Supplementary Information

Molecular Insights into the Reversible Formation of Tau Protein Fibrils

Yin Luo[1], Paul Dinkel[2], Xiang Yu[3], Martin Margittai[2], Jie Zheng[3], Ruth Nussinov[4,5], Guanghong Wei[1], and Buyong Ma[4*]*

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

[2]Department of Chemistry & Biochemistry, University of Denver, Denver, Colorado 80208

[3]Department of Chemical & Biomolecular Engineering, the University of Akron, Akron, Ohio 44325

[4]Basic Science Program, SAIC-Frederick, Inc. Center for Cancer Research Nanobiology Program, Frederick National Lab, Frederick, Maryland 21702

[5]Sackler Inst. of Molecular Medicine Department of Human Genetics and Molecular Medicine Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Computational

Replica exchange molecular dynamics (REMD) simulations

All REMD simulations were performed using the GROMACS-4.5.3 software package [1] with all-atom CHARMM27 force field [2]. The exchange time between the adjacent replicas is 0.2 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[3] 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[4]. Bond lengths within peptides and water molecules were respectively constrained by the LINCS [5] and the SETTLE algorithms [6]. This allows 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^- ions were added to the solution with a ion concentration of 0.1 mol/L. The DSSP program was used to determine the secondary structure [7]. The atom-based contact is defined when carbon atoms of two non-sequential residues' side-chains or main-chains come within 0.54 nm or any other non-hydrogen atoms lies within 0.46 nm [8]. The atom-based contact number is the number of atom-pairs satisfied by the atom-based contact criterion.

K18 monomer and K19 monomer. Each REMD simulation consists of 48 replicas. The simulation temperature ranges from 310 to 430 K with an exponential distribution. Each replica runs for 100 ns (total simulation time is 4.8 μs), and the last 70 ns of the trajectories were used for analysis, which represent 3,336,000 conformations.

K18 octamer and K19 octamer: Each REMD simulation consists of 64 replicas, with the simulation temperature ranging from 310 to 380 K with an exponential distribution. Each replica runs for 11 ns (the total simulation time is 0.706 μs), and the last 5 ns of the trajectories were used for the analysis, which represent 320,000 conformations. For each REMD simulation, there are total 8 initial conformations, 7 of which are constructed using U-turn like structures which have been verified experimentally[9], and the last one is put together by the structure constructed

by homology modeling. Thus, the 64 replicas are initially evenly distributed with 8 sets of 8 conformers from the 8 models. Each conformer in the same set is selected from a short simulation with the temperature rising from 310 K to 600 K, to obtain a decreasing population of ordered β-structures.

Percolation transition of hydration water on the surface of K18 Octamer

To investigate the percolation transition temperature of hydration water around K18 octamer, we calculated the probability distribution of the largest water cluster in the hydration shell at different temperatures (Sup-Figure 2). In the calculation, the last 10 ns data were used. A water molecule was considered as belonging to the hydration shell when the minimal distance between its oxygen atom and the heavy atoms of protein is smaller than the defined shell width D (here, we chose D=4.8 Å) [10]. Water molecules were considered in the same cluster if they are connected by an uninterrupted path of hydrogen-bonds. Two water molecules were considered as hydrogen-bond connected if the distance between their oxygen atoms is less than 3.5 Å and their pair interaction energy was below -2.3 kcal/mol [10]. The probability distribution curve of the largest water cluster in the hydration shell in Sup-Fig. 2 shows a two-peak structure with a left (low S) and right (high S) peak, corresponding to respectively the nonspanning and spanning clusters at T=330 K, indicating that this temperature is close to the percolation threshold [11]. The minimum between the two peaks gives the minimal water cluster size S' (here S'=2500), necessary for the largest cluster to have spanning hydrogen-bond network. The spanning probabilities R at 64 different temperatures were estimated by integrating $P(S_{max})$ for $S_{max} > S'$ (Sup-Fig. 2B). The spanning probability is almost 100% at T=310 K, and drops to zero at T=345 K. The percolation transition temperature is close to T=330 K, with the spanning probability of 50%. The snapshots at T=310, 330, and 380 K were illustrated to show the three largest water clusters and the spanning and nonspanning hydrogen-bond network (Sup-Fig. 2C).

Experimental

Expression, purification, and labeling of tau proteins

Tau constructs K18 and K19 were expressed and purified as previously described[12]. Briefly, expression was performed using *Escherichia coli* (BL21 (DE3)). After 3 hour induction with

IPTG at 37 °C, bacteria were sedimented, pellets resuspended in buffer (500 mM NaCl, 20 mM PIPES, pH 6.4, 1 mM EDTA, 50 mM 2-mercaptoethanol) and stored at -80 °C until later purification. Samples were heated to 80 °C for 20 min, sonicated, and centrifuged at 15,000*g* for 30 min. To the supernatants a 60% weight to volume amount of ammonium sulfate was added. Samples were incubated for 1 hr with gentle rocking and centrifuged 10 minutes at 15,000*g*. Pellets were resuspeded in dH_2O with 2 mM dithiothreitol (DTT), sonicated, passed through a syringe filter, and applied onto a MonoS cation exchange column (GE Healthcare). Tau was eluted with a 50 mM to 1M NaCl gradient and fractions were analyzed for purified protein by SDS-PAGE. Enriched fractions were applied to a Superdex 200 gelfitration column and tau was precipitated overnight at 4 °C with a 4x volumetric excess of acetone (5 mM DTT). Precipitated tau was sedimented, washed and stored in fresh acetone (2 mM DTT). Pellets of K18 were solubilized and, where applicable, acrylodan labeled as previously described[13]. Concentrations of monomeric tau were determined by the BCA method (Pierce).

Aggregation of tau fibrils at high temperature and subsequent cooling

Aggregation reactions were carried out in HEPES buffer (10 mM HEPES, pH 7.4, 0.1 M NaCl or 2.0 M NaCl). Total tau monomer concentrations were 10 μM, with a molar ratio of 98% cystein-free (cysteines replaced with serines) and 2% acrylodan-labeled tau (position 310 ─ or position 309 in control experiment ─ replaced with cysteine and labeled). Heparin (average molecular weight of 4400, Celsus, Cincinnati, OH) when present was 20 μM. Stirring was utilized in all reactions to promote aggregation. The progression of fibrillization was monitored using a Fluorolog 3 fluorometer (Horriba, Jobin). All reactions were allowed to proceed for 80 min at 343 K. Temperature was adjusted using a solid state Pelletier element. Upon cooling from 343 K to 275 K, samples were allowed to equilibrate and were then measured. Excitation of samples occurred at 360 nm with a 5 nm excitation and emission slit width.

Inhibition of heparin-mediated fibril growth by high salt

Tau fibrils were prepared by incubating 25 μM K18 or K19 with 50 μM heparin (10 mM HEPES, pH 7.4) for 3 days at 25 °C under agitating conditions. Fibrils were chilled on ice for 10 min and sonicated for 20 seconds to induce breakage. Monomeric K18 and K19 (10 μM; 98%

cysteine free tau and 2% acrylodan labeled at position 310) were seeded with a 3% molar addition of sonicated fibrils composed of their respective construct. Reaction conditions were maintained at 10 mM HEPES, pH 7.4 in the presence of low (0.1 M NaCl) or high (1 M NaCl) salt. Heparin was added in all reactions at 20 μM. Assembly of tau monomers on fibril seeds was monitored by observing the shift in fluorescence maximum of acrylodan label. Excitation of label occurred at 360 nm with 5 nm excitation and emission slit widths.

Electron microscopy

250-mesh carbon coated copper grids were placed on 10 μL drops of the incubated sample for 40 s, 30 s on 10 μL 2% uranyl acetate, and air-dried on filter paper. Samples were imaged with a Philips/FEI Tecnai-12 electron transmission microscope equipped with a Gatan CCD camera.

Acknowledgements—This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government. This research was supported (in part) by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. J.Z. thanks for financial support from NSF grants (CAREER Award CBET-0952624 and CBET-1158447). G.W. thanks for financial support from National Natural Science Foundation of China (Grant No. 11074047) and Research Fund for the Doctoral Program of Higher Education of China (RFDP-20100071110006). M.M. thanks for financial support from the National Institute of Neurological Disorders and Stroke (Grant Number R01NS076619). The simulations had been performed using the high-performance computational facilities of the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov).

Sup-Table 1: The percentage of secondary structure distributions for K18 and K19 monomers at 310 K.

	K18-monomer				K19-monomer			
	β	Helix	turn	coil	β	Helix	turn	coil
R1N	6.18	1.61	3.31	88.89	2.83	0.43	2.74	94.00
R1M	4.41	7.42	18.69	69.47	2.07	54.21	13.68	30.03
R1C	11.01	8.95	10.09	69.94	19.29	11.23	5.46	64.03
R1T	0.91	10.8	46.25	42.04	4.71	15.15	38.81	41.33
R2N	17.66	5.07	8.51	68.76	--	--	--	--
R2M	3.54	10.42	17.79	68.26	--	--	--	--
R2C	19.01	1.93	6.62	72.45	--	--	--	--
R2T	4.59	7.14	32.55	55.72	--	--	--	--
R3N	21.71	1.00	1.28	76.02	21.92	11.21	8.60	58.27
R3M	3.43	13.37	11.10	72.10	0.65	25.25	12.83	61.27
R3C	9.92	3.64	8.28	78.17	19.91	4.16	8.25	67.68
R3T	2.63	7.93	52.48	36.97	9.37	3.86	26.59	60.18
R4N	11.28	2.27	4.94	81.51	25.84	1.77	6.71	65.68
R4M	0.61	31.93	13.86	53.61	1.36	37.82	9.53	51.28
R4C	4.20	5.48	8.49	81.84	13.44	13.36	8.08	65.12
R4T	2.89	19.65	39.92	37.54	8.03	8.37	33.80	49.80
Tail	4.62	4.73	7.00	83.66	15.44	0.67	1.52	82.37

Sup-Figure 1. (A) Region organization of K18 and K19. Each repeat is divided to 4 regions: N-terminal region (RxN), middle region (RxM), C-terminal region (RxC) and the turn region (RxT), where x = 1, 2, 3, or 4. The last four residues after the R4 repeat is called tail region (T). The first residue index of each region is labeled with the number at the left-bottom corner. The preferred secondary structures are shown with colors: β-structure (red), helix (orange), turn (blue) and random (green). (B) Average inter-region protein-protein contact number for K18 and K19 at T=310 K. The contact numbers for K18 and K19 are both averaged over 70,000 structures.

Sup-Figure 2. Protein solvation energy (per residue) versus simulation temperature for K18 and K19 octamers. (A) Total solvation energy. (B) Protein-solvent interaction energy. (C) Cavity hydration energy (due to excluded volume effect). (D)-(E) Decomposition of protein-solvent interaction energy: electrostatic and VDW interactions.

Sup-Figure 3. K18 fibrils formed in the presence of high salt cold dissociate. Spectra of monomeric tau before (green) and after (blue) 80 min aggregation at 343 K. When aggregated tau was cooled to 275 K, the emission spectrum (purple) red shifted indicating cold dissociation. Protein concentration = 10 μM. NaCl concentration = 2.0 M. These data suggest that under the given experimental conditions 2.0 M NaCl are not sufficient to stabilize the K18 fibril.

Sup-Figure 4. K19 resists aggregation at 343 K. Monomer mixtures containing 98% K19 and 2% acrylodan-labeled K19 were measured before (green spectra) and after (blue spectra) 80 min incubation at 343 K. (A) 10 μM K19, acrylodan label at position 310. (B) 20 μM K19, acrylodan label at position 310. (C) 10 μM K19, acrylodan label at position 309. The lack in blue shift indicates that under the specified conditions K19 does not aggregate. The labeling position was changed in order to ensure that the inability to aggregate was not caused by structural incompatibility with the modification.

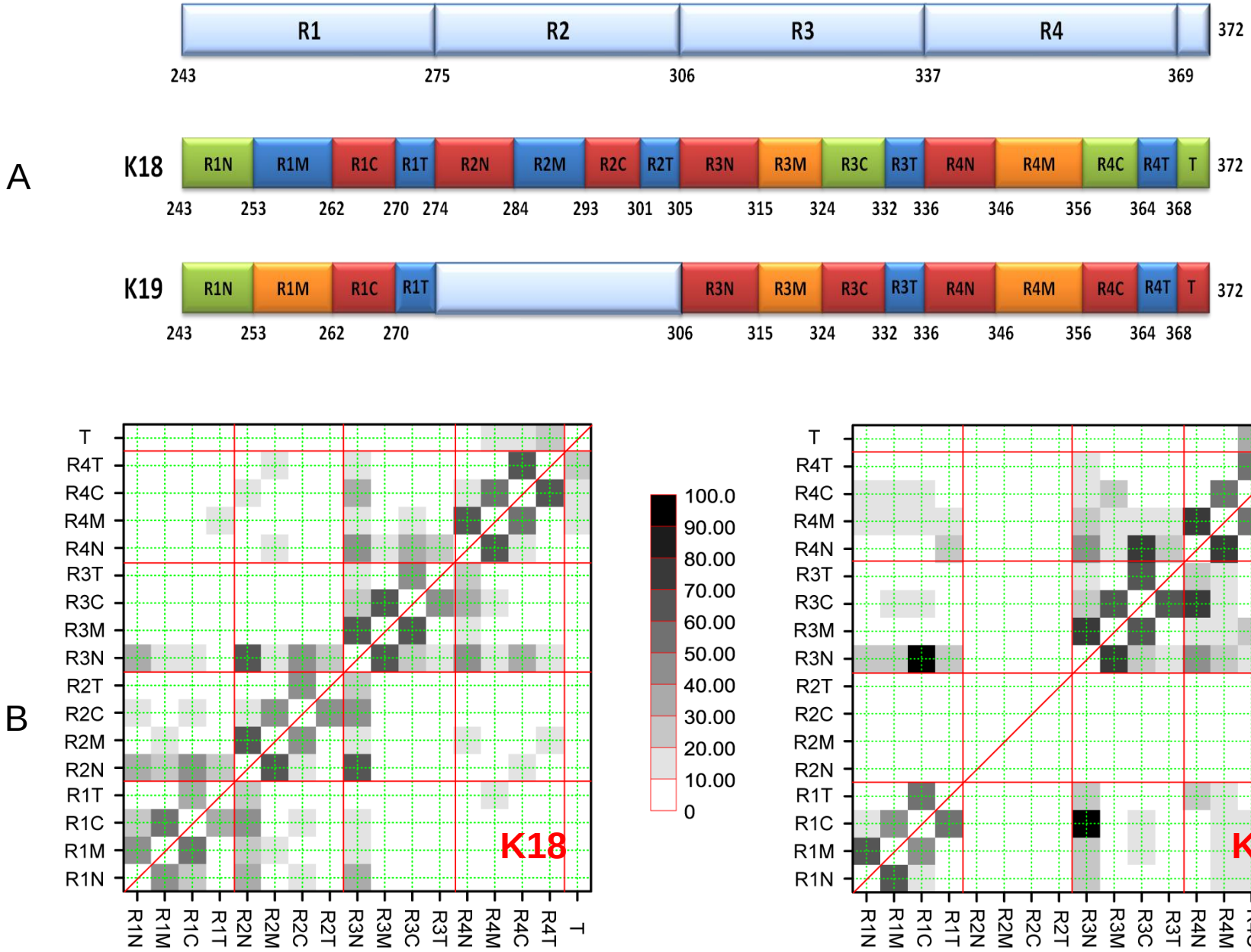
Sup-Figure 5. (A) Probability distribution $P(S_{max})$ of the size S_{max} of the largest water cluster in the hydration shell of the K18 oligomer. (B) Spanning probability *R* at different temperatures. (C)

Snapshots at 6 ns of K18 oligomer system at T=310 K (fully spanning), 330 K (partially spanning), and 380 K (nonspanning). The water molecules within the first hydration shell (D=4.8 Å) are shown, with the molecules in the first, second, and third largest water clusters being colored in red, yellow, and green, respectively. Water molecules in smaller sizes of clusters are represented with line style.

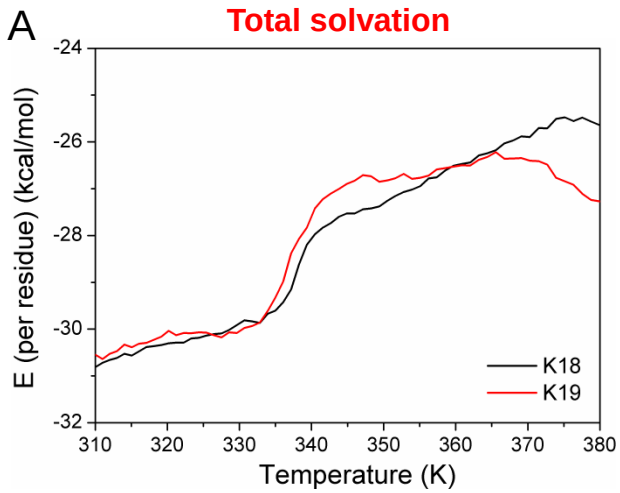
Sup-Figure 6. High salt inhibits heparin-mediated fibril growth of K18 and K19. The growth of acrylodan-labled tau monomers onto fragmented fibril seeds results in a shift in the emission maximum, which is monitored as a function of time. (A) Fibril growth of K18. (B) Fibril growth of K19. All samples contained 3% seeds, 10 μM respective K18 and K19 monomer (2% of monomers were labeled with acrylodan at position 310), and 20 μM heparin. Fibril growth in the presence of 0.1 M NaCl (blue) and 1 M NaCl (red). Values reperesent mean ±s.d. (n=3). The data indicate that heparin-mediated growth is inhibited in the presence of high salt.

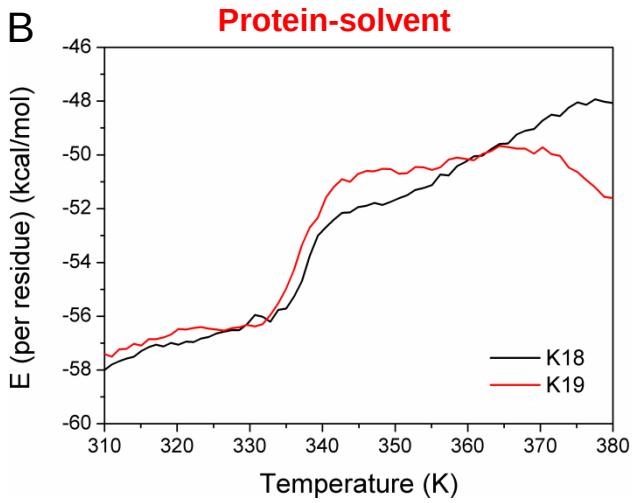
Reference

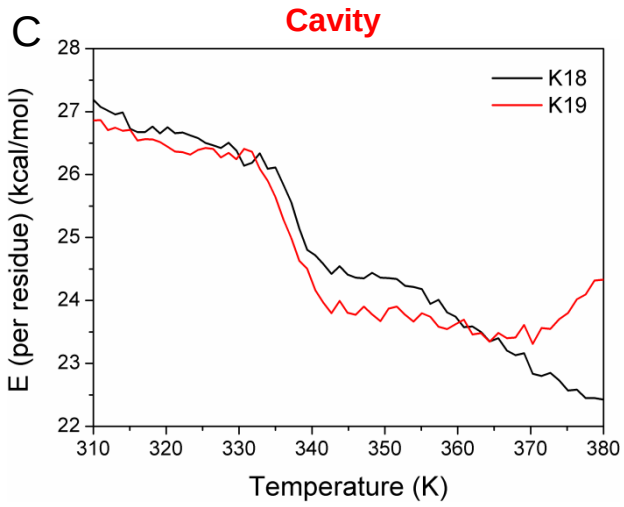
1. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, **4**, 435-447.
2. P. r. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess and E. Lindahl, *Journal of Chemical Theory and Computation*, 2010, **6**, 459-466.
3. G. Bussi, D. Donadio and M. Parrinello, *J Chem Phys*, 2007, **126**, 014101.
4. M. Parrinello and A. Rahman, *J Appl Phys*, 1981, **52**, 7182-7190.
5. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J Comput Chem*, 1997, **18**, 1463-1472.
6. S. Miyamoto and P. A. Kollman, *J Comput Chem*, 1992, **13**, 952-962.
7. W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577-2637.
8. A. Huet and P. Derreumaux, *Biophys J*, 2006, **91**, 3829-3840.
9. A. Siddiqua, Y. Luo, V. Meyer, M. A. Swanson, X. Yu, G. Wei, J. Zheng, G. R. Eaton, B. Ma, R. Nussinov, S. S. Eaton and M. Margittai, *Journal of the American Chemical Society*, 2012, **134**, 10271-10278.
10. I. Brovchenko, A. Krukau, N. Smolin, A. Oleinikova, A. Geiger and R. Winter, *The Journal of Chemical Physics*, 2005, **123**, 224905.
11. A. Oleinikova, N. Smolin, I. Brovchenko, A. Geiger and R. Winter, *The Journal of Physical Chemistry. B*, 2005, **109**, 1988-1998.
12. A. Siddiqua and M. Margittai, *J Biol Chem*, 2010, **285**, 37920-37926.
13. P. D. Dinkel, A. Siddiqua, H. Huynh, M. Shah and M. Margittai, *Biochemistry*, 2011, **50**, 4330-4336.

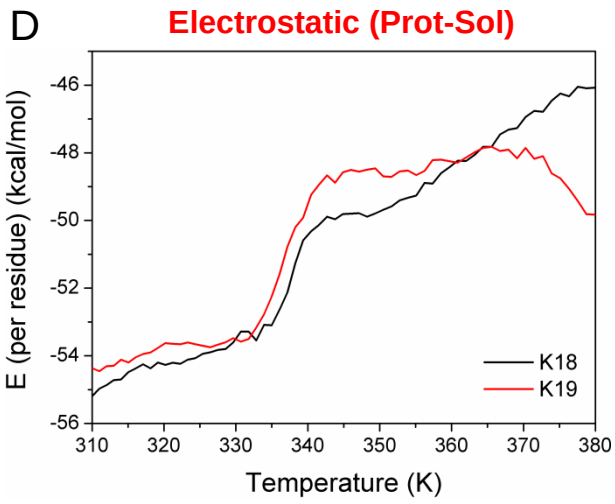

A
R1
R2
R3
R4
372
243
275
306
337
369
K18
R1N R1M R1C R1T R2N R2M R2C R2T R3N R3M R3C R3T R4N R4M R4C R4T T
372
243 253 262 270 274 284 293 301 305 315 324 332 336 346 356 364 368
K19
R1N R1M R1C R1T R3N R3M R3C R3T R4N R4M R4C R4T T
372
243 253 262 270 306 315 324 332 336 346 356 364 368
B
T R4T R4C R4M R4N R3T R3C R3M R3N R2T R2C R2M R2N R1T R1C R1M R1N
K18
R1N R1M R1C R1T R2N R2M R2C R2T R3N R3M R3C R3T R4N R4M R4C R4T T
100.0
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0
T R4T R4C R4M R4N R3T R3C R3M R3N R2T R2C R2M R2N R1T R1C R1M R1N
K19
R1N R1M R1C R1T R2N R2M R2C R2T R3N R3M R3C R3T R4N R4M R4C R4T T

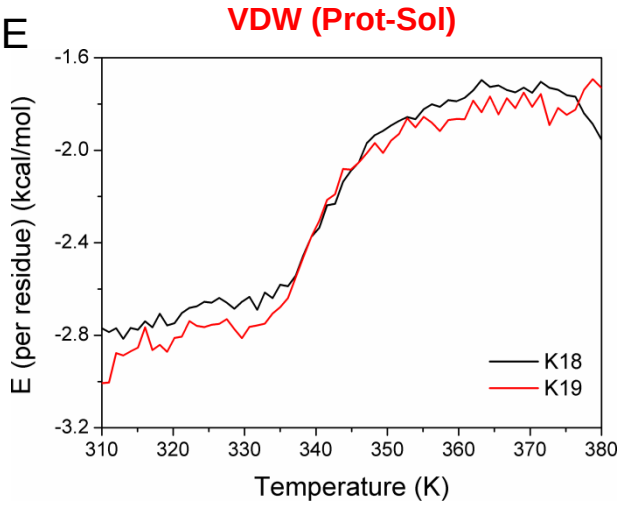
Sup-Figure 1


A
Total solvation
E (per residue) (kcal/mol)
Temperature (K)
K18
K19

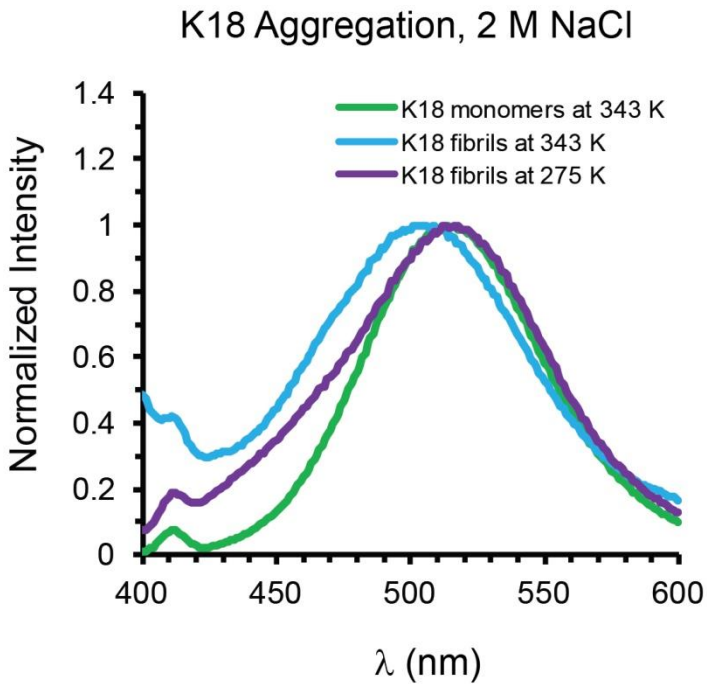

B
Protein-solvent
E (per residue) (kcal/mol)
Temperature (K)
K18
K19


C
Cavity
E (per residue) (kcal/mol)
Temperature (K)
K18
K19

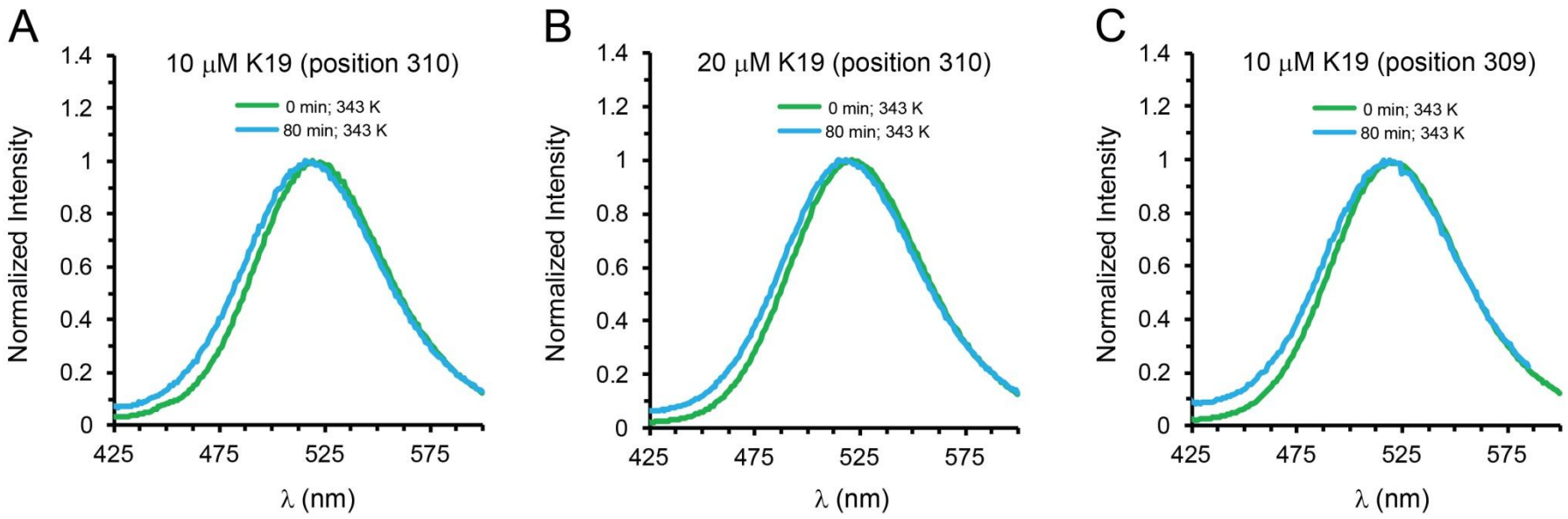

D
Electrostatic (Prot-Sol)
E (per residue) (kcal/mol)
Temperature (K)
K18
K19


E
VDW (Prot-Sol)
E (per residue) (kcal/mol)
Temperature (K)
K18
K19

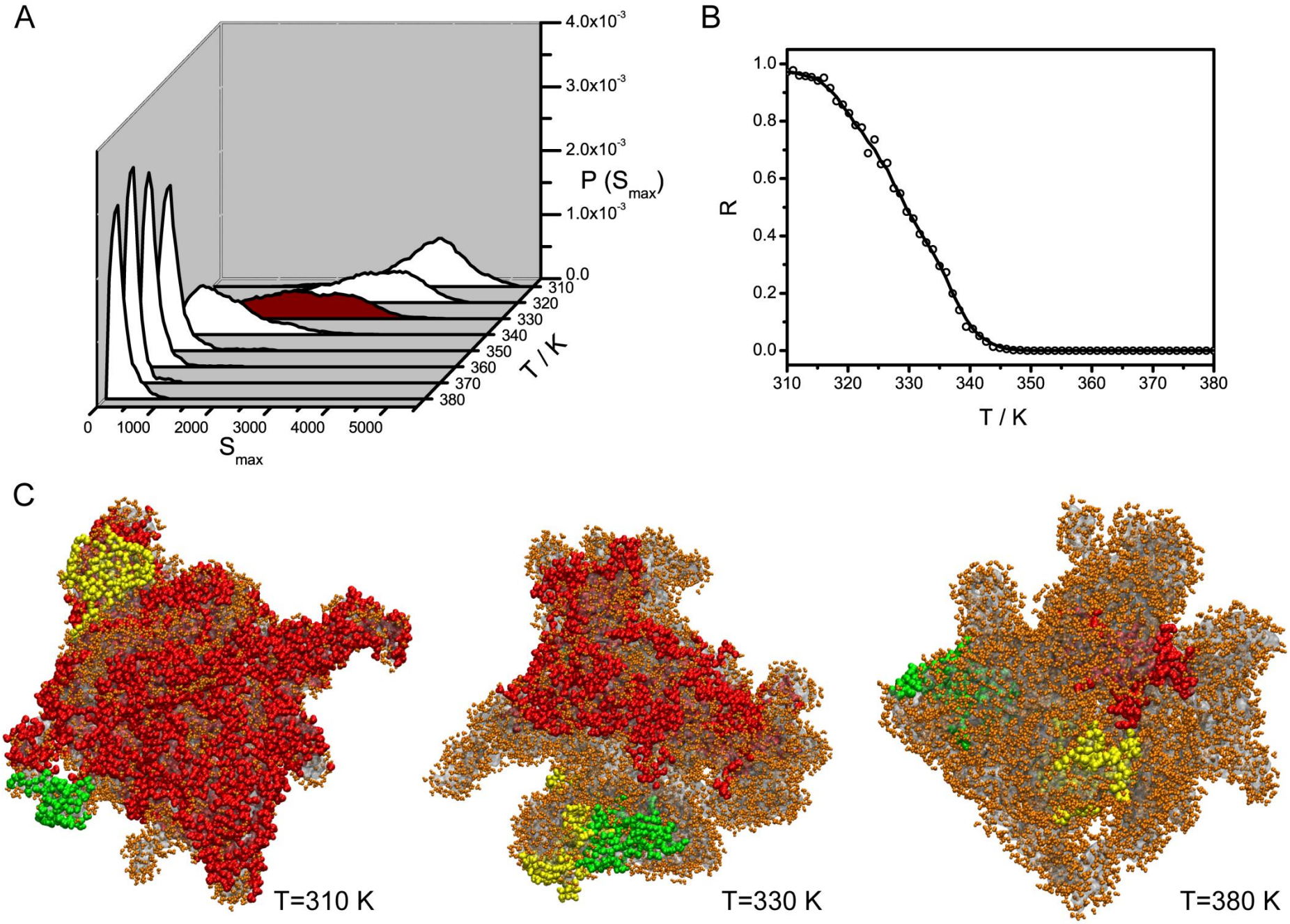
Sup-Figure 2


K18 Aggregation, 2 M NaCl
K18 monomers at 343 K
K18 fibrils at 343 K
K18 fibrils at 275 K
Normalized Intensity
1.4
1.2
1
0.8
0.6
0.4
0.2
0
400
450
500
550
600
λ (nm)

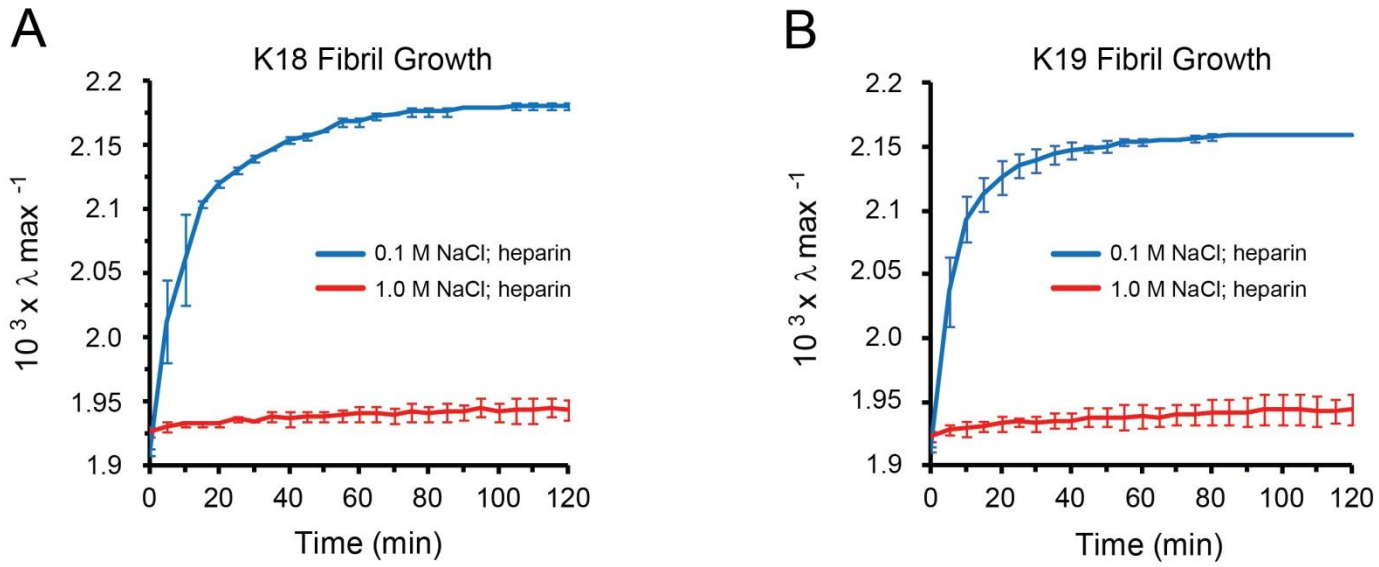
Sup-Figure 3


A
10 μM K19 (position 310)
0 min; 343 K
80 min; 343 K
Normalized Intensity
λ (nm)
B
20 μM K19 (position 310)
0 min; 343 K
80 min; 343 K
Normalized Intensity
λ (nm)
C
10 μM K19 (position 309)
0 min; 343 K
80 min; 343 K
Normalized Intensity
λ (nm)

Sup-Figure 4


A
4.0x10-3
3.0x10-3
2.0x10-3
P (Smax)
1.0x10-3
0.0
310
320
330
340
350
360
370
380
T / K
0
1000
2000
3000
4000
5000
Smax
B
1.0
0.8
0.6
0.4
0.2
0.0
R
310
320
330
340
350
360
370
380
T / K
C
T=310 K
T=330 K
T=380 K

Sup-Figure 5


A
K18 Fibril Growth
10^3 x λ max $^{-1}$
2.2
2.15
2.1
2.05
2.0
1.95
1.9
0
20
40
60
80
100
120
Time (min)
0.1 M NaCl; heparin
1.0 M NaCl; heparin
B
K19 Fibril Growth
10^3 x λ max $^{-1}$
2.2
2.15
2.1
2.05
2.0
1.95
1.9
0
20
40
60
80
100
120
Time (min)
0.1 M NaCl; heparin
1.0 M NaCl; heparin

Sup-Figure 6