Supplementary Information

Disease-associated patterns of acetylation stabilize tau fibril formation

Li Li[1], Binh Nguyen[1], Vishruth Mullapudi[1], Lorena Saelices[1,2], Lukasz A. Joachimiak[1,3,#]

Source data 1. Raw ThT experiments for R1R2, R2R3, R3R4, R1R3 and R4R' chemically modified peptides.

Source data 2. Raw seeding data for experiments for R1R2, R2R3 and R1R3 chemically modified peptides.

Supplementary Tables

3R Tau (263-280) numbering derived from 2N3R tau				
Name	**Acetylation sites**	**K acetylated peptides**	**Name**	**K to Q mutant peptides**
R1R3		TENLKHQPGGGKVQIVYK		
R1R3_1$_{Ac}$	K267	TENLK(Ac)HQPGGGKVQIVYK	R1R3_1$_{Q}$	TENLQHQPGGGKVQIVYK
R1R3_2$_{Ac}$	K274	TENLKHQPGGGK(Ac)VQIVYK	R1R3_2$_{Q}$	TENLKHQPGGGQVQIVYK
R1R3_3$_{Ac}$	K280	TENLKHQPGGGKVQIVYK(Ac)	R1R3_3$_{Q}$	TENLKHQPGGGKVQIVYQ
R1R3_12$_{Ac}$	K267,K274	TENLK(Ac)HQPGGGK(Ac)VQIVYK	R1R3_12$_{Q}$	TENLQHQPGGGQVQIVYK
R1R3_13$_{Ac}$	K267,K280	TENLK(Ac)HQPGGGKVQIVYK(Ac)	R1R3_13$_{Q}$	TENLQHQPGGGKVQIVYQ
R1R3_23$_{Ac}$	K274,K280	TENLKHQPGGGK(Ac)VQIVYK(Ac)	R1R3_23$_{Q}$	TENLKHQPGGGQVQIVYQ
R1R3_123$_{Ac}$	K267,K274,K280	TENLK(Ac)HQPGGGK(Ac)VQIVYK(Ac)	R1R3_123$_{Q}$	TENLQHQPGGGQVQIVYQ

4R Tau (263-280) numbering derived from 2N4R tau				
Name	**Acetylation sites**	**K acetylated peptides**	**Name**	**K to Q mutant peptides**
R1R2		TENLKHQPGGGKVQIINK		
R1R2_1$_{Ac}$	K267	TENLK(Ac)HQPGGGKVQIINK	R1R2_1$_{Q}$	TENLQHQPGGGKVQIINK
R1R2_2$_{Ac}$	K274	TENLKHQPGGGK(Ac)VQIINK	R1R2_2$_{Q}$	TENLKHQPGGGQVQIINK
R1R2_3$_{Ac}$	K280	TENLKHQPGGGKVQIINK(Ac)	R1R2_3$_{Q}$	TENLKHQPGGGKVQIINQ
R1R2_12$_{Ac}$	K267,K274	TENLK(Ac)HQPGGGK(Ac)VQIINK	R1R2_12$_{Q}$	TENLQHQPGGGQVQIINK
R1R2_13$_{Ac}$	K267,K280	TENLK(Ac)HQPGGGKVQIINK(Ac)	R1R2_13$_{Q}$	TENLQHQPGGGKVQIINQ
R1R2_23$_{Ac}$	K274,K280	TENLKHQPGGGK(Ac)VQIINK(Ac)	R1R2_23$_{Q}$	TENLKHQPGGGQVQIINQ
R1R2_123$_{Ac}$	K267,K274,K280	TENLK(Ac)HQPGGGK(Ac)VQIINK(Ac)	R1R2_123$_{Q}$	TENLQHQPGGGQVQIINQ

4R Tau (294-311) numbering derived from 2N4R tau				
Name	**Acetylation sites**	**K acetylated peptides**	**Name**	**K to Q mutant peptides**
R2R3		KDNIKHVPGGGSVQIVYK		
R2R3_1$_{Ac}$	K294	K(Ac)DNIKHVPGGGSVQIVYK	R2R3_1$_{Q}$	QDNIKHVPGGGSVQIVYK
R2R3_2$_{Ac}$	K298	KDNIK(Ac)HVPGGGSVQIVYK	R2R3_2$_{Q}$	KDNIQHVPGGGSVQIVYK
R2R3_3$_{Ac}$	K311	KDNIKHVPGGGSVQIVYK(Ac)	R2R3_3$_{Q}$	KDNIKHVPGGGSVQIVYQ
R2R3_12$_{Ac}$	K294, 298	K(Ac)DNIK(Ac)HVPGGGSVQIVYK	R2R3_12$_{Q}$	QDNIQHVPGGGSVQIVYK
R2R3_13$_{Ac}$	K294, 311	K(Ac)DNIKHVPGGGSVQIVYK(Ac)	R2R3_13$_{Q}$	QDNIKHVPGGGSVQIVYQ
R2R3_23$_{Ac}$	K298, 311	KDNIK(Ac)HVPGGGSVQIVYK(Ac)	R2R3_23$_{Q}$	KDNIQHVPGGGSVQIVYQ
R2R3_123$_{Ac}$	K294,298,K311	K(Ac)DNIK(Ac)HVPGGGSVQIVYK(Ac)	R2R3_123$_{Q}$	QDNIQHVPGGGSVQIVYQ

4R Tau (325-343) numbering derived from 2N4R tau		
Name	**Acetylation sites**	**K acetylated peptides**
R3R4		LGNIHHKPGGGQVEVKSEK
R3R4_1$_{Ac}$	K331	LGNIHHK(Ac)PGGGQVEVKSEK
R3R4_2$_{Ac}$	K340	LGNIHHKPGGGQVEVK(Ac)SEK
R3R4_3$_{Ac}$	K343	LGNIHHKPGGGQVEVKSEK(Ac)
R3R4_123$_{Ac}$	K331, K340, K343	LGNIHHK(Ac)PGGGQVEVK(Ac)SEK(Ac)

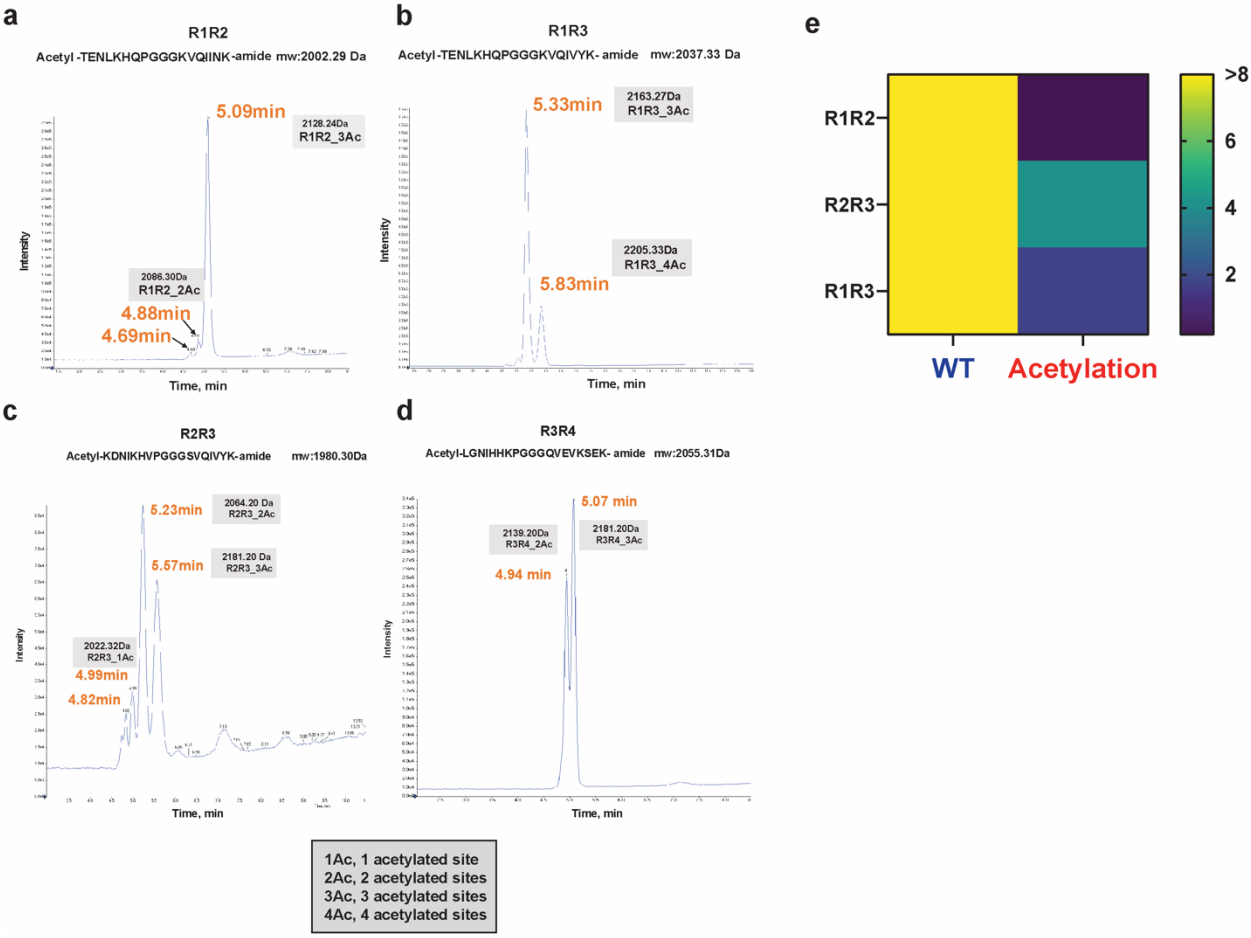
4R Tau (357-374) numbering deriged from 2N4R tau		
Name	**Acetylation sites**	**K acetylated peptides**
4RR'		LDNITHVPGGGNKKIETH
4RR'_1_{Ac}	K369	LDNITHVPGGGN(Ac)KIETH
4RR'_2_{Ac}	K370	LDNITHVPGGGNK(Ac)IETH
4RR'_12_{Ac}	K369, K370	LDNITHVPGGGNK(Ac)K(Ac)IETH

Supplementary Table 1. Nomenclature and sequences for all peptides used in the study. The site of acetylation or glutamine mutation is indicated by "Ac" and "Q", respectively, and is colored in red. All peptides were N-terminally acetylated and C-terminally amidated.

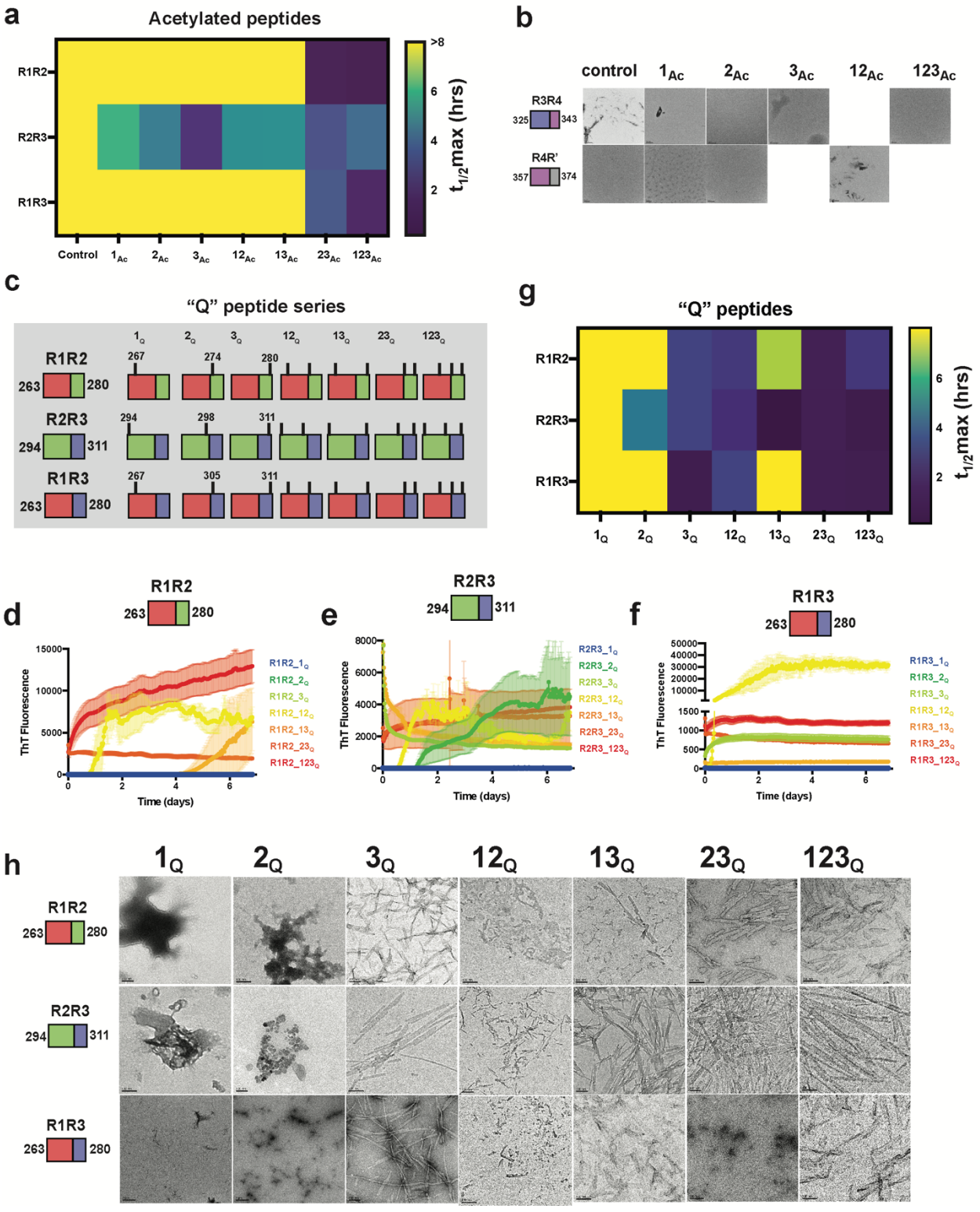
Data collection	
Microscope	Titan Krios
Acceleration Voltage (kV)	300
Detector	K3
Software	SerialEM 3.8
Magnification	105,000x
Pixel size at detector (Å/px)	0.83
Defocus range (μm)	-1.20 to -2.40
Total dose (e^-/ Å^2)	52
Exposure time (sec)	4.50
Number of movie frames	50
Reconstruction	
Usable micrograph	5901
Box size (pixel)	180
Total extracted segments	1,696,844
Number of segments after 2D	125,446
Number of segments after 3D	45,674
Symmetry imposed	C1
Helical rise (Å)	4.75
Helical twist (°)	-1.00
Crossover length (Å)	855
Map sharpening B-factor (Å^2)	-143.25
Map resolution (Å; FSC=0.143)	3.88
Model composition	
Non-hydrogen atoms	4,096
Protein residues	504
Number of chains	4.00
Water	0.00
Ligands (Acetylated)	8.00
Model validation	
Map CC (mask)	0.74
MolProbity score	2.14
Clash score	25.23
R.M.S deviations bonds (Å)	0.007
R.M.S deviations angle (°)	1.291
Rotamer outliers (%)	0.00
Ramachandran plot (favored/allowed/outliers)	(96.30/3.70/0.00)
Cβ outliers	0.00
CaBLAM outliers (%)	2.56%

Supplementary Table 2. Data collection and refinement statistics.

Supplementary Figures and Legends

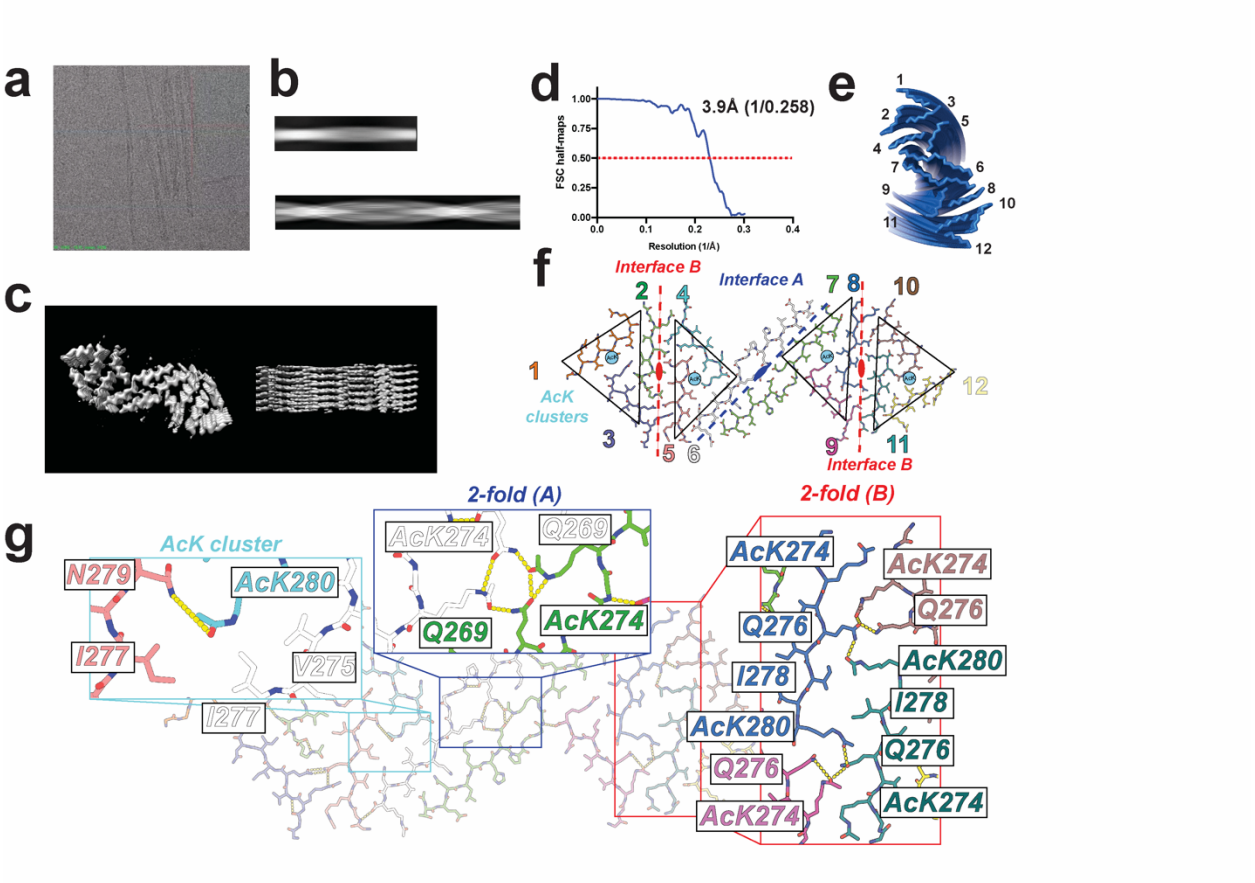


Supplementary Figure 1. Validation of chemical modification of the R1R2, R2R3, R3R4 and R1R3 peptides. Mass spectrometry analysis of chemically acetylated R1R2 (**a**), R1R3 (**b**), R2R3 (**c**) and R3R4 (**d**) peptides. The masses for each peptide species are indicated with annotation of how many lysines are modified. **e.** Estimation of $t_{1/2max}$ from the ThT fluorescence aggregation curves for chemically acetylated vs control peptides shown in Fig. 1d. For peptides that remained flat over the 8 days experiment we estimate the $t_{1/2max} > 8$ (yellow). For peptides that aggregated within 8 days, the values are colored from blue to cyan. Aggregation experiments were performed in triplicate. The data were fit to a non-linear regression model fitting in GraphPad Prism to estimate an average $t_{1/2max}$ with a standard deviation. Error bars represent the standard deviations.

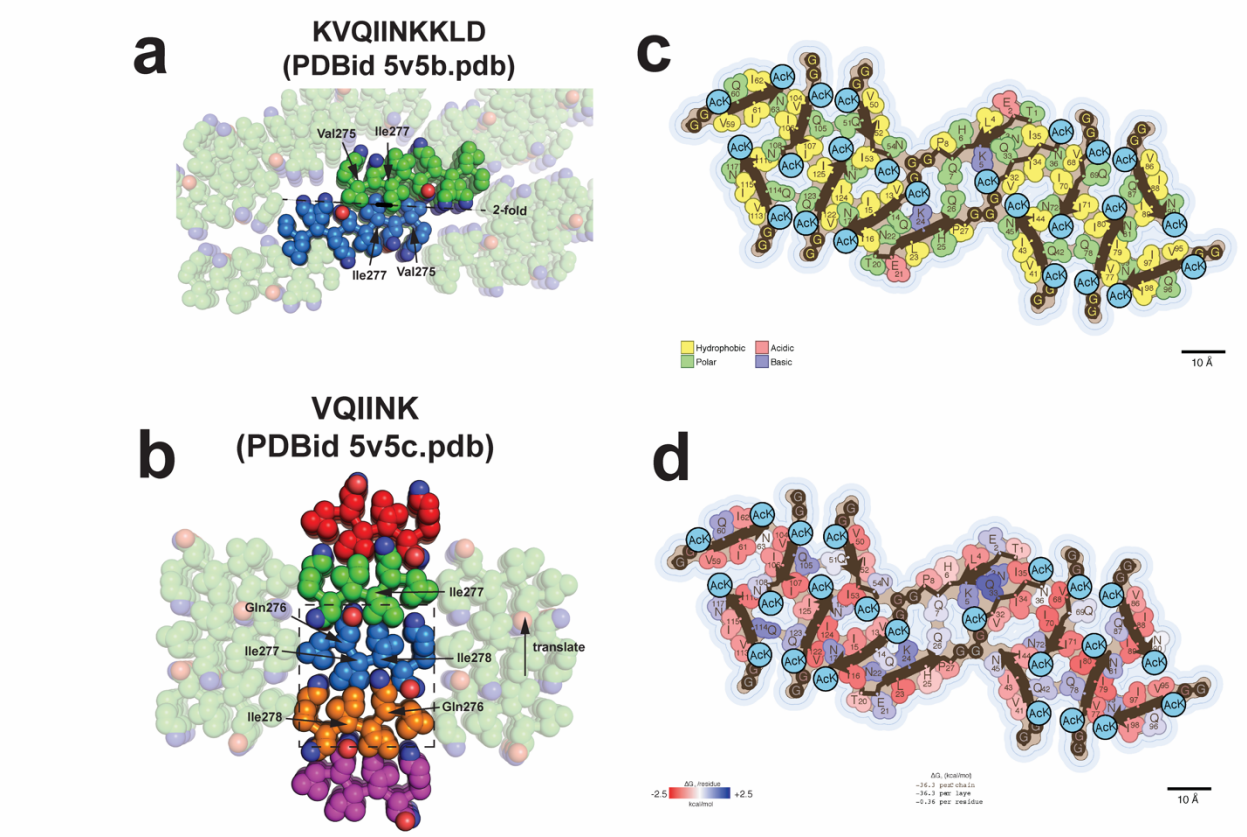


Supplementary Figure 2. Patterns of acetylation and glutamine mutation that drive model peptide aggregation. a. Estimation of $t_{1/2max}$ from the ThT fluorescence aggregation curves for unmodified, mono-, di- and tri-acetylated R1R2, R2R3, R3R4 and R1R3 peptides in Fig. 2c-f. For peptides that remained flat over the 8 days experiment we estimate the $t_{1/2max} > 8$ (yellow).

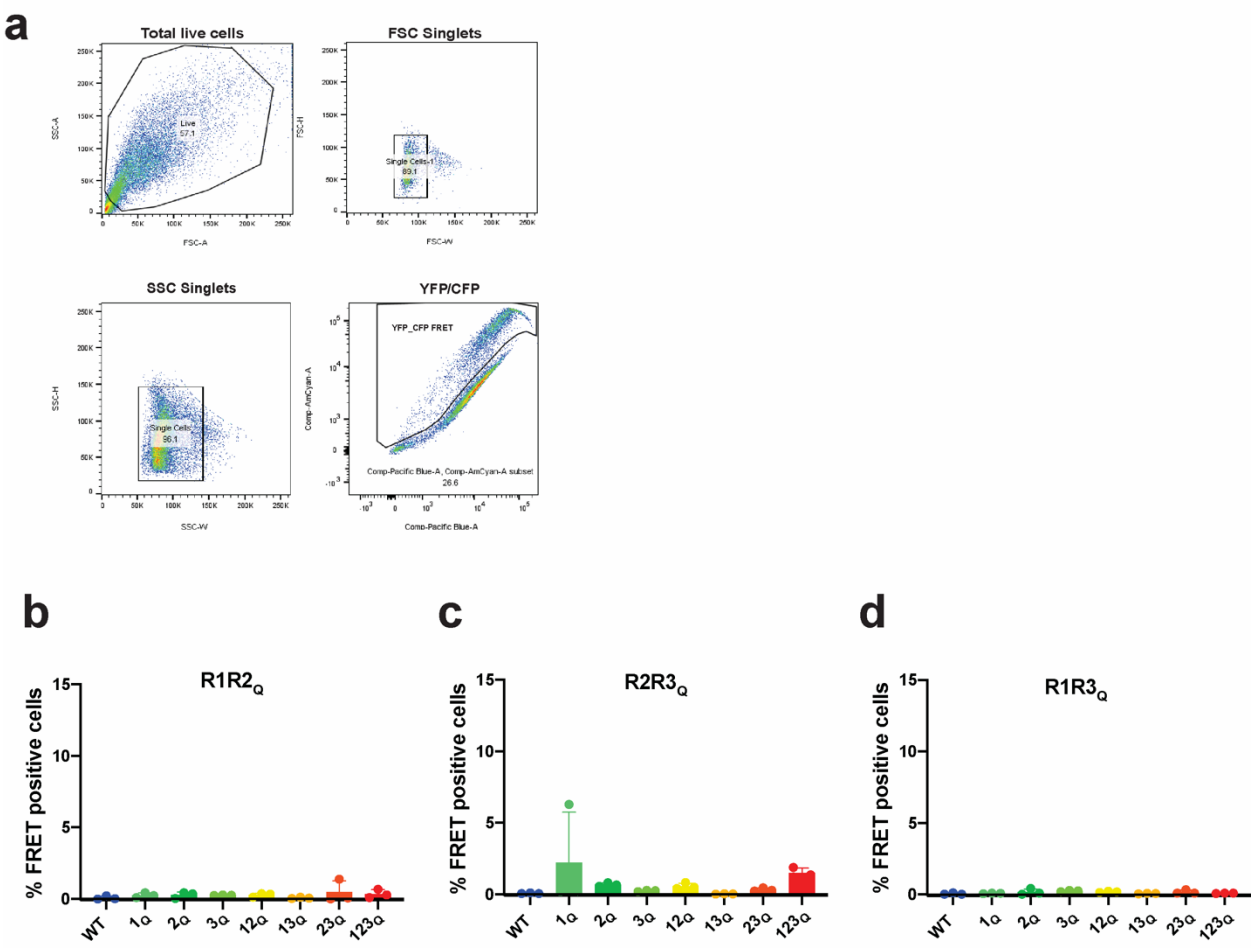
For peptides that aggregated within 8 days, the values are colored from blue to cyan. Aggregation experiments were performed in triplicate. The data were fit to a non-linear regression model fitting in GraphPad Prism to estimate an average $t_{1/2max}$ with a standard deviation. Error bars represent the standard deviations. **b**. TEM images of ThT fluorescence aggregation assay end products from control (WT) and the R3R4 and R4R' acetylated peptides from aggregation experiments shown in Fig. 2f. Scale bars indicate 0.5-1μm. **c**. Illustration of the tau peptide series with all combinatorial glutamine mutations at lysine sites for the R1R2, R2R3 and R1R3 tau peptides. Sequences are colored by repeat domain as in Fig. 1a. Acetylation sites are indicated by ticks above the cartoon for each peptide. ThT fluorescence aggregation experiments of the R1R2 (**d**), R2R3 (**e**) and R1R3 (**f**) unmodified and glutamine mutated series. Curves are colored by the number of modifications from blue (control) to red (all lysines mutated to glutamine modified). Aggregation experiments were performed in triplicate and the averages are shown with standard deviation. The data were fit to a non-linear regression model fitting in GraphPad Prism to estimate an average $t_{1/2max}$ with a standard deviation. **g**. TEM images of ThT fluorescence aggregation assay end products from control and glutamine substituted R1R2, R2R3 and R1R3 peptides from aggregation experiments shown in Supplementary Fig. 2d-f. Scale bars indicate 0.1 μm.



Supplementary Figure 3. Map quality for the R1R2 23$_{Ac}$ fibril structure. a. View of the R1R2 23$_{Ac}$ fibril on a cryo-EM grid. B. 2D projection of the R1R2 23$_{Ac}$ fibril. **c.** 3D map of the R1R2 23$_{Ac}$ fibril contoured at 7 σ**. e.** Fourier Shell Correlation (FSC) curve for refined half-maps of R1R2 23$_{Ac}$. **d.** Illustration of the arrangement of the 12 fragments in a single layer. **f.** Visualization of the symmetric interactions observed within a layer of our fibril assembly comprised of 12 independent chains from 1 to 12. The central interface formed between two extended fragments is annotated as “Interface A” and is defined a 2-fold symmetry axis between monomers 6-7 (blue). “Interface B” is defined by a pseudo 2-fold symmetric interactions between K(Ac)VQIINK(Ac) and K(Ac)VQIINK(Ac) from monomers 2-5 and 8-11 (red). The “interface B” interaction is flanked by a triangular interaction centered on a lysine acetylation site, termed “AcK cluster” formed between monomers 1-2-3 and 10-11-12. Structure is shown as sticks with all-atom and colored by fragment. **g.** illustration of the interactions at the different symmetric interfaces. Interface A is stabilized by three hydrogen bonds formed between two acetylated lysines at 274 and two glutamine 276 within a layer. Interface B is stabilized by symmetric interactions between K(Ac)VQIINK(Ac): K(Ac)VQIINK(Ac) mediated between I278 and I278 nonpolar contacts as as well as hydrogen bonds between acetylated K274 and Q276 from two different chains. AcK cluster is stabilized by interactions between three fragments involving nonpolar contacts between I277, I277, V275 and acetylated K280. Structure is shown as sticks with backbone and side chains.



Supplementary Figure 4. Key stabilizing interactions in R1R2 23$_{Ac}$ are observed in structures of amyloid motif fragments. a. 2-fold symmetric interactions observed in X-ray structures of KVQIINKKLD (PDB id 5v5b.pdb) stabilized by interactions between V275 and I277. Structures are shown in spheres (all-atom) representation and the key interactions are highlighted. The central dimer is highlighted and the peripheral interactions are shown with transparency. **b.** X-ray structures of VQIINK (PDB id 5v5c.pdb) reveal a screw axis core dimer interaction (dashed box) that is reproduced via translation in the lattice. This interaction is stabilized by contacts between I278 and I278 with interlayer Q279 hydrogen bonding. The core of the interactions are shown as spheres (all-atoms) and colored by chain. Peripheral interactions are shown with transparency. **c**. Single layer view of the R1R2 23Ac structure highlighting the amino acid property composition across the 12 fragments. Nonpolar, polar, acidic, basic, and acetyl-lysine are colored yellow, green, red, blue and baby blue, respectively. **d**. Solvation energy estimation of per residue contribution to the stability of the monolayer. Destabilizing residues are colored in blue (+2.5 kcal/mol) and stabilizing residues are colored red (-2.5 kcal/mol).



Supplementary Figure 5. Cell-based evaluation of acetylated or glutamine substituted tau peptides. a. Flow cytometry strategy to quantify FRET between intracellular mClover3 and mCerulean tau aggregates. Quantification of FRET signal across control and glutamine substituted "Q" peptides transduced into the tau biosensor cells: R1R2 (**b**), R2R3 (**c**) and R1R3 (**d**). Bar plots are colored as in Supplementary Fig. 2. Data is shown as averages across three experiments with error bars representing a 95% CI of each condition.