# Supplementary Information

## Structural Insight into Tau Protein's Paradox of Intrinsically Disordered Behavior, Self-acetylation Activity, and Aggregation

Yin Luo,[†] Buyong Ma,*[,‡] Ruth Nussinov, [‡,§] and Guanghong Wei*[,†]

[†] State Key Laboratory of Surface Physics, Key Laboratory for Computational Physical Sciences (MOE), and Department of Physics, Fudan University, Shanghai, P.R.China

[‡] Basic Science Program, Leidos Biomedical Research, Inc. Cancer and Inflammation Program, National Cancer Institute, Frederick, Maryland 21702, USA

[§] Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

**This supplementary material contains the details of initial state modeling, REMD simulations, analysis methods, some details of results, three tables, and 10 supplementary figures.**

### Initial state of K18 and K19 monomers

For each REMD simulation, 24 different starting states were used: five of them were generated using I-TASSER[1-3]--an online server based the iterative threading assembly refinement algorithm; ten of them were generated using the online server QUARK[4]--a template-free protein structure prediction program using continuous structure fragments and optimized knowledge-based fore field; five of them are generated using Robetta[5]--an online server based on comparative modeling or de novo Rosetta fragment insertion method; one of them was generated using the SAM-T06[6-7]--an online protein structure prediction server based on hidden Markov models (HMM-based); while the other three are modeled with U-turn structures[8]. Previous MD-simulation and double electron−electron resonance spectroscopy studies suggested that the amyloid filaments of K18 and K19 are ensemble of U-turn motifs[8-9]. To examine whether such

U-turn motifs are populated in the monomeric forms of K18 and K19, we selected three conformations with U-turn motifs as the starting states of REMD simulations. Figures S1 shows the 24 initial states of REMD runs for (A) K18 and (B) K19.

**REMD simulations**

We explored the conformational space of tau K18 and K19 monomers using replica exchange molecule dynamics (REMD) simulations[10-12] with explicit water. Explicit-solvent REMD simulation has been shown to be able to reveal the conformational organization of small disordered histone tails[13]. Two REMD simulations were performed using the GROMACS-4.5.3 software package[14] with all-atom CHARMM27 force field[15]. The simulations were conducted in the NPT ensemble using 48 replicas. The exchange time between two adjacent replicas is 1 ps and the average acceptance ratio is 16%. In the REMD simulations, the temperatures were maintained at the chosen values using the velocity rescaling method[16] with a coupling constant of 0.1 ps. The pressure was kept at 1.0 atm with a coupling constant of 1 ps via the Parrinello-Rahman barostat[17]. Bond lengths within peptides and water molecules were respectively constrained by the LINCS[18] and the SETTLE[19] algorithms. This allowed an integration time step of 2 fs. A twin-range cutoff of 0.9/1.4 nm was applied for van der Waals interactions and the particle mesh Ewald method was used to calculate the electrostatic interactions with a real space cutoff of 0.9 nm. All simulations were performed using periodic boundary conditions. The protein molecules were explicitly solvated in a TIP3P water box with a minimum distance of 1.0-1.8 nm from any edge of the box to any protein atom. The Na+ and Cl- were added to the solution with an ion concentration of 0.1 M.

For each REMD simulation, the 24 different starting conformations were evenly distributed in 48 different temperatures (replicas) ranging from 310-430 K. Each replica was run for 100 ns (a total simulation time is 4.8 μs), and the last 70 ns of trajectories at T=310 K were used for analysis, which represent 70,000 conformations (one conformation was taken every 1 ps). In the rmsd-based cluster analysis, only 14,000 conformations (one conformation was taken every 5 ps) were used as the computer memory will be exceeded if the matrix is larger than 14,000×14,000. In this study, except for the rmsd-based cluster analysis, we used 70,000 conformations for all other analyses. We checked the final structures at t=100 ns, and basically excluded the possibility that the final structures are biased by the initial conformations (see Fig. S1).

## Analysis

All results reported in this study refer to the REMD sampling collected at 310 K. The intramolecular interactions were analyzed by the probability of residue-residue contact maps. Here, a contact is defined when carbon atoms of two non-sequential residues come within 0.54 nm or any other non-hydrogen atoms of two non-sequential residues lie within 0.46 nm[20]. A salt bridge between a pair of oppositely charged residues is considered to be formed if the centroids of the side-chain charged groups in oppositely charged residues lie within 0.4 nm of each other[21].

The DSSP program was used to determine the secondary structure[22]. The SPARTA program is used to predict the $C_\alpha$ chemical shift[23]. The free energy landscape was constructed using the relation $-RT \ln H(x,y)$, where $H(x,y)$ was the histogram of two selected reaction coordinates, x and y. In the analysis of spatial distribution of helical and β-sheet structures, the x coordinate is the distance of the helical structure from the protein's centroid and the y coordinate is the distance of the β-sheet structure from the protein's centroid. In the analysis of hydrophobic and hydrophilic solvent accessible surface area (SASA) distribution, the x coordinate is the hydrophobic SASAs, and the y coordinate is the hydrophilic SASAs. RMSD-based cluster analysis was performed to investigate the atomic structures of K18 and K19. With a $C_\alpha$-RMSD cutoff of 0.4 nm, the K18 and K19 monomers were separated into 289 and 262 clusters, respectively.

## Convergence of REMD simulations

The convergence of the two REMD simulations for K18 and K19 was verified by monitoring the conformational evolution in temperature space, by comparing: the probability of secondary structure contents, the probability density function (PDF) of radius gyration (Rg) and solvent accessible surface area (SASA) of K18 and K19, and the alpha-helix, β-strand, turn and coil percentages of both K18 and K19 as a function of amino acid residues within two different time intervals (30-65 and 65-100 ns), and by comparing our calculated secondary chemical shifts of $C_\alpha$ atoms with those from previous NMR experiments[24]. As seen from Fig. S2, the highly structured conformers of K18 and K19 can evolve from random or partially structured conformations. Figures S3 and S4 shows that the probability of secondary structure contents, the PDF of Rg and SASA of K18 and K19, and the secondary structure percentages of each residue within the two different time intervals are quite similar. We have calculated the number of clusters within the two time windows for both K18 and K19 using a $C_\alpha$-RMSD cutoff of 0.4 nm.

The numbers of clusters are listed in Table S1 and they are close to each other. We have also calculated the $C_\alpha$-RMSD of each of the first eight clusters in Figure 2 with each of the 24 initial structures after equilibration. Table S2 lists the smallest $C_\alpha$-RMSD value of the first eight clusters with one of the 24 initial states. It can be seen form Table S2 that all of the clusters of both K18 and K19 have the smallest $C_\alpha$-RMSD of larger than 0.55 nm and 7/6 out of the 8 clusters of K18/K19 have the smallest $C_\alpha$-RMSD of larger than 1.1 nm. These data demonstrate that the simulation results are not biased by the initial states. The comparisons between our predicted and experimentally-measured $C_\alpha$ secondary shifts show a Pearson correlation coefficient of 0.614 for K18 and 0.487 for K19 (Fig. S5), indicating that our simulation results are consistent with those from previous experiments[24]. A recent REMD study of a 20-residue IDP fragment using an aggregated simulation time of 12.4-μs produced calculated $C_\alpha$ secondary chemical shifts which have a correlation coefficient of 0.73 with experimental values[25]. We note that comparison of secondary chemical shift values is more rigorous than that of direct chemical shifts. The equilibrium structures of α-synuclein multimers have been simulated with REMD with implicit solvent model, and the correlation of the calculated chemical shift with experimental value was 0.991[26]. We also achieved excellent agreement with experimental $C_\alpha$ chemical shifts with a correlation coefficient of 0.989 for K18 and 0.984 for K19 (see Fig. S5). Taken together, our data indicate converged sampling of conformational space of the large K18 and K19 proteins.

**α-helical and β-sheet structures in K18 and K19 monomers**

Conversion of α-helix into β-sheet is frequently observed in the process of amyloid formation, as in the case of native insulin[27], designed peptides[28], amyloid β-protein (Aβ)[29] and more. We found the populations of helical structures are higher than β-structures in each repeat. In K18, the probabilities to form helix and β-sheet structures are respectively 11.15% and 7.35% for R1, 16.76% and 12.59% for R2, 14.28% and 10.35% for R3, and 22.41% and 4.52% for R4. In K19, the probabilities for R1 to form helical structure and β-structure are 25.41% and 8.31%, respectively, 18.70% (helical) and 14.40% (β) for R3, where 22.37% (helical) and 11.54% (β) for R4.

In K18, the non-local β-sheets can be divided into five major types (Figure 2(B)): (i) parallel β-sheet formed by the N-terminal region of R1 and C-terminal region of R2 (B1a-B1b in C2); (ii) parallel β-sheet formed by the N-terminal region of R1 and R2 (B1a-B1b in C3 and B1a-B1b in

C5); (iii) parallel β-sheet formed by the N-terminal of R2, R3 and R4 (B1a-B1b-B1c in C7); (iv) anti-parallel β-sheet formed by the middle of R1 and R4 (B2a-B2b in C3). Here, Bna-Bnb-Bnc (n=1, 2, 3) represents a β-sheet consisting of three β-strands labelled as a, b, and c, and the two adjacent letters represent two neighboring β-strands. In K19, the non-local β-sheets have five major types (Figure 2(D)): (i) parallel β-sheet formed by the middle of R1 and the C-terminal of R3 (B2a-B2b in C7); (ii) parallel β-sheet formed by the C-terminal of R1 and the middle of R4 (B1c-B1d in C7); (iii) parallel β-sheet formed by the N-terminal of R3 and the middle of R3 (B1a-B1b in C8); (iv) anti-parallel β-sheet formed by the N-terminal of R3 and R4 (B1b-B1c in C3). Overall, K19 forms larger β-sheets than K18, for example, β-sheet with four strands B1a-B1b-B1c-B1d is seen in clusters 3 and 7, while in K18, the largest β-sheet is three-stranded β-sheet in clusters 3, 5, and 8. In addition, β-strands in K19 are usually longer than those in K18.

**Contact probability maps and the spatial distribution of the secondary structures**

By examining the global residue-residue and side-chain--side-chain (SC-SC) contact probability, we found that there are distinct differences between K18 and K19 conformations (Fig. S7(A) and (B)). In K18, repeat R1 has strong interaction with R2, and R3 has strong interaction with R4, consistent with previous studies[30]. However, in K19, R1 interacts with both R3 and R4. Common interaction patterns are also seen in K18 and K19. In both isoforms, the C-terminal of each repeat and the N-terminal of the following repeat, separated by PGGG, show a contact pattern perpendicular to the diagonal line, corresponding to a β-hairpin structure, indicating that the domain-domain interactions are not completely random. We found that β-hairpin motifs instead of U-turn motifs are populated in the K18 and K19 monomers. The spatial distributions of the secondary structures of K18 and K19 show that most structured regions are far from the protein centroid (Fig. S7 (C) and (D)), indicating non-compact structural features of the IDP protein. The global minimum of the structured regions is located around 0.2-1.0 nm from the protein centroid, suggesting that structured regions also tend to locate close to each other, but are not in the centroid of K18 /K19. Thus, the spatial distributions of the secondary structures of K18 and K19 provided further evidence for both structured and disordered natures of tau proteins.

**K18 and K19 monomers have different salt-bridge distributions**

Structural details of K18 and K19 fibrils have been revealed through combined experimental and computational approaches[8-9, 31], indicating the existence of similar but polymorphic core

structures for both K18 and K19 tau filaments. Solid state NMR also found polymorphic core structures in the K19 fibril[32]. These studies indicated that the third repeat in K19 forms the rigid core, and Cys322-Gly323 could form kinks that disrupt the β-strands[8-9, 32]. The seeding barrier between K18 tau and K19 tau indicated that there are differences in fibril structures[33] and in monomeric properties.

Interactions between different repeats in K18 and K19 were mainly stabilized by electrostatic interactions. They are likely to promote fibril growth via peptide addition along the fibril axis[8-9, 31]. Since tau is a highly charged protein, with 21 positively charged residues and 11 negatively charged residues in K18, and with 16 positively charged residues and 9 negatively charged residues in K19, it is interesting to investigate the salt-bridge distributions in the monomers. We calculated all the salt bridges formed in K18 and K19 at T=310 K, using the data from 30 ns to 100 ns with a total of 70,000 conformations. Figure S10 shows the histogram of the number of salt bridges with respect to the number of intervening residues (NI). Here, the number of intervening residues is the number of residues separating the two salt-bridging residues in the sequence. Most of the salt bridges are formed between sequentially close residues in both K18 and K19 ($NI \leq 11$). In K18 (Figure S10 (A)), there are three dominant long-range salt bridge distributions with the peak around NI=27, 41, and 66. However, in K19 (Fig. S10 (B)), there are several long-range salt bridge distributions with the peak located at NI=27, 36, 43, 50, 54, 58 and 73.

We also calculated the average number of intra-repeat and inter-repeat salt bridges (Table S3). Most of the local salt bridges ($NI \leq 11$) are formed within R3, correlated with the high propensity of helical structures. In K18, the average numbers of inter-repeat salt bridges between R1-R2 and R3-R4 are larger than other repeat pairs, consistent with the contact probability map of K18 showing two well-separated contact regions seen in Fig. S7: R1-R2 and R3-R4. In K19, the average number of inter-repeat salt bridges between R1-R4 is much larger than those in R1-R3 and R3-R4, keeping a compact conformation. In the fibrils, the inter-repeat contacts are R1-R2-R3-R4 for K18 and R1-R3-R4 for K19[8]. Thus, the addition of K18 monomer to the fibril does not require large domain reorganization, while K19 needs to break the R1-R4 contact for polymerization. Experimentally, K19 is slower to form amyloid fibrils in the absence of heparin catalysis[31]. The reorganization of salt bridges could be one of the factors to slow fibril formation in K19.

## Three tables:

Table S1. The number of clusters within the two time windows (30-65 ns and 65-100 ns) for both K18 and K19 using a $C_\alpha$-RMSD cutoff of 0.4 nm. *P* refers to the probability of the conformations in a cluster among all the conformations.

| | Time interval | All clusters | Clusters with *P*>0.1% | Clusters with *P*>1% |
|---|---|---|---|---|
| K18 | 30-65 ns | 171 | 80 | 26 |
| | 65-100 ns | 132 | 69 | 23 |
| K19 | 30-65 ns | 160 | 89 | 23 |
| | 65-100 ns | 111 | 64 | 17 |

Table S2: The smallest $C_\alpha$-RMSD value of the first eight clusters with one of the 24 initial states.

| K18 | | | K19 | | |
|---|---|---|---|---|---|
| cluster name | name of initial state | RMSD (nm) | cluster name | name of initial state | RMSD (nm) |
| C1 | Q2 | 1.29 | C1 | Q2 | 1.11 |
| C2 | Q10 | 1.31 | C2 | Q9 | 1.16 |
| C3 | R1 | 0.81 | C3 | R5 | 0.86 |
| C4 | IT3 | 1.37 | C4 | IT2 | 1.26 |
| C5 | U1 | 1.45 | C5 | IT2 | 1.27 |
| C6 | IT3 | 1.43 | C6 | Q2 | 1.04 |
| C7 | IT2 | 1.14 | C7 | IT3 | 0.57 |
| C8 | Q1 | 1.44 | C8 | R3 | 1.17 |

Table S3. The average number of intra-repeat and inter-repeat salt bridges.

| System | R1 | R2 | R3 | R4 | R1-R2 | R1-R3 | R1-R4 | R2-R3 | R2-R4 | R3-R4 |
|---|---|---|---|---|---|---|---|---|---|---|
| K18 | 0.92 | 0.65 | 1.32 | 3.47 | 0.43 | 0.26 | 0.37 | 0.13 | 0.28 | 0.61 |
| K19 | 0.97 | - | 0.58 | 3.60 | - | 0.21 | 1.02 | - | - | 0.42 |

## 10 Supplementary Figures

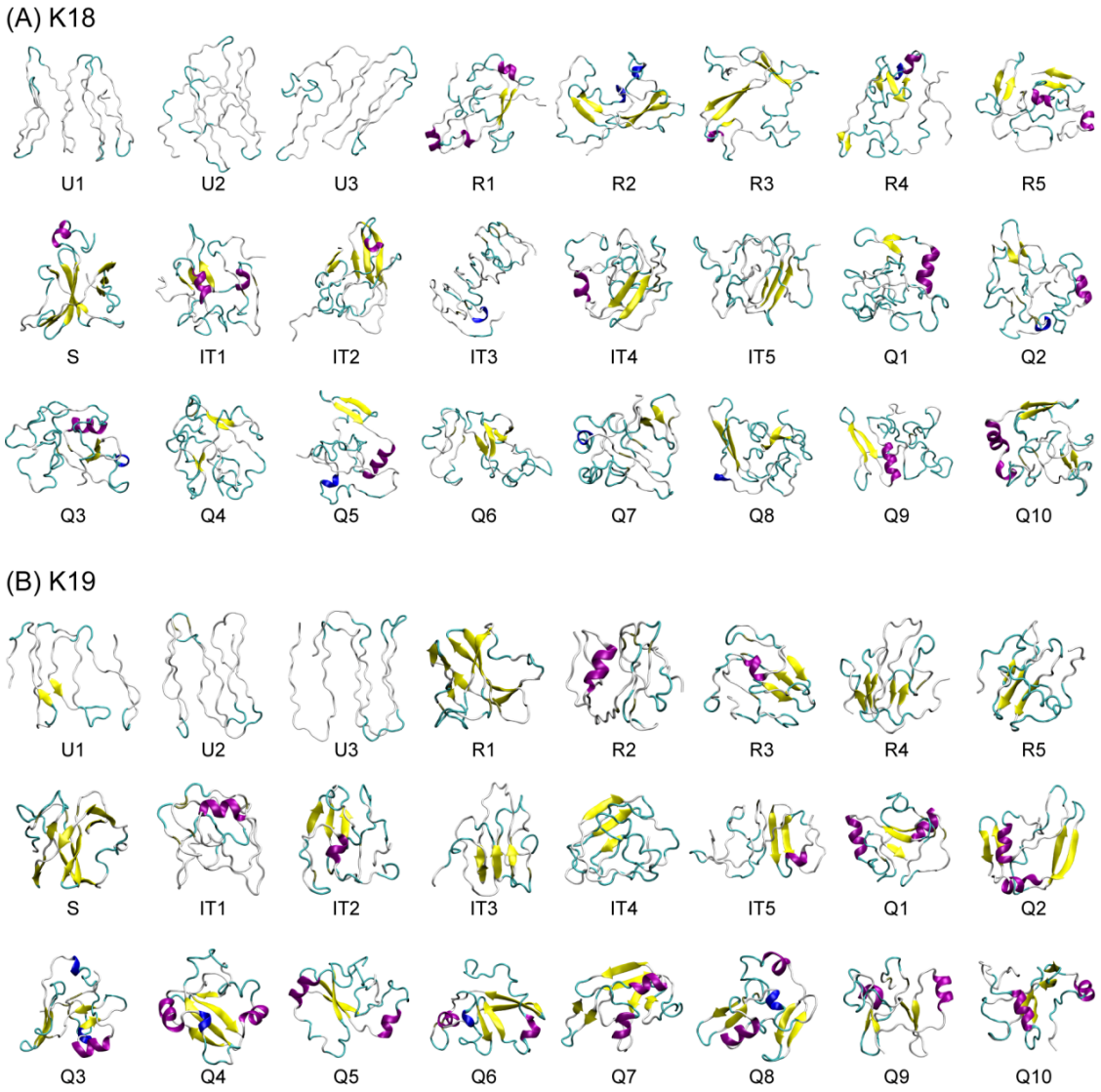


**Figure S1**. The 24 initial states of REMD runs for (A) K18 and (B) K19, with 3 conformations with U-turn structures and 21 conformations predicted by SAM-T06, ROBETTA, QUARK and I-TASSER. Here, U stands for 'U-turn-like', R stands for 'Robetta', S stands for 'SAM-T06', IT stands for 'I-TASSER', Q stands for 'QUARK'.

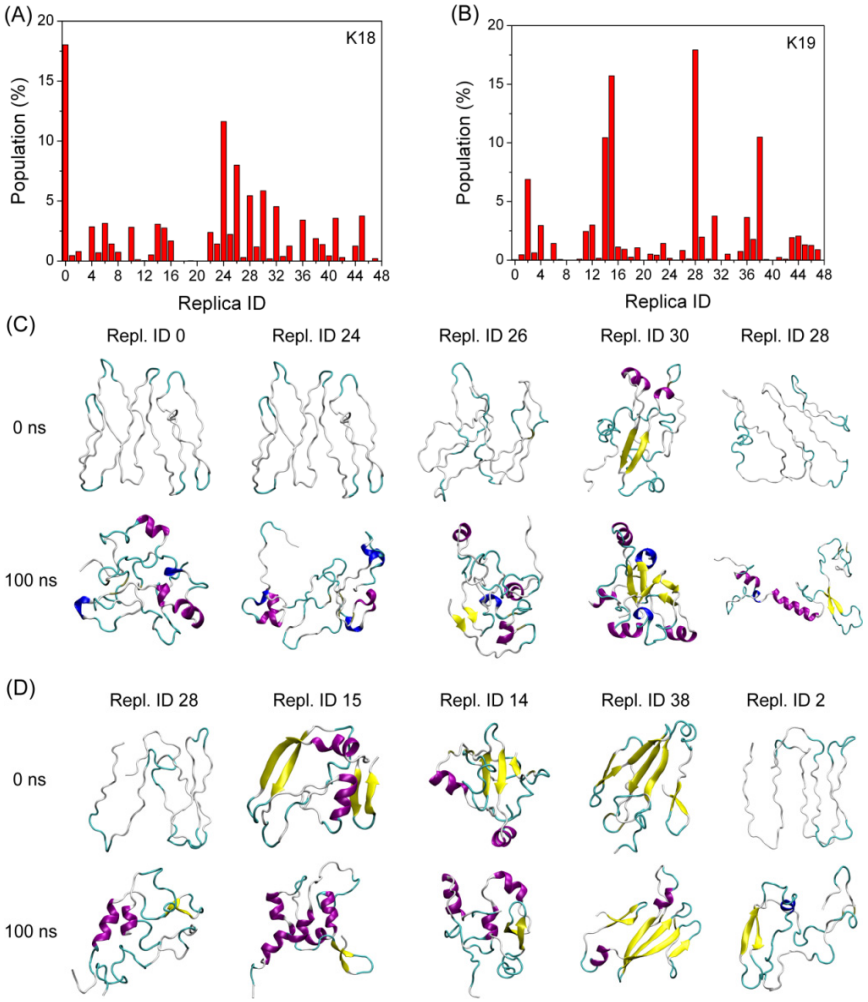


**Figure S2**. The populations of the conformations of each replicas at T=310 K for (A) K18 and (B) K19. The top five replicas at t=0 ns and t=100 ns for (C) K18 and (D) K19. During the simulations, the most structured conformers occupied the lowest temperature (310 K). As shown in Fig. S2(A)-(B), most replicas have certain persisting populations at 310 K, especially those replicas initially set at high temperature, indicating sufficient sampling of the replicas in temperature space. We selectively analyze the top five replicas with the highest persisting populations at 310 K. For K18, the conformations in replicas 0, 24, 26, and 28 are initially random coiled and that in replica 30 is partially structured. However, at the end of the simulation, all of these replicas have evolved into partially structured conformations, and the conformation in replica 30 is highly structured (Fig. S2(C)). For K19, two of the top five replicas are initially random and three are partially structured (Fig. S2(D)). Similar to K18, the initial conformations of K19 also experience random coil to α-helix/β-sheet (or vice versa) transitions (for example, Repl. ID 28 and Repl ID 2 in Fig. S2(D)). These results indicate K18 and K19 are not trapped in local energy minimum.

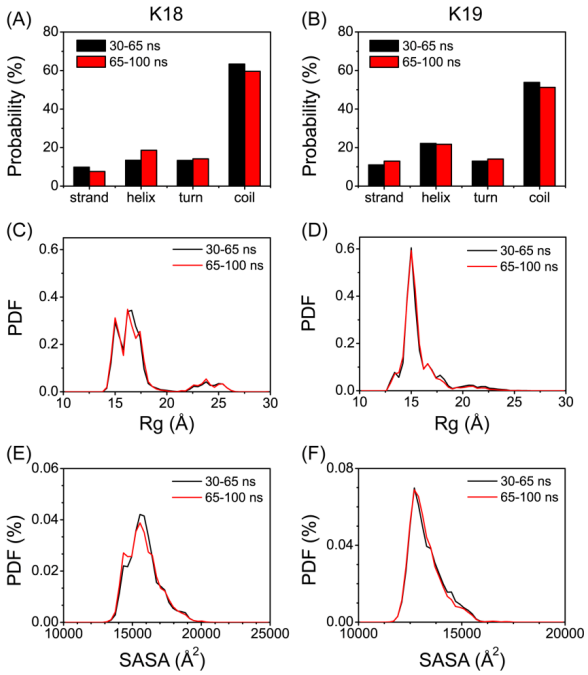


**Figure S3**: Convergence check for the REMD runs of K18 and K19 at 310 K. (A-B): Probability of secondary structures (including strand, helix, turn, and coil). (C-D): Probability density function (PDF) of radius of gyration (Rg) of K18 and K19 monomers. (E-F): PDF of the solvent accessible surface area (SASA) of K18 and K19 monomers within the time intervals of 30-65 ns and 65-100 ns.

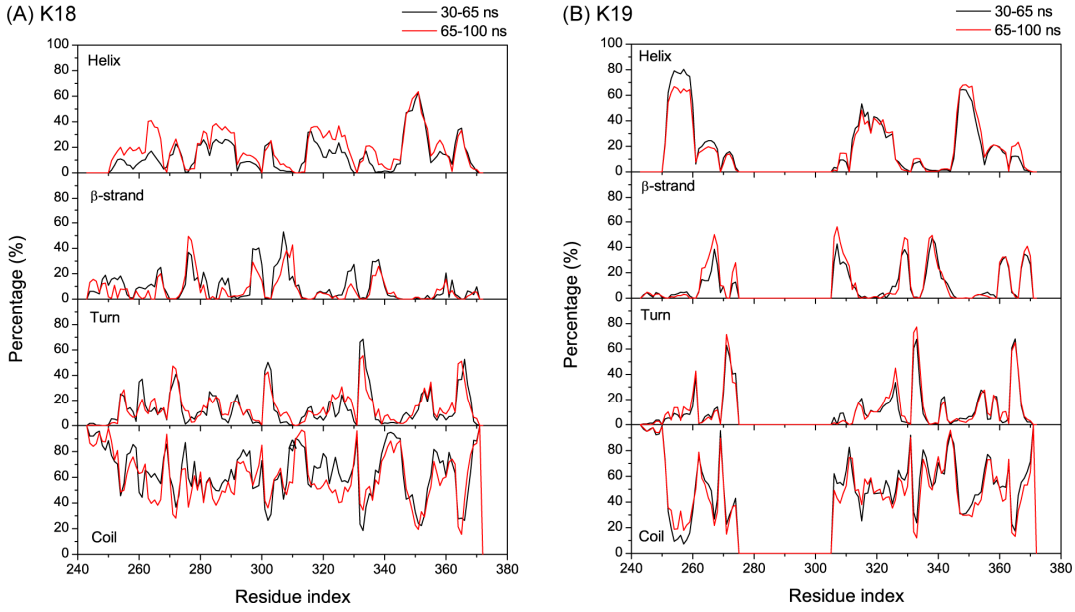


**Figure S4**. Convergence check for the REMD runs of K18 and K19 at 310 K. The alpha-helix, β-strand, turn and coil percentages of both K18 (A) and K19 (B) as a function of amino acid residues within the two time windows: 30-65 ns and 65-100 ns.

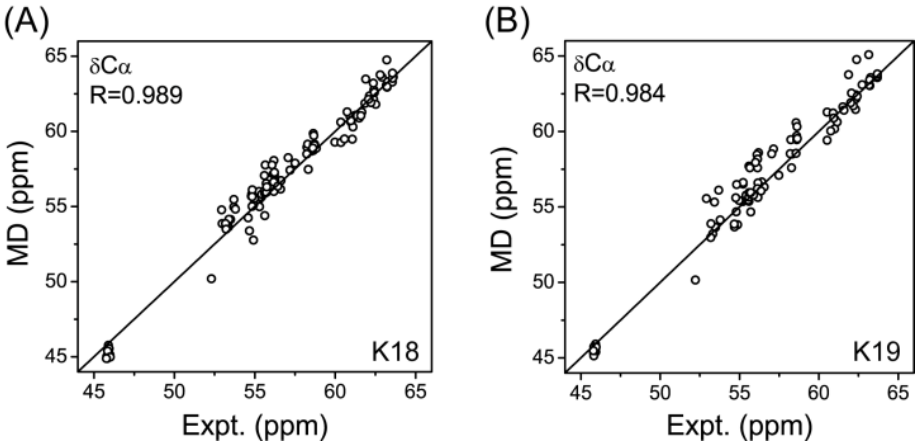


**Figure S5**. Scatter plots comparing Expt and MD-predicted chemical shifts of $C_\alpha$ atom for (A) K18 and (B) K19. The Pearson correlation coefficients (R) between experimental-measured and MD-generated chemical shifts are indicated.

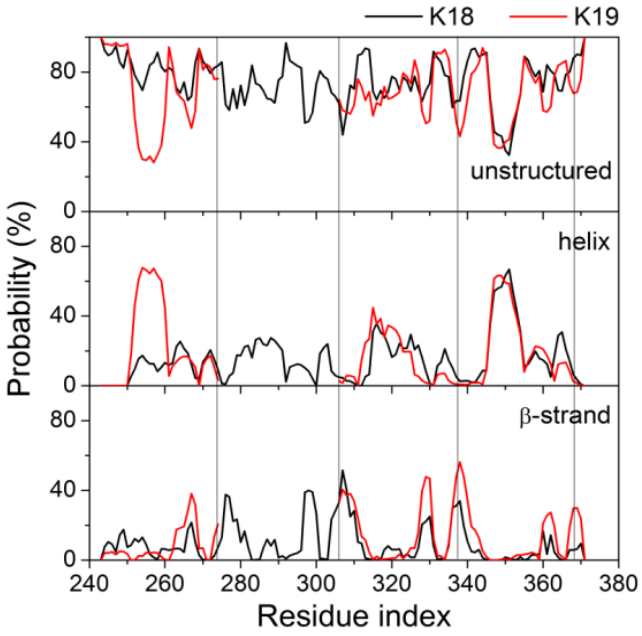


**Figure S6**. Secondary structure analysis of K18 and K19 monomer in aqueous solution at T=310 K. In both monomers, β-sheet regions are mostly located on both sides of the PGGG regions at the termini of each repeat, while α-helix regions are mainly in the middle of each repeat. Compared with K18, K19 has higher β-sheet propensity, especially in R3 and R4 regions, and higher α-helix propensity in R1 region. Each repeat has a similar β-α-β-turn structure pattern, but with different probabilities. The distribution of secondary structures in both K18 and K19 is similar to those obtained by NMR characterization[24, 31, 34].

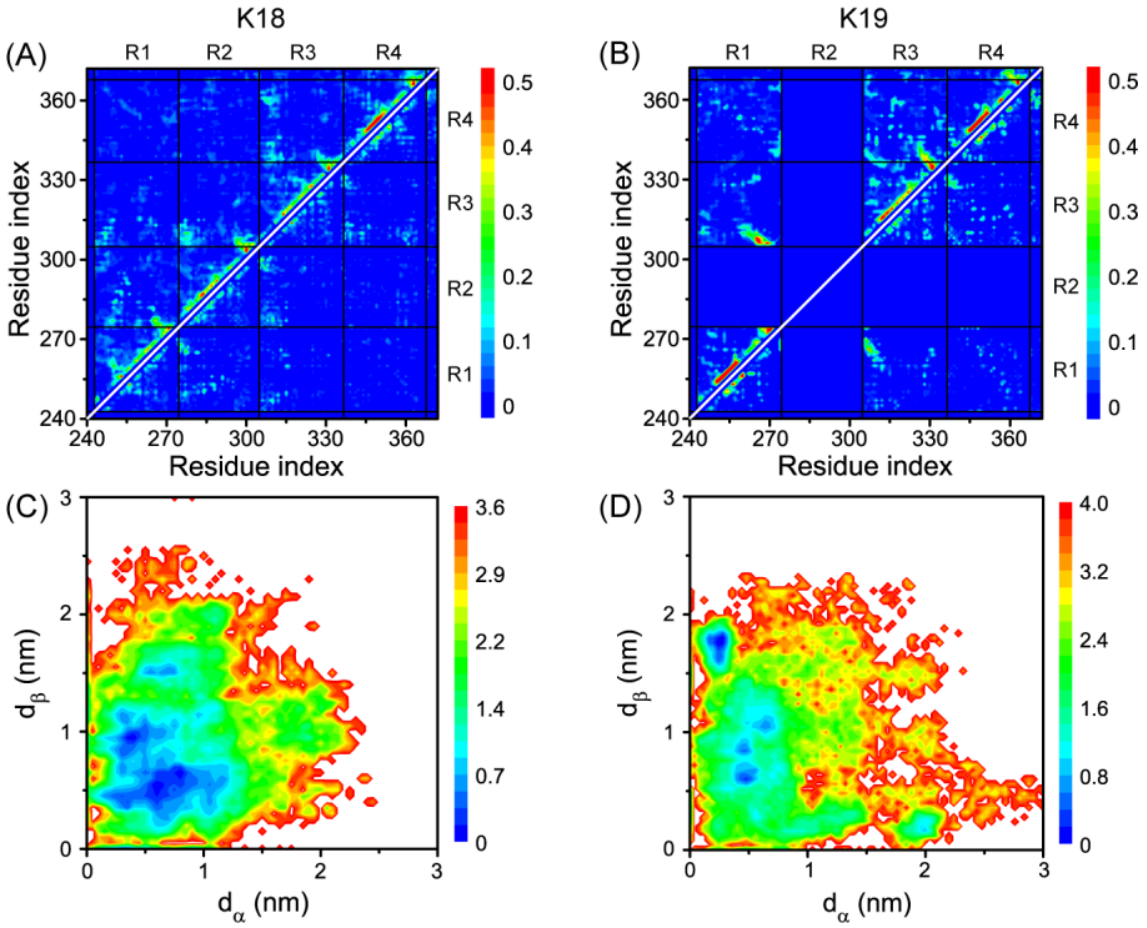


**Figure S7**. Contact probability maps (A, B) and the distribution of the secondary structures (C, D) indicate that domain-domain arrangement is not random. In the contact probability map for K18 (A) and K19 (B), the upper and lower triangular parts represent the residue-residue and side chain-side chain contacts, respectively. The contacts between residue i and residue i+1 and those between residue i and residue i+2 are not considered since the probabilities of these contacts are always close to 1 regardless of the secondary structure. The contacts along the diagonal line correspond to helical structures. Linear shapes far from the diagonal line correspond to parallel and anti-parallel β-sheet structures. For each conformer, we calculate the distributions of the distance between the centroid of the α-helix/β-sheet to the protein centroid. The distributions of the secondary structures of K18 (C) and K19 (D) indicate that locally structured regions are not in the center of mass for the IDP tau protein.

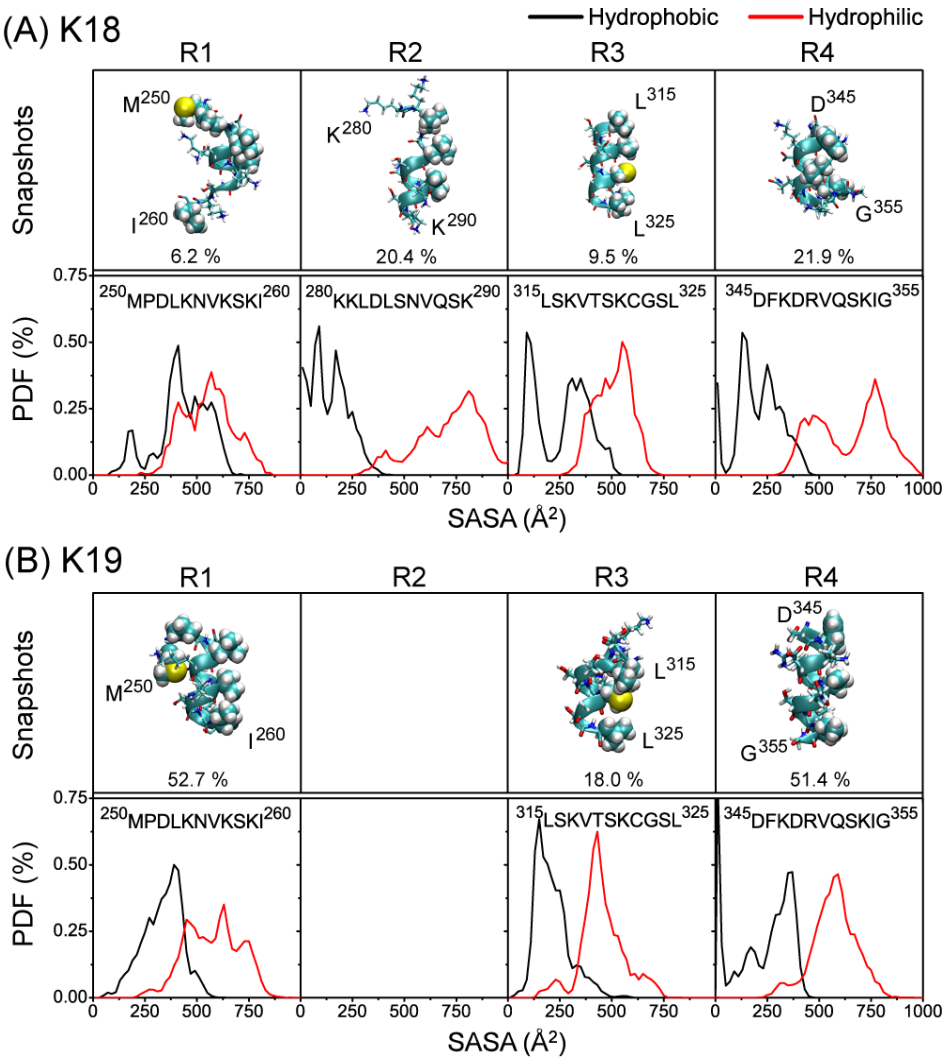


**Figure S8**. Representative structure of the most populated helical conformation in each repeat of K18/K19 and the probability distribution of hydrophilic and hydrophobic SASAs of the helix. Helical structures are mainly located in the middle region of each repeat: i.e. $^{250}$MPDLKNVKSKI$^{260}$ in R1, $^{280}$KKLDLSNVQSK$^{290}$ in R2, $^{315}$LSKVTSKCGSL$^{325}$ in R3, and $^{345}$DFKDRVQSKIG$^{355}$ in R4. We performed a RMSD-based cluster analysis with a backbone-RMSD cutoff of 3 Å for each repeat to identify the most populated helix in each repeat. The probability curve of hydrophobic SASAs in R3 of K18 and that in R4 of K19 both display a peak at a larger SASA=370 Å$^2$, very close to the hydrophilic SASA peak around 500 Å$^2$, implying the amphiphilic feature of the helix.

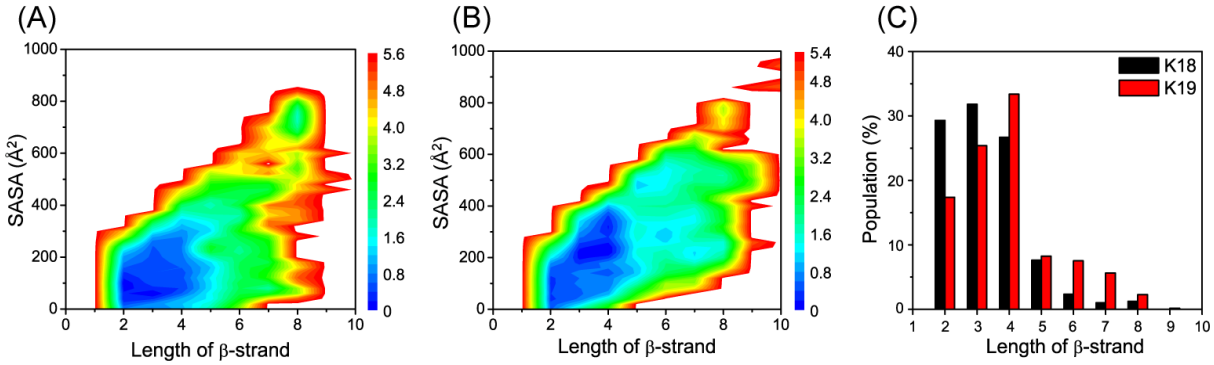


**Figure S9**. Free energy landscape of (A) K18 and (B) K19 as a function of β-strand length and SASA, and (C) the probability of β-strand length. Our analysis indicated that short -sheet stretches linked by turn regions frequently exist in the K18 and K19 monomers. Such stretches often cooperatively trigger fibril formation in long amyloidogenic sequences[[35]]. As seen from Figure S9, the most populated β-strand lengths in both K18 and K19 range from 2 to 5. The large variation of SASA from 0~400 Å2 indicates that these stretches can be buried or exposed to solvent. In the fibril structural models of K18 and K19 with a fully extended repeat[8-9, 31], the long β-strand stretch can span 20 residues. It is interesting to see what re-arrangements are needed to change the short β-sheet stretches to fully extended repeats. In our simulations, we found that there are some residual β-structures located in the C-terminal of R1, R2, R3 and R4, and the N-terminal of R2, R3, and R4, most of which form local β-hairpin structures between the C-terminal of one repeat and the N-terminal of the next. However, in the first eight clusters, we found some non-local β-sheet structures linking more than two repeats. In K18, the β-sheets B1a-B1b-B1c in C3 (5.67%) and C8 (2.64%) linking R1-R2-R3 together and the β-sheet B1a-B1b-B1c in C7 (2.79%) linking R2-R3-R4, located outside of the compact structure, are partially exposed to the solvent and keep the elongated faces towards to the solvent (Fig. 2(A)). In K19, the β-sheet B1a-B1b-B1c-B1d in C3 (6.94%) and C7 (2.37%) linking R1-R3-R4 also exist (Fig. 2(C)). Therefore, the topologies of the short β-stretches indicate that they may need to be reorganized upon β-stretch elongation.

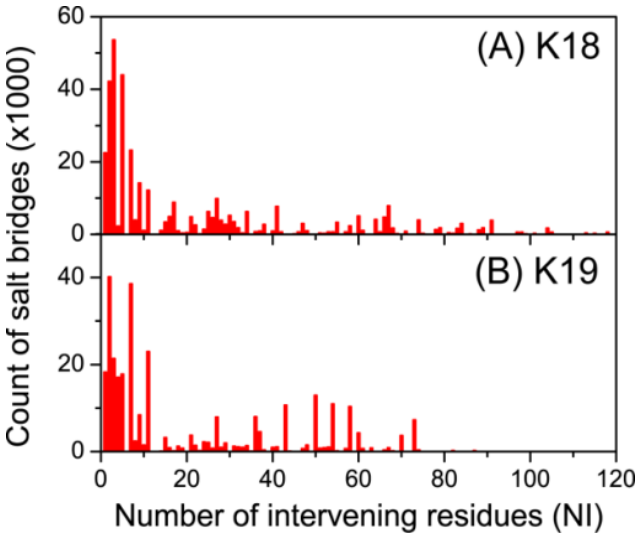


**Figure S10**. Histograms show the count of salt bridges with respect to the number of intervening residues for K18 and K19 at 310 K. The number of intervening residues is the number of residues separating the two salt-bridging residues in the sequence.

## References


(1) Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **2010,** *5*, 725-738.

(2) Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **2008,** *9*, 40.

(3) Roy, A.; Yang, J.; Zhang, Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.* **2012,** *40*, W471-477.

(4) Xu, D.; Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Struct., Funct., Bioinf.* **2012,** *80*, 1715-1735.

(5) Kim, D. E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **2004,** *32*, W526-531.

(6) Karplus, K.; Karchin, R.; Draper, J.; Casper, J.; Mandel-Gutfreund, Y.; Diekhans, M.; Hughey, R. Combining local-structure, fold-recognition, and new fold methods for protein structure prediction. *Proteins* **2003,** *53 Suppl 6*, 491-496.

(7) Karplus, K.; Katzman, S.; Shackleford, G.; Koeva, M.; Draper, J.; Barnes, B.; Soriano, M.; Hughey, R. SAM-T04: what is new in protein-structure prediction for CASP6. *Proteins: Struct., Funct., Bioinf.* **2005,** *61 Suppl 7*, 135-142.

(8) 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*, 14950-14959.

(9) Siddiqua, A.; Luo, Y.; Meyer, V.; Swanson, M. A.; Yu, X.; Wei, G.; Zheng, J.; Eaton, G. R.; Ma, B.; Nussinov, R.; Eaton, S. S.; Margittai, M. Conformational basis for asymmetric

seeding barrier in filaments of three- and four-repeat tau. *J. Am. Chem. Soc.* **2012,** *134*, 10271-10278.
(10) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999,** *314*, 141-151.
(11) Hansmann, U. H. E.; Okamoto, Y. Prediction of peptide conformation by multicanonical algorithm: New approach to the multiple-minima problem. *J. Comput. Chem.* **1993,** *14*, 1333-1338.
(12) Hansmann, U. H. E. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997,** *281*, 140-150.
(13) Potoyan, D. A.; Papoian, G. A. Energy landscape analyses of disordered histone tails reveal special organization of their conformational dynamics. *J. Am. Chem. Soc.* **2011,** *133*, 7405-7415.
(14) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008,** *4*, 435-447.
(15) Bjelkmar, P. r.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **2010,** *6*, 459-466.
(16) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126*, 014101.
(17) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981,** *52*, 7182-7190.
(18) 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*, 1463-1472.
(19) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992,** *13*, 952-962.
(20) Day, R.; Bennion, B. J.; Ham, S.; Daggett, V. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J. Mol. Biol.* **2002,** *322*, 189-203.
(21) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. Folding funnels and binding mechanisms. *Protein Eng.* **1999,** *12*, 713-720.
(22) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** *22*, 2577-2637.
(23) Shen, Y.; Bax, A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **2007,** *38*, 289-302.
(24) Mukrasch, M. D.; Biernat, J.; von Bergen, M.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. Sites of tau important for aggregation populate {beta}-structure and bind to microtubules and polyanions. *J. Biol. Chem.* **2005,** *280*, 24978-24986.
(25) Wang, Y.; Chu, X.; Longhi, S.; Roche, P.; Han, W.; Wang, E.; Wang, J. Multiscaled exploration of coupled folding and binding of an intrinsically disordered molecular recognition element in measles virus nucleoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110*, E3743-3752.
(26) Gurry, T.; Ullman, O.; Fisher, C. K.; Perovic, I.; Pochapsky, T.; Stultz, C. M. The dynamic structure of alpha-synuclein multimers. *J. Am. Chem. Soc.* **2013,** *135*, 3865-3872.
(27) Yamamoto, S.; Watarai, H. Raman optical activity study on insulin amyloid- and prefibril intermediate. *Chirality* **2012,** *24*, 97-103.

(28) Singh, Y.; Sharpe, P. C.; Hoang, H. N.; Lucke, A. J.; McDowall, A. W.; Bottomley, S. P.; Fairlie, D. P. Amyloid formation from an alpha-helix peptide bundle is seeded by 3(10)-helix aggregates. *Chemistry* **2011,** *17*, 151-160.
(29) Kirkitadze, M. D.; Condron, M. M.; Teplow, D. B. Identification and characterization of key kinetic intermediates in amyloid beta-protein fibrillogenesis. *J. Mol. Biol.* **2001,** *312*, 1103-1119.
(30) Mukrasch, M. D.; Bibow, S.; Korukottu, J.; Jeganathan, S.; Biernat, J.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. Structural polymorphism of 441-residue tau at single residue resolution. *PLoS Biol.* **2009,** *7*, e1000034.
(31) Luo, Y.; Dinkel, P.; Yu, X.; Margittai, M.; Zheng, J.; Nussinov, R.; Wei, G.; Ma, B. Molecular insights into the reversible formation of tau protein fibrils. *Chem. Commun. (Cambridge, U. K.)* **2013,** *49*, 3582-3584.
(32) Daebel, V.; Chinnathambi, S.; Biernat, J.; Schwalbe, M.; Habenstein, B.; Loquet, A.; Akoury, E.; Tepper, K.; Muller, H.; Baldus, M.; Griesinger, C.; Zweckstetter, M.; Mandelkow, E.; Vijayan, V.; Lange, A. beta-Sheet core of tau paired helical filaments revealed by solid-state NMR. *J. Am. Chem. Soc.* **2012,** *134*, 13982-13989.
(33) Dinkel, P. D.; Siddiqua, A.; Huynh, H.; Shah, M.; Margittai, M. Variations in filament conformation dictate seeding barrier between three- and four-repeat tau. *Biochemistry* **2011,** *50*, 4330-4336.
(34) Ozenne, V.; Schneider, R.; Yao, M.; Huang, J. R.; Salmon, L.; Zweckstetter, M.; Jensen, M. R.; Blackledge, M. Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J. Am. Chem. Soc.* **2012,** *134*, 15138-15148.
(35) Hu, L.; Cui, W.; He, Z.; Shi, X.; Feng, K.; Ma, B.; Cai, Y. D. Cooperativity among short amyloid stretches in long amyloidogenic sequences. *PLoS One* **2012,** *7*, e39369.