correlates with their aggregation propensities (Figure S2B). Further, the number of hydrogen bonds formed agrees well with our ThT experimental data for the aggregation of PHF6 capping variants without heparin: Ac-PHF6-NH_2 > Ac-PHF6 > PHF6-NH_2 ≅ PHF6.

Morphology of the PHF6 aggregates:

Next, we studied the variation in the radius of gyration, R_g, of the clusters versus cluster size. We fit R_g using Equation (1)

$$R_g = A N^m \quad (1)$$

where A is an undetermined prefactor and the exponent m is determined by plotting R_g versus N in logarithmic scale (Figure 4). The fractal dimension (d_f)[73,91] of the clusters is defined as $m=1/d_f$. We found that R_g follows Eqn (1) with distinct exponents for different capping conditions. The corresponding exponents for Ac-PHF6, PHF6-NH_2, and PHF6 are 0.38 (d_f=2.6), 0.31 (d_f=3.2), and 0.43 (d_f=2.3) respectively. Interestingly, we found that Ac-PHF6-NH_2 follows two distinct power laws for two different cluster size regimes: From N=1 to 5 m = 0.43 and from N=8 to 12 m = 0.82 (Figure 4A). The increase in R_g with N saturates at N=7 before the discontinuous change in the power law at N>8. This crossover behavior of Ac-PHF6-NH_2 as N increases is further discussed below utilizing the three-dimensional structures of the peptide aggregates. It is also discussed how the structures of the uncapped PHF6 aggregates are significantly different than the clusters of Ac-PHF6-NH_2 despite of the fact that the growth exponent for Ac-PHF6-NH_2 (N=1 to 5) is similar to that of PHF6.

In order to understand the shape of the peptide clusters, we have further calculated the asphericity of the peptide aggregates. By calculating the corresponding radii of gyration (R_1, R_2 and R_3) along the principal axes of the aggregates, the asphericity (D) can be calculated as given in Equation (2)

$$D = 1 - 3\frac{\min(R_1,R_2,R_3)}{R_1+R_2+R_3} \quad (2)$$

where min(R_1, R_2,R_3) denotes the minimum among R_1, R_2 and R_3. An asphericity of 0 corresponds to a spherical cluster. The probability distributions of the asphericty of the peptide clusters with the size of the clusters (N) for the different capping conditions are plotted in Figure S3. We find that for all PHF6 capping variants, the asphericity of the aggregates reduces with the increase in the number of the peptide monomers in the aggregates and the larger aggregates (N>4) assume near-spherical shapes. However, for Ac-PHF6-NH_2, an abrupt increase in the asphericity is observed for the clusters of size N>8. This discontinuity in the asphericity is consistent with the crossover behavior of the power law of R_g with the cluster growth at N>8.

Three-dimensional structures of the peptides and their aggregates: We have further calculated the secondary structures of the peptides where the most stable cluster for each cluster size are

chosen for the calculations. In Figure S4A in the Supporting Information, we plot the average number of amino acids with secondary β-structures in a cluster along with the cluster size N, where the β-sheet and β-bridge structures are considered together. We plot the same in Figure S4B where the alpha-helical structures are added to the data presented in S4A. For Ac-PHF6, a moderate increase in the β-content is observed with the increasing cluster size. Both PHF6 and PHF6-NH_2 do not form large clusters. However, smaller oligomers (N<=7) observed for all peptides and for PHF6-NH_2 in particular, show β-structures. No significant alpha-helical structures are found for Ac-PHF6, PHF6-NH_2 and PHF6. Interestingly, for Ac-PHF6-NH_2, the number of β-structures first increases with N till N=4, then decreases till the cross-over at N=7. In Figure 5A-G, we show representative snapshots of the most stable clusters of AcPHF6-NH_2 peptides with different cluster sizes. Upon comparing the structures of the dimers, trimers and the tetramers, we find that the growth of the peptide aggregates (N<=4) occurs with a semi-planar geometry with the tetramers assuming curved β-structures. At N=5, the β-structure is distorted and the cluster assumes a more spherical structure. Figure S4B in the Supporting Information is almost identical to S4A except for Ac-PHF6-NH_2 at N=7, where the peptides assume prominent alpha-helical structures. The overall shape of the cluster at N=7 is found to be cylindrical (Figure 5E). At this cluster size, the three-dimensional growth of the Ac-PHF6-NH_2 aggregate saturates as observed from the plot between Rg and N in Figure 4. The fully-capped Ac-PHF6-NH_2 peptides are significantly more structured with β-content at higher order oligomers (N>8). From the structures of the larger aggregates (N>10), we find that the larger clusters are composed of two smaller moieties comprised of 7-8 peptides and 3-4 peptides respectively. The overall growth of the clusters beyond N=8 tends to be linear which explains the discontinuity in the asphericity of the aggregates. It must also be noted that the increase in Rg with N for N>8 corresponds to a fractal dimension of 1.2 and a fractal dimension close to 1 corresponds to linear growth of a cluster.[74] The linear growth of the peptide aggregate may potentially lead to longer fibril formation. In comparison, the Ac-PHF6 variant of the peptide aggregates into amorphous distorted spherical clusters for N>4 (Figure S5), although the individual peptides show β-structures (Figure S4). Additional snapshots of the smaller clusters (N<4) formed by the PHF6 and the PHF6-NH_2 peptides are shown in Figures S6 and S7, respectively.

Taken together, our MD results correlate well with our CD experiments where a mixture of β-structures and alpha-helical structures for Ac-PHF6-NH_2 was observed. The β-structures arise

predominantly from the higher order oligomers. During the transition between amorphous lower order oligomers and structured higher order oligomers, the peptides assume alpha-helical structures. Our MD results also indicate that β-structures may form in the lower order oligomers of Ac-PHF6 and PHF6-NH_2, which further explains why an enhancement in the β-sheet content was observed with heparin for these two peptides. However, in contrast to the CD experiments with heparin, no alpha-helical structures for Ac-PHF6 were observed in the MD simulations.

Molecular mechanism that underlies PHF6 aggregation:

A detailed analysis of the hydrogen bond formation between the individual amino acids provides significant insight into the mechanism of self-assembly. The probability distributions of the hydrogen bond contacts (considering the backbone and the side chain atoms together), shown in Figure 6(A-D), indicate the mutual affinities and the cross-peptide linkages between the amino acid residues in the peptide aggregates. For Ac-PHF6-NH_2, strong interactions were observed between the polar residues glutamine (Q) and tyrosine (Y), and also between the hydrophobic residues valine (V) and isoleucine (I) (Figure 6A). The interactions were strongest between the Qs and between Q and Y. Unlike Ac-PHF6-NH_2, Ac-PHF6 has a negatively charged C-terminus. However, the positively charged side chain of K, adjacent to the negatively charged C-terminus facilitates a strong interaction between these lysine residues when the C-terminus is uncapped. This observation is consistent with the formation of parallel dimers in the Ac-PHF6 peptides (Figure S5A). In addition, for Ac-PHF6, we observed interaction of all residues with residue Q, presumably due to its longer side chain (Figure 6C). The strong interaction between the Qs was also observed for Ac-PHF6-NH_2 (Figure 6A) and both these peptides had higher propensity towards the formation of larger aggregates.

In PHF6, a strong interaction of lysine (K) with both K and V was observed (Figure 6B). The positively charged N-terminus and the negatively charged C-terminus of an adjacent peptide result in anti-parallel peptides for dimeric and trimeric PHF6 (Figure S6A, S6B). However, the positively charged lysine side chain adjacent to the negatively charged C-terminus facilitates interactions between the C-terminus and the side chain of the lysine on an adjacent peptide resulting in parallel tetrameric aggregates (Figure S6C). Like PHF6, PHF6-NH_2 has a positive charge on the N-terminus, however this peptide experiences strong repulsion between the N-terminus and the positively charged lysine side chain near the C-terminus. This repulsion results in minimal attraction between the Q and V residues on adjacent peptides (Figure 6D). The positive net charge

on the PHF6-NH_2 peptide might also hinder the formation of the larger aggregates due to Coulombic repulsion.

Taken together, our simulation data emphasize the importance of the positively charged lysine residue and the charge on the N-terminus. Whenever N-terminus is positively charged the lysine residue at the C-terminus experiences repulsion, whether or not the C-terminus is free. Additionally, our MD simulations suggest that the strong interactions among polar residues and hydrophobic residues are drivers for fibrillation as these interactions are observed for Ac-PHF6-NH_2 and Ac-PHF6, both of which convert to amyloid fibrils. However, the morphology of the resulting fibrils is significantly different for the Ac-PHF6-NH_2 and the Ac-PHF6 peptides (Figure 3). Therefore, our MD data also suggest that the terminal charges not only influence the fibrillation propensity of PHF6 peptide but also affect the overall fibrillar morphology and their formation mechanism.

The terminal capping also influences the fibrillation of a cytotoxic 31-residue long repeat 3 (R3) fragment of Tau

Terminal capping also had a significant influence on the fibrillation of the full length R3 repeat fragment (Scheme 1). *Without heparin*, capped R3 fragment Ac-R3-NH_2 (acetylated at the N-terminus and amidated at the C-terminus) began to show an increase in ThT fluorescence after 14-16 hours, but uncapped R3 (free N- and C-termini) did not show any increase within 24 hours of continuous monitoring (Figure 7A). Thus, our ThT kinetics data for R3 variants is consistent with PHF6 where uncapped or singly capped variants did not show any increase in ThT within 24 hours while fully capped PHF6 (Ac-PHF6-NH_2) showed an immediate increase in ThT fluorescence. However, after 2 weeks, ThT fluorescence was similar for both R3 variants (inset in Figure 7A and Figure S8A) but the morphologies of fibrils were found to be highly influenced by terminal charges. The uncapped R3 showed the formation of fibrils that resembled the fibrils formed by Ac-PHF6 (Figure 7B). The fibrils formed by Ac-R3-NH_2 were similar to Ac-PHF6-NH_2 fibrils (Figure 7C). Interestingly, *with heparin*, unlike PHF6, both Ac-R3-NH_2 and R3 showed an immediate increase in ThT fluorescence (Figure 7D) and after 2 weeks ThT fluorescence was similar for both variants (Figure S8B). Also, both variants showed the formation of amyloid fibrils (Figure 7E, 7F). The R3 peptide has a net positive charge and therefore we believe that by screening the positive charge, heparin is able to reduce the interpeptide repulsion that might be slowing its aggregation in the absence of heparin.

SUMMARY

In summary, we have shown that aggregation of the important PHF6 peptide fragment of the intrinsically disordered protein Tau is highly dependent on the charges on its termini. The introduction or removal of charges not only affects the fibrillation propensity and final fibrillar morphology but also alters the mechanism of aggregation as revealed by our MD simulations. While the charges on both termini are important, the charge on the N-terminus particularly plays an important role in PHF6 aggregation. Additionally, the positively charged residue K present at the C-terminus appears to be critical. It is highly likely that replacing this residue with an uncharged or negatively charged residue could increase the fibrillation propensity of all PHF6 capping variants. In fact, it has been shown previously that the substitution of K by glutamic acid (E) resulted in gel formation, suggesting that the peptide rapidly formed amyloid.[45] Interestingly, a recent study on the Tau filaments isolated from AD patients has shown that the residue K at position 311 in PHF6 is acetylated and/or ubiquitinated.[92] The acetylation of residue K would result in the neutralization and promote Tau aggregation. Our results that reflect a strong influence of charges and a critical role of positively charged residue K on Tau aggregation could account for the effect of such post translational modifications on final fibrillar structure. Furthermore, our results underscore the importance of the polar residues Q and Y, and hydrophobic residues V and I in PHF6 fibrillation. The significance of the VYK sequence within PHF6 in the context of amyloid formation has been highlighted previously.[45] Moreover, it is interesting to note that even for full length repeat 3 within Tau (R3), which is much longer than PHF6, the effect of capping the termini was fairly significant. In fact, it is possible that modifying the termini of amyloid associated peptides might provide a promising handle for controlling the fibrillation propensities of amyloidogenic protein/peptide fragments. Depending upon the amino acid sequence and charges on the peptide, the modification on either terminus might result in the production of non-amyloid forming species that are unable to readily convert to fibrils. These species might then act as potent inhibitors of full-length protein fibrillation by preventing the interactions necessary for amyloid formation.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the National Science Foundation (NSF) under grant CHE-1565941 (M.T.B.) and MCB-1716956 (J.-E.S.), the Allan Nishimura Summer Research Fund (KLC) and the Stauffer Summer Chemistry Endowment (SLC). J.-E.S. also acknowledges the Center for Scientific Computing at the California Nanosystems Institute (NSF Grant CNS-1725797) and the National Institutes of Health (NIH Grant R01-GM118560-01A, PI: Kevin Plaxco). The computer simulations were performed on the Stampede 2 supercomputer located at the Texas Advanced Computing Center through the Extreme Science and Engineering Discovery Environment (XSEDE) grant TG-MCA05S027 from NSF. We thank Prof. Songi Han for giving us access to the plate reader for ThT fluorescence assays and Yimei Chen of the Electron Microscopy Facility at the University of Chicago for help collecting TEM images.

SUPPORTING INFORMATION

The supporting information is available free of charge on the ACS publication website. Additional circular dichroism data (CD) (Figure S1), MD simulations (Figure S2-S7) and ThT fluorescence (Figure S8) data are presented.

REFERENCES

(1) Iadanza, M. G.; Jackson, M. P.; Hewitt, E. W.; Ranson, N. A.; Radford, S. E. A New Era for Understanding Amyloid Structures and Disease. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (12), 755–773.

(2) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27–68.

(3) Riek, R.; Eisenberg, D. S. The Activities of Amyloids from a Structural Perspective. *Nature* **2016**, *539* (7628), 227–235.

(4) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The Amyloid State and Its Association with Protein Misfolding Diseases. *Nat. Rev. Mol. Cell Biol.* **2014**, *15* (6), 384–396.

(5) Eisenberg, D.; Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **2012**, *148* (6), 1188–1203.

(6) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.

(7) Chiti, F. ; Stefani, M. ; Taddei, N. ; Ramponi, G. ; Dobson, C. M. Rationalization of the Effects of Mutations on Peptide and Protein Aggregation Rates. *Nature* **2003**, *424* (6950), 805-808.

(8) Dobson, C. M. Protein Misfolding, Evolution and Disease. *Trends Biochem. Sci.* **1999**, *24* (9), 329–332.

(9) Tartaglia, G. G.; Vendruscolo, M. The Zyggregator Method for Predicting Protein Aggregation Propensities. *Chem. Soc. Rev*. **2008**, *37* (7), 1395–1401.

(10) López de la Paz, M.; Serrano, L. Sequence Determinants of Amyloid Fibril Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (1), 87–92.

(11) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T. et al. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447* (7143), 453-457.

(12) Weihofen, A.; Martoglio, B. Intramembrane-Cleaving Proteases: Controlled Liberation of Proteins and Bioactive Peptides. *Trends Cell Biol.* **2003**, *13* (2), 71–78.

(13) Brown, M. S.; Ye, J.; Rawson, R. B.; Goldstein, J. L. Regulated Intramembrane Proteolysis: a Control Mechanism Conserved from Bacteria to Humans. *Cell* **2000**, *100* (4), 391-398.

(14) Zhang, Z.; Song, M.; Liu, X.; Kang, S. S.; Kwon, I.-S.; Duong, D. M.; Seyfried, N. T.; Hu, W. T.; Liu, Z.; Wang, J.-Z. et al. Cleavage of Tau by Asparagine Endopeptidase Mediates the Neurofibrillary Pathology in Alzheimer's Disease. *Nat. Med.* **2014**, *20* (11), 1254–1262.

(15) Gamblin, T. C.; Chen, F.; Zambrano, A.; Abraha, A.; Lagalwar, S.; Guillozet, A. L.; Lu, M.; Fu, Y.; Garcia-Sierra, F.; LaPointe, N. et al. Caspase Cleavage of Tau: Linking Amyloid and Neurofibrillary Tangles in Alzheimer's Disease. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (17), 10032–10037.

(16) Ilitchev, A. I.; Giammona, M. J.; Olivas, C.; Claud, S. L.; Cantrell, K. L. L.; Wu, C.; Buratto, S. K.; Bowers, M. T. Hetero-Oligomeric Amyloid Assembly and Mechanism: Prion Fragment PrP(106−126) Catalyzes the Islet Amyloid Polypeptide β-Hairpin. *J. Am. Chem. Soc.* **2018**, *140* (30), 9685-9695.

(17) Stöhr, J.; Wu, H.; Nick, M.; Wu, Y.; Bhate, M.; Condello, C.; Johnson, N.; Rodgers, J.; Lemmin, T.; Acharya, S. et al. A 31-Residue Peptide Induces Aggregation of Tau's Microtubule-Binding Region in Cells. *Nat. Chem.* **2017**, *9* (9), 874-881.

(18) Zheng, X.; Wu, C.; Liu, D.; Li, H.; Bitan, G.; Shea, J.-E.; Bowers, M. T. Mechanism of C-Terminal Fragments of Amyloid β-Protein as Aβ Inhibitors: Do C-Terminal Interactions Play a Key Role in Their Inhibitory Activity? *J. Phys. Chem. B* **2016**, *120* (8), 1615–1623.

(19) Andreasen, M.; Skeby, K. K.; Zhang, S.; Nielsen, E. H.; Klausen, L. H.; Frahm, H.; Christiansen, G.; Skrydstrup, T.; Dong, M.; Schiøtt, B. et al. The Importance of Being Capped: Terminal Capping of an Amyloidogenic Peptide Affects Fibrillation Propensity and Fibril Morphology. *Biochemistry* **2014**, *53* (44), 6968–6980.

(20) Sabaté, R.; Espargaró, A.; de Groot, N. S.; Valle-Delgado, J. J.; Fernàndez-Busquets, X.; Ventura, S. The Role of Protein Sequence and Amino Acid Composition in Amyloid Formation: Scrambling and Backward Reading of IAPP Amyloid Fibrils. *J. Mol. Biol.* **2010**, *404* (2), 337–352.

(21) Fishwick, C. W. G.; Beevers, A. J.; Carrick, L. M.; Whitehouse, C. D.; Aggeli, A.; Boden, N. Structures of Helical-Tapes and Twisted Ribbons: The Role of Side-Chain Interactions on Twist and Bend Behavior. *Nano Lett.* **2003**, *3* (11), 1475-1479.

(22) Rg Gsponer, J.; Haberthü, U.; Caflisch, A. The Role of Side-Chain Interactions in the Early Steps of Aggregation: Molecular Dynamics Simulations of an Amyloid-Forming Peptide from the Yeast Prion Sup35. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (9), 5154-5159.

(23) Marchut, A. J.; Hall, C. K. Side-Chain Interactions Determine Amyloid Formation by Model Polyglutamine Peptides in Molecular Dynamics Simulations. *Biophys. J.* **2006**, *90* (12), 4574-4584.

(24) Wei, G.; Shea, J.-E. Effects of Solvent on the Structure of the Alzheimer Amyloid-beta (25-35) Peptide. *Biophys. J.* **2006**, *91* (5), 1638-1647.

(25) Shen, C.-L.; Murphy, R. M. Solvent Effects on Self-Assembly of Beta-Amyloid Peptide. *Biophys. J.* **1995**, *69* (2), 640-651.

(26) Lee, M.; Choi, H.; Na, S. Effects of End-Terminal Capping on Transthyretin (105-115) Amyloid Protofibrils Using Steered Molecular Dynamics. *J. Nanomat.* **2016**, 1-10.

(27) Tao, K.; Wang, J.; Zhou, P.; Wang, C.; Xu, H.; Zhao, X.; Lu, J. R. Self-Assembly of Short Aβ(16−22) Peptides: Effect of Terminal Capping and the Role of Electrostatic Interaction. *Langmuir* **2011**, *27* (6), 2723–2730.

(28) Goedert, M. Tau Filaments in Neurodegenerative Diseases. *FEBS Lett*. **2018,** *592* (14), 2383-2391.

(29) Guo, T.; Noble, W.; Hanger, D. P. Roles of Tau Protein in Health and Disease. *Acta Neuropathologica*. **2017**, *133* (5), 665-704.

(30) Goedert, M.; Spillantini, M. G. Propagation of Tau Aggregates. *Molecular Brain*. **2017**, *10* (1), 1-9.

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

(32) Jouanne, M.; Rault, S.; Voisin-Chiret, A. S. Tau Protein Aggregation in Alzheimer's Disease: An Attractive Target for the Development of Novel Therapeutic Agents. *Eur. J. Med. Chem*. **2017**, *139*, 153-167.

(33) Wang, Y.; Mandelkow, E. Tau in Physiology and Pathology. *Nat. Rev. Neurosci*. **2016**, *17* (1), 5-21.

(34) Jeganathan, S.; Von Bergen, M.; Brutlach, H.; Steinhoff, H.-J.; Mandelkow, E. Global Hairpin Folding of Tau in Solution. *Biochemistry* **2006**, *45* (7), 2283-2293.

(35) Cleveland, D. W.; Hwo, S.-Y.; Kirschner, M. W. Physical and Chemical Properties of Purified Tau Factor and the Role of Tau in Microtubule Assembly. *J. Mol. Biol.* **1977**, *116* (2), 227-247.

(36) 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* (4), 519-526.

(37) Buée, L.; Delacourte, A. Comparative Biochemistry of Tau in Progressive Supranuclear Palsy, Corticobasal Degeneration, FTDP-17 and Pick's Disease. *Brain Pathol.* **1999**, *9* (4), 681–693.

(38) Mandelkow, E. M.; Mandelkow, E. Biochemistry and Cell Biology of Tau Protein in Neurofibrillary Degeneration. *Cold Spring Harb. Perspect. Med*. **2012**, *2* (7), 1–25.

(39) Goedert, M.; Jakes, R.; Spillantini, M. G.; Hasegawa, M.; Smith, M. J.; Crowther, R. A. Assembly of Microtubule-Associated Protein Tau into Alzheimer-like Filaments Induced by Sulphated Glycosaminoglycans. *Nature* **1996**, *383* (6600), 550-553.

(40) Kampers, T.; Friedhoff, P.; Biernat, J.; Mandelkow, E. M.; Mandelkow, E. RNA Stimulates Aggregation of Microtubule-Associated Protein Tau into Alzheimer-like Paired Helical Filaments. *FEBS Lett.* **1996**, *399* (3), 344-349.

(41) Peterson, D. W.; Zhou, H.; Dahlquist, F. W.; Lew, J. A Soluble Oligomer of Tau Associated

with Fiber Formation Analyzed by NMR. *Biochemistry* **2008**, *47* (28), 7393-7404.

(42) Li, W.; Lee, V. M. Y. Characterization of Two VQIXXK Motifs for Tau Fibrillization in Vitro. *Biochemistry* **2006**, *45* (51), 15692-15701.

(43) 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. W. Cryo-EM Structures of Tau Filaments from Alzheimer's Disease. *Nature* **2017**, *547* (7662), 185–190.

(44) 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. U. S. A.* **2000**, *97* (10), 5129-5134.

(45) Goux, W. J.; Kopplin, L.; Nguyen, A. D.; Leak, K.; Rutkofsky, M.; Shanmuganandam, V. D.; Sharma, D.; Inouye, H.; Kirschner, D. A. The Formation of Straight and Twisted Filaments from Short Tau Peptides. *J. Biol. Chem.* **2004**, *279* (26), 26868-26875.

(46) Quinn, J. P.; Corbett, N. J.; Kellett, K. A. B.; Hooper, N. M. Tau Proteolysis in the Pathogenesis of Tauopathies: Neurotoxic Fragments and Novel Biomarkers. *J. Alzheimer's Dis.* **2018**, *63* (1), 13–33.

(47) Von Bergen, M.; Barghorn, S.; Biernat, J.; Mandelkow, E. M.; Mandelkow, E. Tau Aggregation Is Driven by a Transition from Random Coil to Beta Sheet Structure. *Biochim. Biophys. Acta.* **2005**, *1739* (2-3), 158-166.

(48) Tomoo, K.; Yao, T. M.; Minoura, K.; Hiraoka, S.; Sumida, M.; Taniguchi, T.; Ishida, T. Possible Role of Each Repeat Structure of the Microtubule-Binding Domain of the Tau Protein in in Vitro Aggregation. *J. Biochem.* **2005**, *138* (4), 413-423.

(49) Frenkel-Pinter, M.; Richman, M.; Belostozky, A.; Abu-Mokh, A.; Gazit, E.; Rahimipour, S.; Segal, D. Selective Inhibition of Aggregation and Toxicity of a Tau-Derived Peptide Using Its Glycosylated Analogues. *Chemistry* **2016**, *22* (17), 5945-5952.

(50) Jiji, A. C.; Shine, A.; Vijayan, V. Direct Observation of Aggregation-Induced Backbone Conformational Changes in Tau Peptides. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (38), 11562-11566.

(51) Rojas Quijano, F. A.; Morrow, D.; Wise, B. M.; Brancia, F. L.; Goux, W. J. Prediction of Nucleating Sequences from Amyloidogenic Propensities of Tau-Related Peptides. *Biochemistry* **2006**, *45* (14), 4638-4652.

(52) Santa-María, I.; Pérez, M.; Hernández, F.; Muñoz, V.; Moreno, F. J.; Avila, J. In Vitro Tau Fibrillization: Mapping Protein Regions. *Biochim. Biophys. Acta - Mol. Basis Dis.* **2006**, *1762* (7), 683-692.

(53) Li, W.; Sperry, J. B.; Crowe, A.; Trojanowski, J. Q.; Smith, A. B.; Lee, V. M. Y. Inhibition of Tau Fibrillization by Oleocanthal via Reaction with the Amino Groups of Tau. *J. Neurochem.* **2009**, *110* (4), 1339-1351.

(54) Feinstein, H.E.; Benbow, S.J.; LaPointe, N.E.; Patel, N.; Ramachandran, S.: Do, T. D.; Gaylord, M.R.; Huskey, N.E.; Dressler, N.; Korff, M. et al. Oligomerization of the Microtubule Associated Protein Tau is Mediated by its N-Terminal Sequences: Implications for Normal and Pathological Tau Action. *J. Neurochem.* **2016**, *137* (6), 939-954.

(55) Ganguly, P; Do, T. D.; Larini, L.; LaPointe, N. E.; Sercel, A. J.; Shade, M.F.; Feinstein, S. C.; Bowers, M. T.; Shea, J.-E. Tau Assembly: The Dominant Role of PHF6 (VQIVYK) in Microtubule Binding Region Repeat R3. *J. Phys. Chem. B* **2015**, *119* (13), 4582-4593.

(56) Do, T. D.; Economou, N. J.; Chamas, A.; Buratto, S. K.; Shea, J.-E.; Bowers, M. T. Interactions between Amyloid-β and Tau Fragments Promote Aberrant Aggregates: Implications for Amyloid Toxicity. *J. Phys. Chem. B* **2014**, *118* (38), 11220-11230.

(57) Larini, L.; Gessel, M. M.; LaPointe, N. E.; Do, T. D.; Bowers, M. T.; Feinstein, S. C.; Shea, J.-E. Initiation of Assembly of Tau(273-284) and its ΔK280 Mutant: An Experimental and Computational Study. *Phys. Chem. Chem. Phys.* **2013**, *15* (23), 8916-8928.

(58) Hess, B.; Kutzner, C.; Spoel, D. van der; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435-447.

(59) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225-11236.

(60) Rizzo, R. C.; Jorgensen, W. L. OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999**, *121* (20), 4827-4836.

(61) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105* (28), 6474-6487.

(62) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison

of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(63) Weerasinghe, S.; Smith, P. E. A Kirkwood–Buff Derived Force Field for Sodium Chloride in Water. *J. Chem. Phys.* **2003**, *119* (21), 11342–11349.

(64) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52* (2), 255–268.

(65) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697.

(66) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(67) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.

(68) Verlet, L. Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159* (1), 98–103.

(69) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.

(70) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(71) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13* (8), 952–962.

(72) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577-2637.

(73) Isella, L.; Drossinos, Y. Langevin Agglomeration of Nanoparticles Interacting via a Central Potential. *Phys. Rev. E* **2010**, *82*, 011404.

(74) Feder, J. The Cluster Fractal Dimension. In: Fractals. Physics of Solids and Liquids. *Springer: Boston,* **1988***;* 31-40.

(75) Ganguly, P.; Schravendijk, P.; Hess, B.; van der Vegt, N. F. A. Ion Pairing in Aqueous Electrolyte Solutions with Biologically Relevant Anions. *J. Phys. Chem. B* **2011**, *115* (13), 3734–3739.

(76) Ganguly, P.; van der Vegt, N. F. A. Convergence of Sampling Kirkwood–Buff Integrals of

Aqueous Solutions with Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (3), 1347–1355.

(77) Levine, H. Thioflavine T Interaction with Synthetic Alzheimer's Disease β-Amyloid Peptides: Detection of Amyloid Aggregation in Solution. *Protein Sci.* **2008**, *2* (3), 404–410.

(78) Goux, W. J. The Conformations of Filamentous and Soluble Tau Associated with Alzheimer Paired Helical Filaments. *Biochemistry* **2002**, *41* (46), 13798-13806.

(79) Battisti, A.; Ciasca, G.; Grottesi, A; Bianconi, A.; Tenenbaum, A. Temporary Secondary Structures in Tau, an Intrinsically Disordered Protein. *Mol. Simul.* **2012**, *38* (7), 525-533.

(80) Luo, Y.; Ma, B.; Nussinov, R.; Wei, G. Structural Insight into Tau Protein's Paradox of Intrinsically Disordered Behavior, Self-Acetylation Activity, and Aggregation. *J. Phys. Chem. Lett.* **2014**, *5* (17), 3026-3031.

(81) Siddiqua, A.; Luo, Y.; Meyer, V.; Swanson, M. A.; Yu, X.; Wei, G.; Zheng, J.; Eaton, G. R.; Ma, B.; Nussinov, R. et al. Conformational Basis for Asymmetric Seeding Barrier in Filaments of Three- and Four-Repeat Tau. *J. Am. Chem. Soc.* **2012**, *134* (24), 10271-10278.

(82) Yu, X.; Luo, Y.; Dinkel, P.; Zheng, J.; Wei, G.; Margittai, M.; Nussinov, R.; Ma, B. Cross-seeding and conformational Selection between Three- and Four-repeat Human Tau Proteins. *J. Biol. Chem.* **2012**, *287* (18), 14950-14959.

(83) Meyer, V.; Dinkel, P. D.; Luo, Y.; Yu, X.; Wei, G.; Zheng, J.; Eaton, G. R.; Ma, B.; Nussinov, R.; Eaton, S. S. et al. Single Mutations in Tau Modulate the Populations of Fibril Conformers through Seed Selection. *Angew. Chem., Int. Ed.* **2014**, *53* (6), 1590-1593.

(84) Li, X.; Dong, X.; Wei, G.; Margittai, M.; Nussinov, R.; Ma, B. The Distinct Structural Preferences of Tau Protein Repeat Domains. *Chem. Commun.* **2018**, *54* (45), 5700-5703.

(85) Matthes, D.; Gapsys, V.; Daebel, V.; de Groot, B. L. Mapping the Conformational Dynamics and Pathways of Spontaneous Steric Zipper Peptide *Oligomerization*. *PLoS ONE* **2011**, *6* (5), e19129.

(86) Cheon, M.; Chang, I.; Hall, C. K. Influence of Temperature on Formation of Perfect Tau Fragment Fibrils Using PRIME20/DMD Simulations. *Protein Sci.* **2012**, *21* (10),1514-1527.

(87) Li, D.-W.; Mohanty, S.; Irbä ck, A.; Huo, S. Formation and Growth of Oligomers: A Monte Carlo Study of an Amyloid Tau Fragment. *PLoS Comput. Biol.* **2008**, *4* (12), e1000238.

(88) Levine, Z. A.; Larini, L.; LaPointe, N. E.; Feinstein, S. C.; Shea, J.-E. Regulation and

Aggregation of Intrinsically Disordered Peptides. Proc. *Natl. Acad. Sci. U. S. A.* **2015**, *112* (9), 2758-2763.

(89) Derreumaux, P.; Man, V. H.; Wang, J.; Nguyen, P. H. Tau R3-R4 Domain Dimer of the Wild Type and Phosphorylated Ser356 Sequences. I. In Solution by Atomistic Simulations. *J. Phys. Chem. B* **2020**, *124* (15), 2975-2983.

(90) Carballo-Pacheco, M.; Strodel, B. GROMACS 4: Advances in the Simulation of Protein Aggregation at the Atomistic Scale. *J. Phys. Chem. B* **2016**, *120* (12), 2991-2999.

(91) Kolb, M.; Botet, R.; Jullien, R. Scaling of Kinetically Growing Clusters. *Phys. Rev. Lett.* **1983**, *51*, 1123-1126.

(92) Arakhamia, T.; Lee, C. E.; Carlomagno, Y.; Duong, D. M.; Kundinger, S. R.; Wang, K.; Williams, D.; DeTure, M.; Dickson, D. W.; Cook, C. N. et al. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell* **2020,** *180* (4) , 1-12.

Figure Captions

Figure 1. ThT fluorescence for all capping variants of Tau PHF6 fragment (A, B) without heparin, and (C, D) with heparin. The data shown in A and C were acquired for 24 hours in kinetic mode and the data shown in B and D were acquired after 2 weeks of incubation at 37°C. The fully capped PHF6 fragment (Ac-PHF6-NH_2) is shown in olive. The fragment acetylated at the N-terminus (Ac-PHF6) and the fragment amidated at the C-terminus (PHF6- NH_2) are shown in red and blue, respectively. The peptide with both termini free is shown in black.

Figure 2. Circular Dichroism (CD) spectra for all capping variants (A) without heparin, and (B) with heparin. The data were collected after 2 weeks of incubation at 37°C. The fully capped PHF6 fragment (Ac-PHF6-NH_2) is shown in olive. The fragment acetylated at the N-terminus (Ac-PHF6) and the fragment amidated at the C-terminus (PHF6- NH_2) are shown in red and blue, respectively. The peptide with both termini free is shown in black.

Figure 3. Transmission electron microscopy (TEM) images of aggregates formed by all capping variants of Tau PHF6 fragment (A-D) without heparin, and (E-H) with heparin. All images acquired after 2 weeks of incubation at 37°C. The straight filament like morphologies are highlighted with cyan arrows and the twists within typical paired helical filaments (PHFs) are highlighted with pink arrows. The scale bar is 50 nm for all of the images corresponding to a 137,200x magnification.

Figure 4. Plots of log radius of gyration (R_g) of the peptide clusters with the log of thesize of the clusters (N) in terms of the number of the peptides in the respective clusters. Dashed lines correspond to a linear fit to the data. The“m” value represents the corresponding exponent of the power law relation between R_g and N (see text for details). For the Ac-PHF6 peptide, the data for N=3 has been omitted in the linear fit. The maximum error in $R_g < 0.01$nm.

Figure 5 Representative snapshots of the most stable peptide clusters of the respective cluster sizes for the Ac-PHF6-NH2 peptide are shown (A-G). The numbers in the circles for each panel represent the number of the peptides in the respective aggregates. The backbones of the peptides

are shown in the “ribbon” representation and the side chain atoms are shown in the “licorice” representation. The side chains are not shown in the middle and the lower panel of E) and in the lower panel of G), allow for clear visualization of the backbone structures.

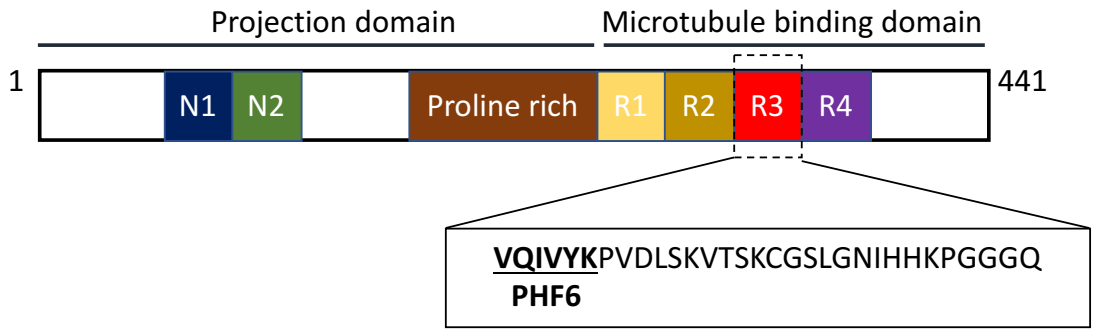
Figure 6. The probability distributions of the inter-residue hydrogen bonds between the various amino acid residues are shown in A) Ac-PHF6-NH2, B) PHF6, C) Ac-PHF6 and D) PHF6-NH2. The relative intensity of each interaction is given by reference to the color bar on the right hand side of each figure.

Figure 7. ThT fluorescence (A) and TEM images (B,C) for R3 capping variants without heparin, and (D-F) with heparin are shown. The data shown in A and D were acquired for 24 hours in kinetic mode. The inset in panel A shows ThT fluorescence data acquired for uncapped and capped variants after 2 weeks. TEM images were acquired after 2 weeks of incubation at 37°C. The scale bar is 50 nm for all of the images corresponding to a 137,200x magnification. The fully capped R3 fragment (Ac-R3-NH_2) is shown in red and the fragment with both termini free (R3) is shown in black.

Figures

Scheme 1.

(A) Schematic of the longest isoform of human Tau. The R3 repeat sequence shown, along with the PHF6 fragment, is present in isoforms of Tau and is considered important for Tau filament assembly.



(B) The capping variants of the R3 and PHF6 fragment used in this study:

R3: $^{+}H_3N$-VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ-COO^{-}

Ac-R3-NH_2: CH_3CONH-VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ-$CONH_2$

PHF6: $^{+}H_3N$-VQIVYK-COO^{-}

Ac-PHF6: CH_3CONH-VQIVYK-COO^{-}

PHF6-NH_2: $^{+}H_3N$-VQIVYK-$CONH_2$

Ac-PHF6-NH_2: CH_3CONH-VQIVYK-$CONH_2$

Table 1. A comparative overview of our experimental results obtained for four capping variants of PHF6. The (-) and (+) signs in the heparin (-/+) column indicate with or without heparin added, while in the rest of the table they are used to compare the extent of fibrillation (ThT) / fibril formation (TEM) / secondary structure changes (CD) during fibrillation, where (-) stands for no change, (+) stands for an appreciable change and (++) indicates a large extent of change.

Technique	Heparin (-/+)	Capping Variants			
		PHF6	Ac-PHF6	PHF6-NH_2	Ac-PHF6-NH_2
ThT Fluorescence	-	-	+	-	++
	+	-	+	-	++
TEM	-	-	+	-	++
	+	-	++	+	++
CD	-	-	-	-	++
	+	-	+	+	++

Figure 1.

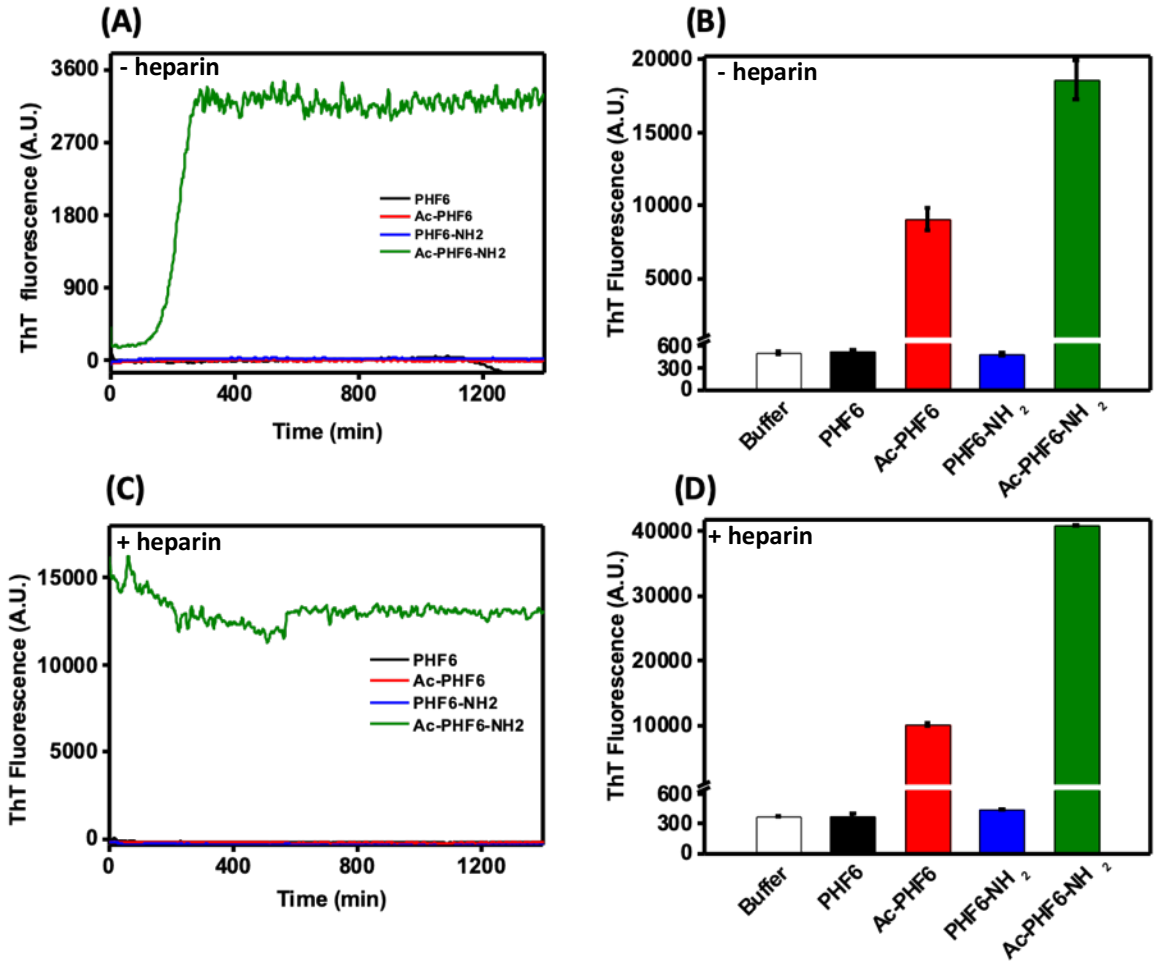


Figure 2.

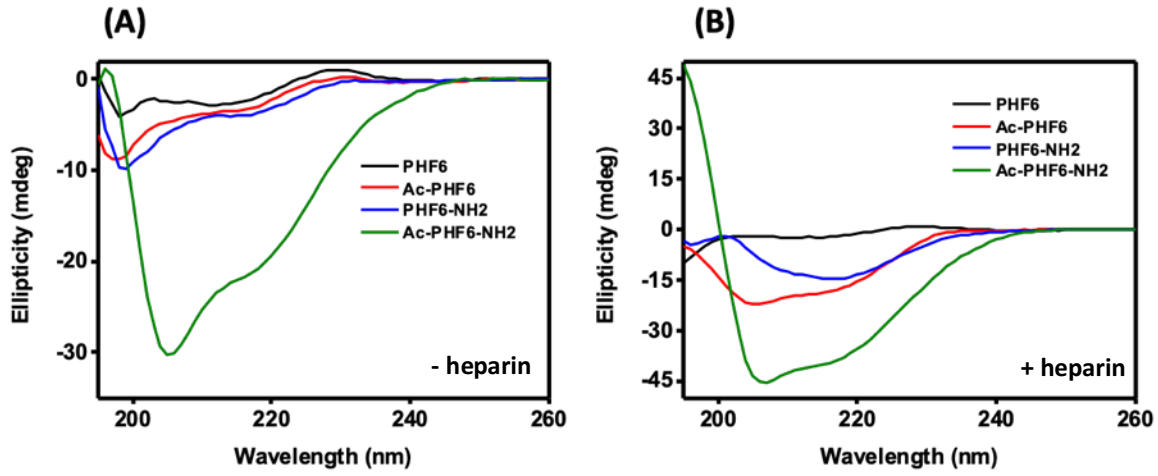


Figure 3.

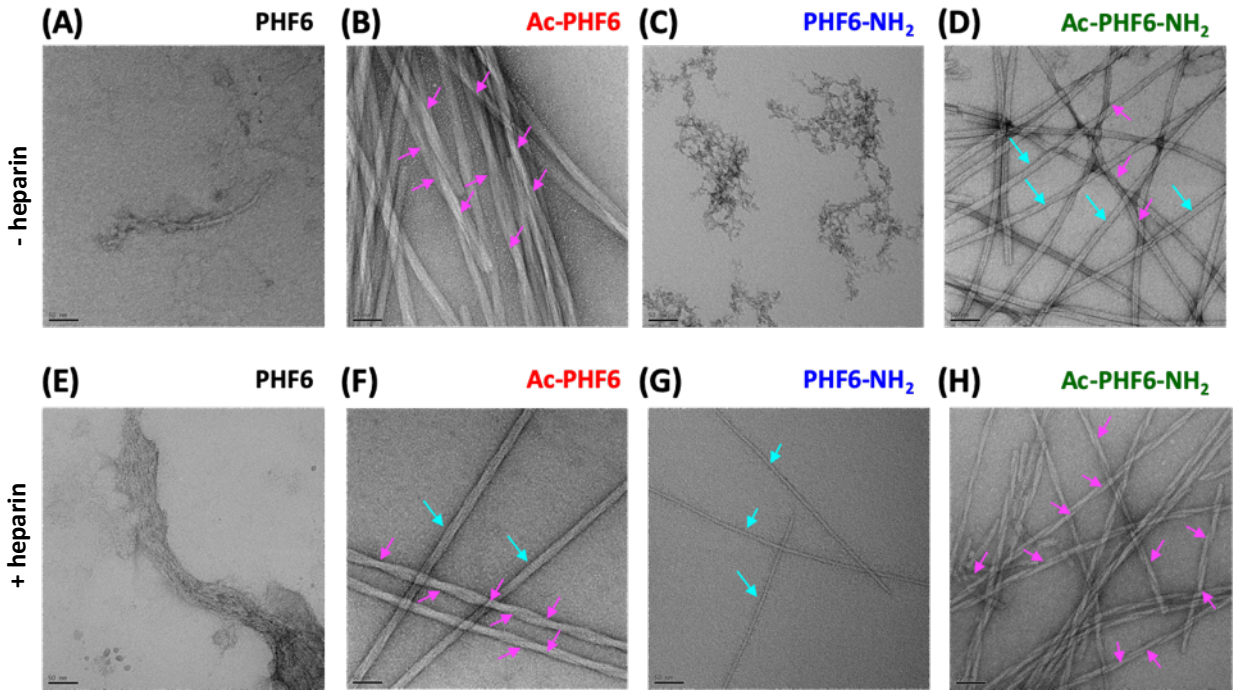

(A) PHF6
(B) Ac-PHF6
(C) PHF6-NH2
(D) Ac-PHF6-NH2
- heparin
(E) PHF6
(F) Ac-PHF6
(G) PHF6-NH2
(H) Ac-PHF6-NH2
+ heparin

Figure 4.

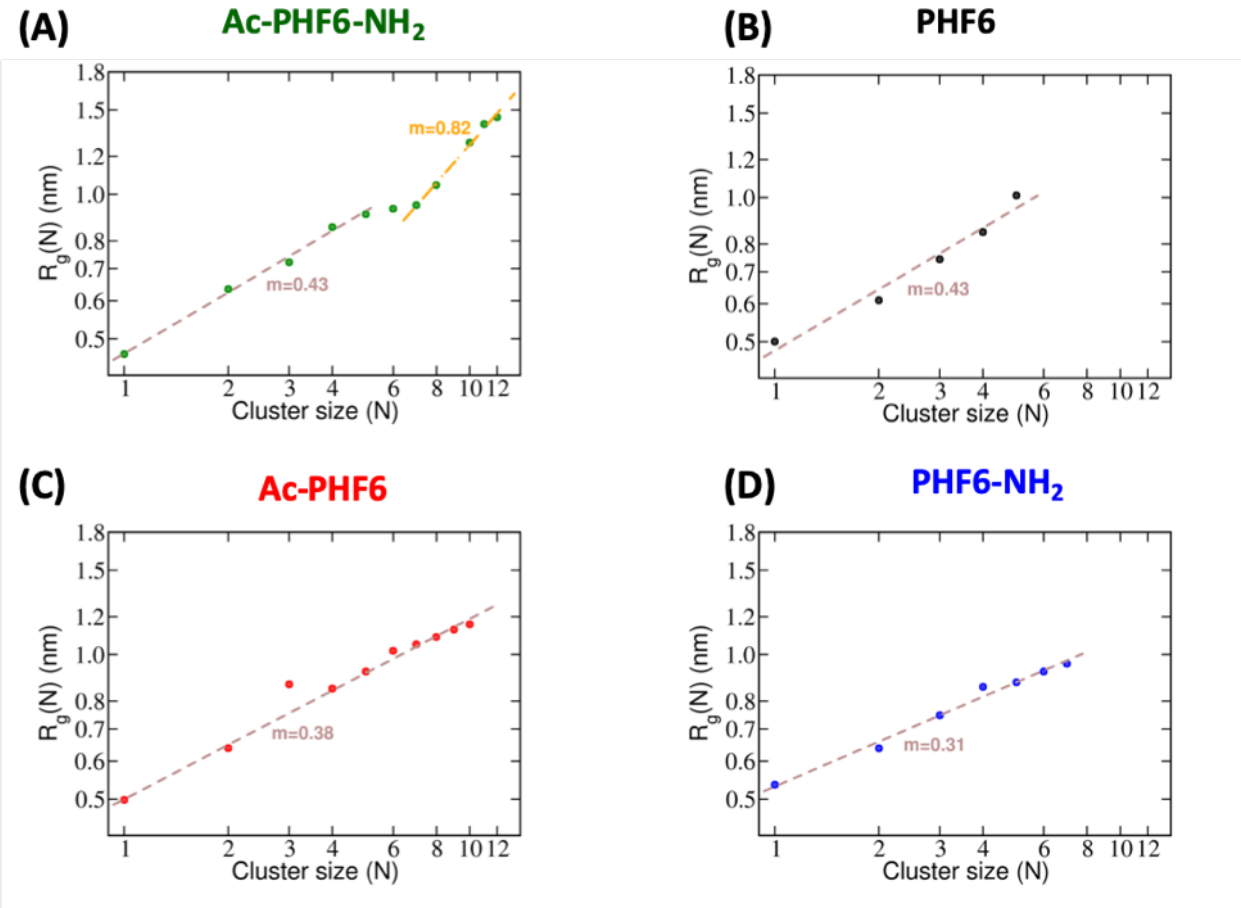

(A)
Ac-PHF6-NH2
m=0.82
m=0.43
(B)
PHF6
m=0.43
(C)
Ac-PHF6
m=0.38
(D)
PHF6-NH2
m=0.31
Rg(N) (nm)
Cluster size (N)

Figure 5.

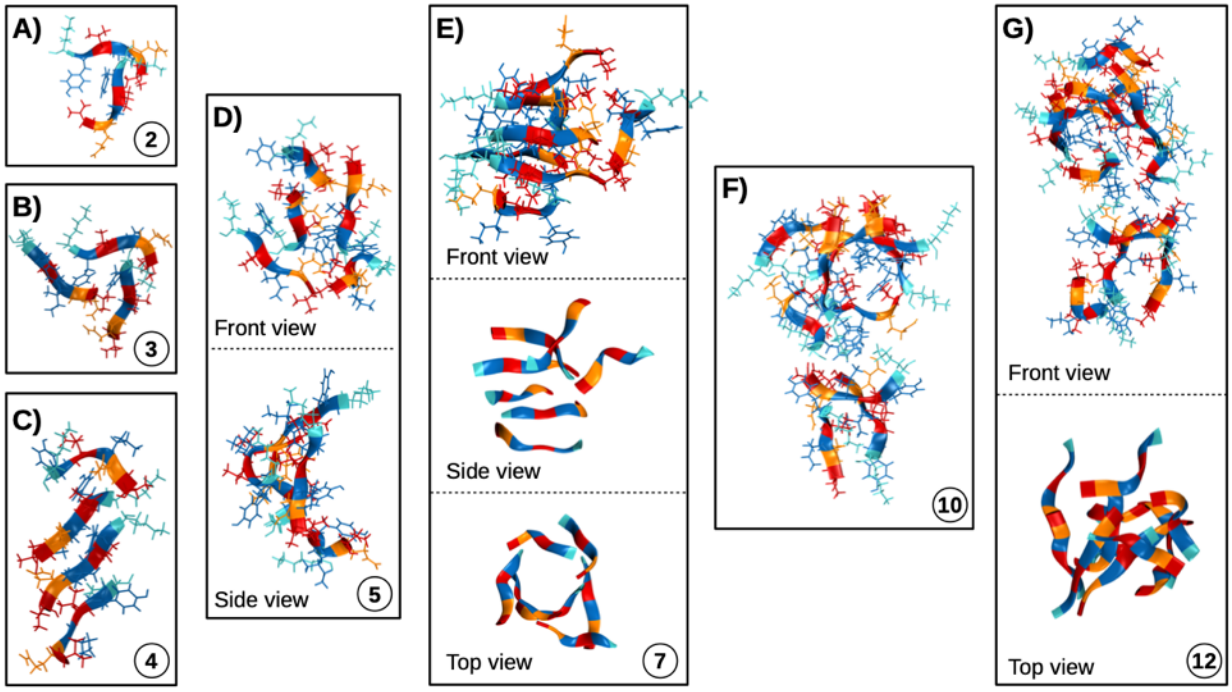

A)
2
B)
3
C)
4
D)
Front view
Side view
5
E)
Front view
Side view
Top view
7
F)
10
G)
Front view
Top view
12

Figure 6.

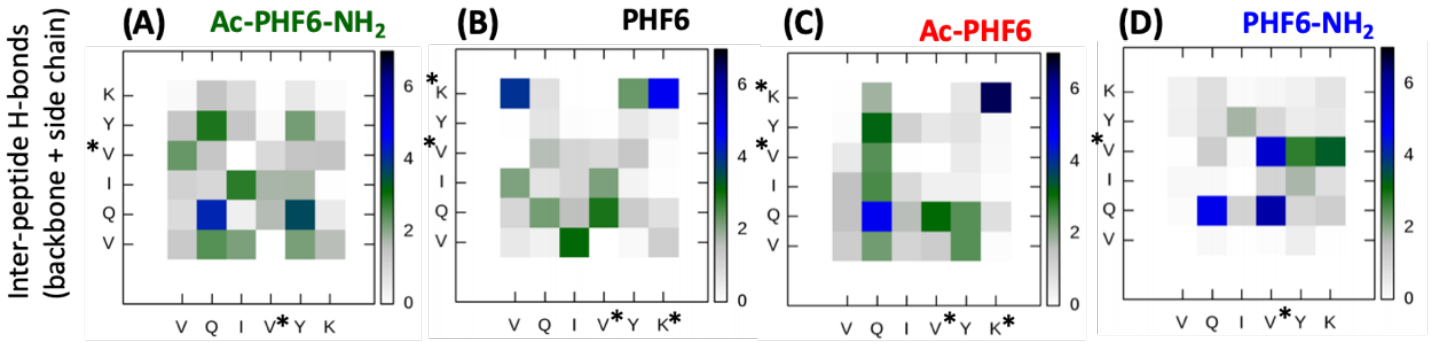
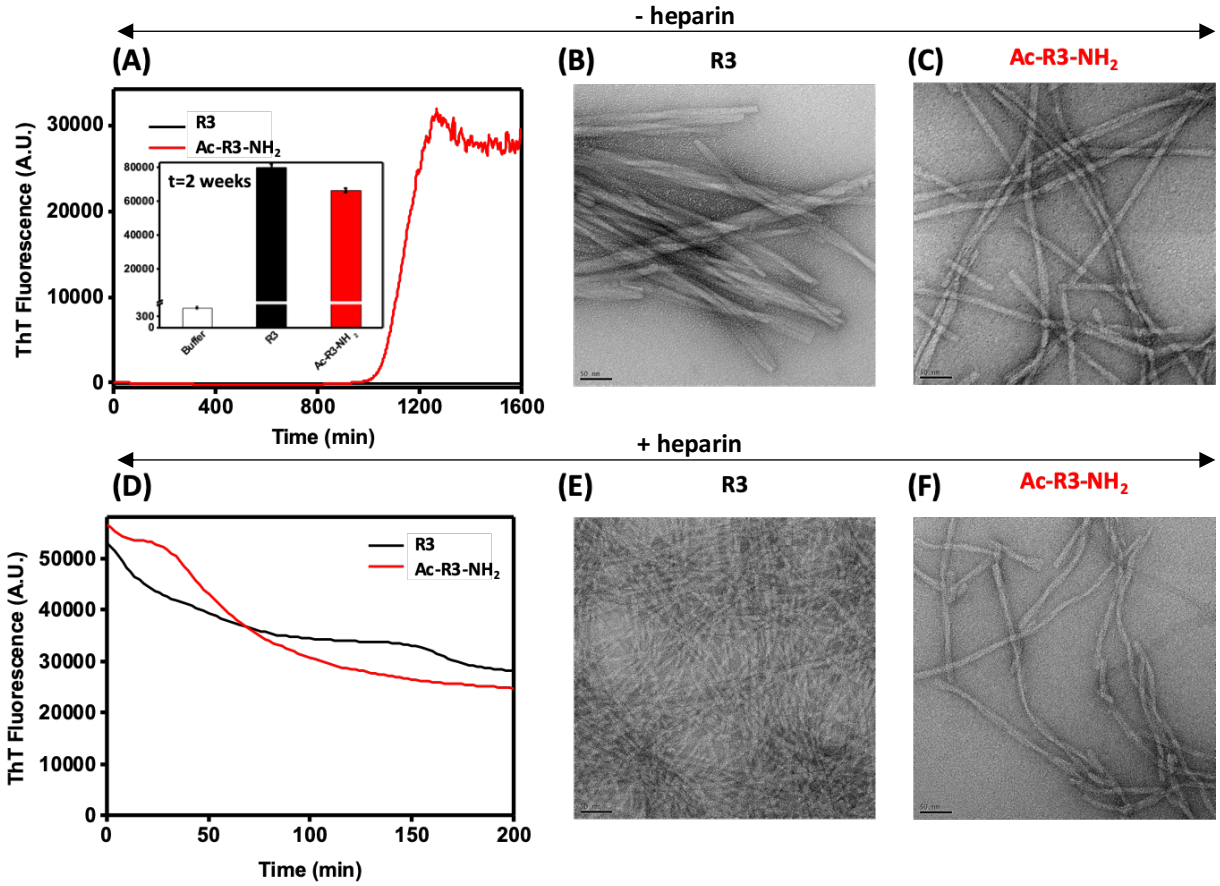

Inter-peptide H-bonds
(backbone + side chain)
(A) Ac-PHF6-NH2
(B) PHF6
(C) Ac-PHF6
(D) PHF6-NH2
K Y V I Q V
V Q I V* Y K
0 2 4 6

Figure 7.


- heparin
(A)
(B)
R3
(C)
Ac-R3-NH2
ThT Fluorescence (A.U.)
30000
20000
10000
0
R3
Ac-R3-NH2
t=2 weeks
80000
60000
40000
20000
300
0
Buffer
R3
Ac-R3-NH2
0
400
800
1200
1600
Time (min)
+ heparin
(D)
(E)
R3
(F)
Ac-R3-NH2
ThT Fluorescence (A.U.)
50000
40000
30000
20000
10000
0
R3
Ac-R3-NH2
0
50
100
150
200
Time (min)