Supporting Information:

Spontaneous Structural Transition in Phospholipid-Inspired Aromatic Phosphopeptide Nanostructures

Michal Pellach[†], Yoav Atsmon-Raz[§,‡], Eyal Simonovsky[§,‡], Hugo Gottlieb[⊥], Guy Jacoby[♣], Roy Beck[♣], Lihi Adler-Abramovich[†], Yifat Miller[§,‡,], Ehud Gazit[†,#,*]*

[†]Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[§]Department of Chemistry, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel

[♣]The Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel

[‡]Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel

[⊥]Department of Chemistry, Bar-Ilan University, Ramat Gan 52900, Israel

[#]Department of Materials Science and Engineering Iby and Aladar Fleischman Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel

Table S1. NMR Data, L-phosphopeptide – TFE-D2, 313 K

H	δ	interresidue NOE interactions[a]
CH_3CO	1.88	Phe-1-NH
Phe-1-NH	6.79	Phe-2-NH
Phe-1-α	4.55	
Phe-1-β [b]	2.94, 3.06	
Phe-2-NH	6.99	
Phe-2-α	4.63	Val-3-NH
Phe-2-β [b]	3.01, 3.11	Val-3-NH
Val-3-NH	7.13	Lys-4-δ
Val-3-α	4.11	Lys-4-NH
Val-3-β	2.05	
Val-3-γ	0.92, 0.94	Lys-4-NH
Lys-4-NH	7.64	Pro-5-β
Lys-4-α	4.79	
Lys-4-β	1.86	
Lys-4-γ	2.17	
Lys-4-δ	1.56	
Lys-4-ε	2.95, 3.10	
Lys-4-NH_3^+	7.17	
(L)Pro-5-α	4.26	Val-8-γ
(L)Pro-5-β	2.02, 2.43	pSer-6-NH
(L)Pro-5-γ	2.21, 2.40	pSer-6-NH
(L)Pro-5-δ	3.84	

H	δ	interresidue NOE interactions[a]
pSer-6-NH	7.86	Val-8-α
pSer-6-α	4.30	
pSer-6-β	4.40	
Glu-7-NH	7.75	Val-8-NH
Glu-7-α	4.35	
Glu-7-β	2.26, 2.43	
Glu-7-γ	2.51, 2.62	
Val-8-NH	7.31	
Val-8-α	3.85	Phe-9-NH
Val-8-β	2.02	Phe-9-NH, Phe9-*ortho*
Val-8-γ	0.72, 0.86	
Phe-9-NH	7.57	
Phe-9-α	4.47	
Phe-9-β [b]	2.78, 2.95	
Phe-10-NH	7.48	
Phe-10-α	4.71	$CONH_2$
Phe-10-β [b]	3.12, 3.32	
$CONH_2$	6.08, 6.90	

a. NOE interactions only presented ***once*** (forward into the chain).
b. *Ortho* protons: 7.55 (Phe-1); 7.13 (Phe-2); 7.07 (Phe-9); 7.34 (Phe-10)

Table S2. NMR Data, *D-phosphopeptide* – TFE-D_2, 313 K

H	δ	interresidue NOE interactions[a]
CH_3CO	1.90	Phe-1-NH
Phe-1-NH	6.79	Phe-2-NH
Phe-1-α	4.95	
Phe-1-β	2.92, 3.03	
Phe-2-NH	8.18	
Phe-2-α	4.74	
Phe-2-β	2.95, 3.04	
Val-3-NH	7.10	Lys-4-δ
Val-3-α	4.63	Lys-4-NH
Val-3-β	2.01	
Val-3-γ	0.95, 0.96	Lys-4-NH, Pro5-δ
Lys-4-NH	8.60	Glu-7-β
Lys-4-α	4.83	Pro-5-δ
Lys-4-β	2.00, 2.03	
Lys-4-γ	160, 1.33	
Lys-4-δ	1.75	
Lys-4-ε	3.04, 3.11	
Lys-4-NH_3^+	?	
(D)Pro-5-α	4.41	
(D)Pro-5-β	2.11, 2.28	
(D)Pro-5-γ	2.02, 2.23	
(D)Pro-5-δ	3.68	

H	δ	interresidue NOE interactions[a]
pSer6-NH	8.85	
pSer6-α	4.30	Glu-7-NH
pSer6-β	4.48	
Glu-7-NH	8.18	Val-8-NH
Glu-7-α	4.41	
Glu-7-β	2.08, 2.32	
Glu-7-γ	2.54, 2.70	
Val-8-NH	7.19	
Val-8-α	4.47	Phe-9-NH
Val-8-β	1.90	Phe-9-NH
Val-8-γ	0.78, 0.86	Phe-9-NH
Phe-9-NH	8.37	
Phe-9-α	4.68	Phe-10-NH
Phe-9-β	2.81, 3.03	Phe-10-NH
Phe-10-NH	7.42	
Phe-10-α	4.86	
Phe-10-β	2.99, 3.19	
$CONH_2$	6.00, 7.17	

a. NOE interactions only presented ***once*** (forward into the chain).

Table S3. NMR Data, *L-phosphopeptide* – DMSO-D_6, 313 K

H	δ [b]	interresidue NOE interactions [a]
CH_3CO	1.71	Phe-1-NH
Phe-1-NH	7.92	
Phe-1-α	4.46	
Phe-1-β	2.68, 3.92	
Phe-2-NH	8.01	
Phe-2-α	4.58	
Phe-2-β	2.83, 3.03	Val-3-NH
Val-3-NH	7.72	Lys-4-δ
Val-3-α	4.18	Lys-4-NH
Val-3-β	1.95	Lys-4-NH
Val-3-γ	0.83, 0.84	Lys-4-NH, Lys4-α
Lys-4-NH	8.09	
Lys-4-α	4.57	Pro-5-δ
Lys-4-β	1.47	
Lys-4-γ	1.95,1.3-1.6	Pro-5-γ
Lys-4-δ	1.3-1.6	Pro-5-γ
Lys-4-ε	*ca*. 2.75	
Lys-4-NH_3^+		
(L)Pro-5-α	4.19	
(L)Pro-5-β	2.12	
(L)Pro-5-γ	1.81, 1.96	
(L)Pro-5-δ	3.61, 3.77	

H	δ [b]	interresidue NOE interactions [a]
pSer-6-NH		
pSer-6-α	4.04	Glu-7-NH
pSer-6-β	4.23	
Glu-7-NH	7.67	
Glu-7-α	4.27	Val-8-NH
Glu-7-β	1.81, 2.03	
Glu-7-γ	2.31	
Val-8-NH	7.39	
Val-8-α	4.00	
Val-8-β	1.91	Phe-9-NH
Val-8-γ	0.69, 0.71	Phe-9-NH
Phe-9-NH	7.80	
Phe-9-α	4.41	$CONH_2$
Phe-9-β	2.84, 3.00	
Phe-10-NH		
Phe-10-α	4.45	
Phe-10-β	2.74	
$CONH_2$	6.98, 7.05	

a. NOE interactions only presented ***once*** (forward into the chain).
b. Some chemical shifts could not be detected, in most cases due to broadening of signals.

Table S4. NMR Data, *D-phosphopeptide* – DMSO-D_6, 318 K – "80%" conformer

H	δ [b]	interresidue NOE interactions[a]
CH_3CO	1.71	
Phe-1-NH		
Phe-1-α		
Phe-1-β		
Phe-2-NH	8.45	
Phe-2-α	4.78	Val-3-NH, Val3-γ
Phe-2-β	2.75, 2.93	
Val-3-NH	8.37	
Val-3-α	4.69	
Val-3-β	1.95	
Val-3-γ	0.85, 0.86	Pro-5-γ
Lys-4-NH		
Lys-4-α	4.70	Pro-5-δ
Lys-4-β	1.55, 1.77	Pro-5-δ
Lys-4-γ		
Lys-4-δ		
Lys-4-ε	2.71	
Lys-4-NH_3^+	7.58	
(D)Pro-5-α	4.30	
(D)Pro-5-β	1.73, 2.22	
(D)Pro-5-γ	1.89, 2.09	
(D)Pro-5-δ	3.61	

H	δ [b]	interresidue NOE interactions[a]
pSer-6-NH		
pSer-6-α	4.04	
pSer-6-β	4.45	
Glu-7-NH	8.28	
Glu-7-α	4.67	
Glu-7-β	2.3 – 2.8	
Glu-7-γ	2.3 – 2.8	
Val-8-NH	8.00	
Val-8-α	4.42	
Val-8-β	1.71	Phe-9-NH
Val-8-γ	0.68, 0.70	Phe-9-NH
Phe-9-NH		
Phe-9-α		
Phe-9-β		
Phe-10-NH		
Phe-10-α		
Phe-10-β		
$CONH_2$	6.93, 7.26	

a. NOE interactions only presented ***once*** (forward into the chain).
b. Several chemical shifts could not be detected, in most cases due to broadening of signals.

Table S5. NMR Data, *D-phosphopeptide* – DMSO-D_6, 318 K – "20%" conformer

H	δ[b]	interresidue NOE interactions[a]
CH_3CO		
Phe-1-NH		
Phe-1-α		
Phe-1-β		
Phe-2-NH		
Phe-2-α	4.57	Val-3-NH
Phe-2-β		
Val-3-NH	7.73	
Val-3-α	4.21	
Val-3-β	1.94	
Val-3-γ	0.81, 0.83	
Lys-4-NH	8.13	
Lys-4-α	4.19	
Lys-4-β	1.3 – 1.8	
Lys-4-γ	1.3 – 1.8	
Lys-4-δ	1.3 – 1.8	
Lys-4-ε	2.74	
Lys-4-NH_3^+		
(D)Pro-5-α		
(D)Pro-5-β		
(D)Pro-5-γ		
(D)Pro-5-δ		

H	δ[b]	interresidue NOE interactions[a]
pSer-6-NH		
pSer-6-α		
pSer-6-β		
Glu-7-NH	7.93	
Glu-7-α	4.31	Val-8-α
Glu-7-β	2.2 – 2.9	
Glu-7-γ	2.2 – 2.9	
Val-8-NH	7.82	
Val-8-α	4.07	
Val-8-β	1.89	
Val-8-γ	0.65, 0.73	
Phe-9-NH		
Phe-9-α		
Phe-9-β		
Phe-10-NH		
Phe-10-α		
Phe-10-β		
$CONH_2$		

a. NOE interactions only presented ***once*** (forward into the chain).
b. Many chemical shifts could not be detected, in most cases due to broadening of signals.

Figure S1

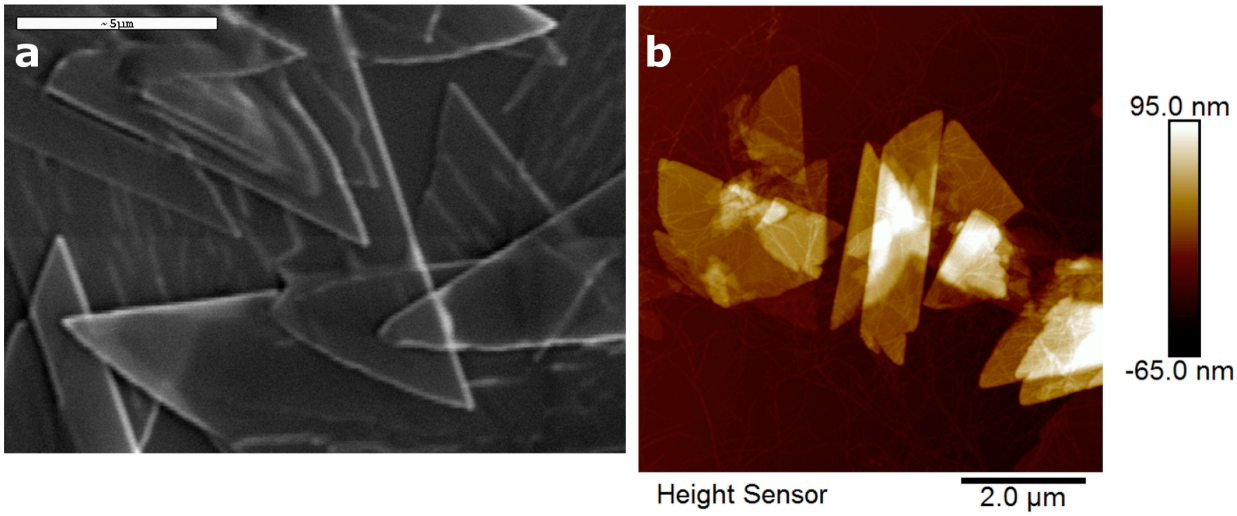


Figure S1. SEM image (a) and AFM image (b) of D-phosphopeptide self-assembled half-elliptical plates. The plates appear to have fabric-like flexibility, adopting the shape of the structures beneath, and a thickness of approximately 10 ± 2 nm. The stacking of the structures made precise measurements of thickness difficult. There is possible variation of thickness within each plate, which can be explained by the MD simulations, if "slices" of the fibrils are not identical in length.

Figure S2

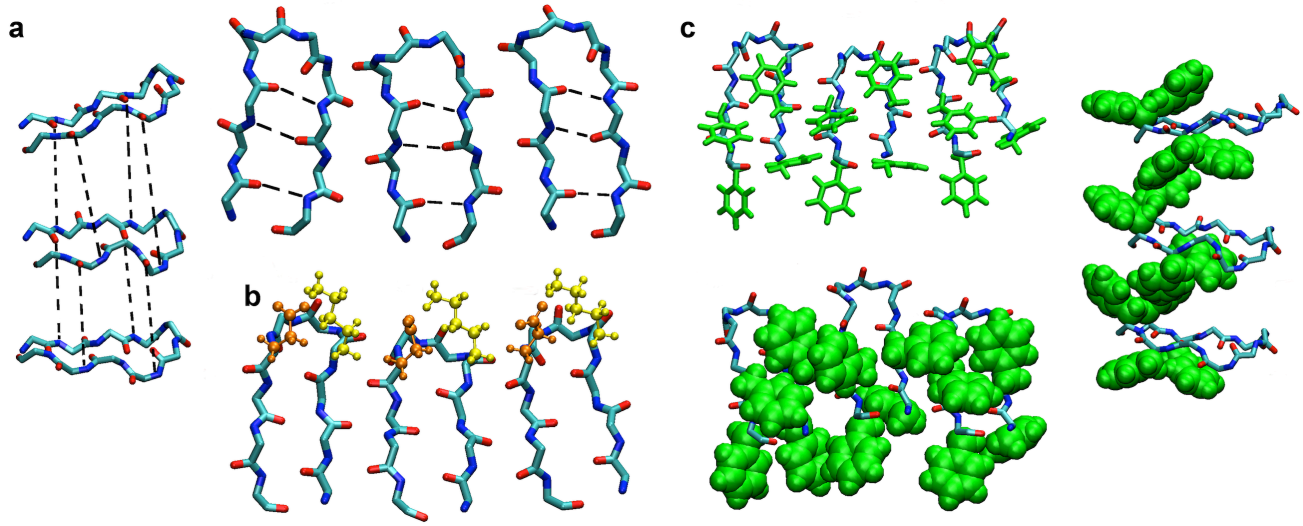


Figure S2. Representation of intra- and intermolecular interactions formed by the D-phosphopeptide (Model A). (a) Hydrogen bonding interactions (b) Salt-brigde interactions between Lys-4 and Glu-7 (c) aromatic interactions. These interactions stabilize the intermediate fibrillar structures, and most likely also the final supramolecular structures.

Figure S3

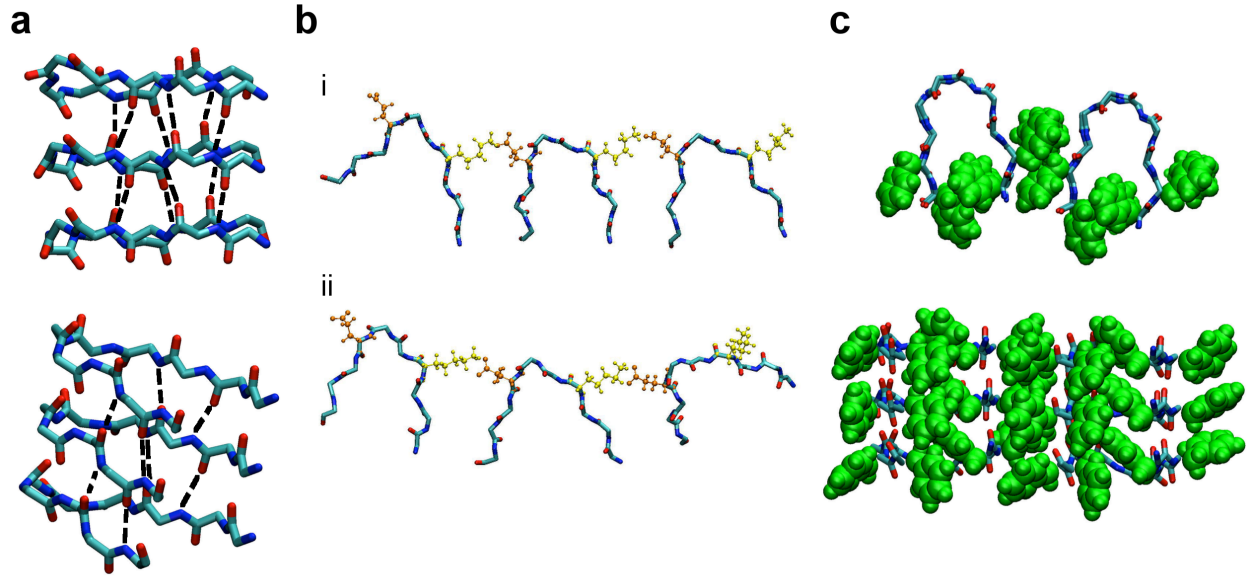


Figure S3. Representation of intra- and intermolecular interactions formed by the L-phosphopeptide (Model B). (a) Hydrogen bonding interactions (b) Salt-brigde interactions between Lys-4 and Glu-7, when pSer faces inwards of the tubular fibril structures (i) and with pSer facing outwards of the tubular fibril structures (ii). (c) aromatic interactions. These interactions stabilize the intermediate fibrillar structures, and most likely also the final supramolecular structures.

Figure S4

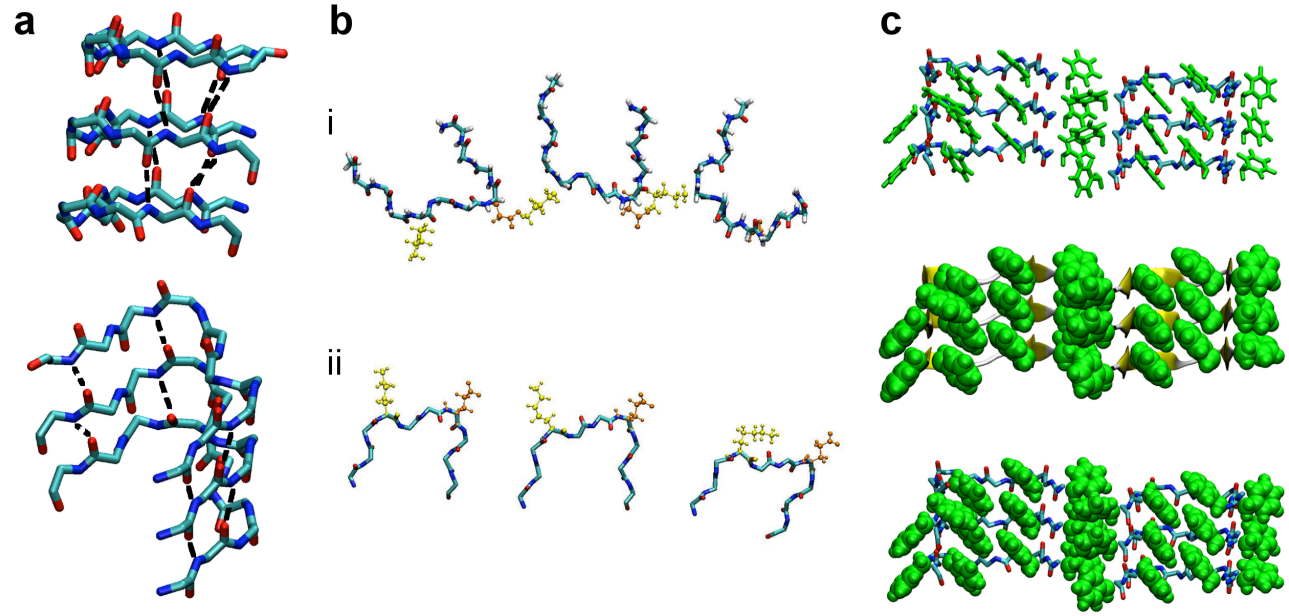


Figure S4. Representation of selected intra- and intermolecular interactions formed by the L-phosphopeptide (Model C). (a) Hydrogen bonding interactions (b) Salt-brigde interactions between Lys-4 and Glu-7, when pSer faces inwards of the tubular fibril structures (i) and with pSer facing outwards of the tubular fibril structures (ii). (c) aromatic interactions. These interactions stabilize the intermediate fibrillar structures, and most likely also the final supramolecular structures.

Figure S5

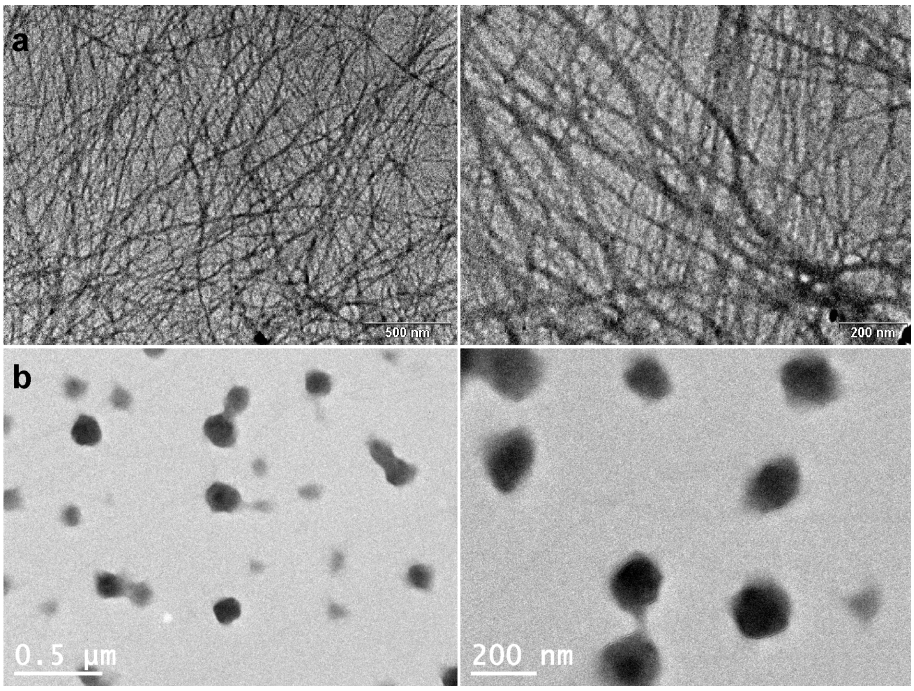


Figure S5. TEM images from non-phosphorylated analogues of the phosphopeptides: (a) "D-peptide" and (b) "L-peptide". Under the same self-assembling conditions as the phosphopeptides (solvent switch from HFIP to water), the D-peptide analog forms a gel in water, as a result of a dense mesh of fine fibers. The L-peptide forms nanoparticles with undefined boundaries.

Computation Details

1. Construction of the 'building block' monomeric models of L-phosphopeptide and D-phosphopeptide

One of the most challenging issues when investigating the self-assembly of peptides at the atomic resolution is to predict the 'building block' structural model of the monomeric peptide. In the current study, the CD measurements have shown for both L-phosphopeptide and D-phosphopeptide that the monomeric peptides have a secondary structure of β-sheet. Furthermore, the NMR experiments indicated that these two peptides are folded; suggesting that the 'building block' of the peptides is not a secondary structure of one long β-strand. We therefore proposed that each of the monomeric 'building blocks' L-phosphopeptide and D-phosphopeptide will be folded into β-sheet-turn-β-sheet.

1.1. Construction of the 'building block' D-phosphopeptide

According to the CD spectrum of the D-phosphopeptide the shape of the spectrum indicates that this peptide has a secondary structure of β-hairpin with β-turn type II'. This is supported by previous study[1]. Furthermore, Schneider and coworkers proposed such a secondary structure for the MAX1 peptide in which the DPro is located in a β-turn type II'.[2] We therefore modelled the monomeric D-phosphopeptide Model A as a β-hairpin, in which the residues DPro, pSer and Glu are located in the turn domain, while the hydrophobic residues are located in the two β-strands of the β-hairpin structure (Fig. 2b in main text).

1.2. Construction of the 'building block' L-phosphopeptide

Since the self-assembly of the L-phosphopeptide and D-phosphopeptide into the supra-structure yield to two different morphologies, we suggest that the monomeric L-phosphopeptide and D-phosphopeptide 'building blocks' also structurally differ in the folded state. The β-hairpin secondary structures are organized as β-sheet-turn-β-sheet, but also β-arch secondary structures are organized as β-sheet-turn-β-sheet. The common structure of β-arch appears in amyloids, such as Aβ amyloid (a peptide that is related to Alzheimer's disease). Two ssNMR structural Aβ amyloid models had been proposed that are mainly differ in the turn shape. First model proposed by Tycko and coworkers and the second by Lührs et al. In the current study, we suggest to superimpose the residues of the L-phosphopeptide in the backbone of each one of the structural Aβ amyloid models. In Model B, we superimposed the residues in which Lys4-LPro5-pSer6-Glu7 are located in the turn domain of the β-arch of Tycko's Aβ backbone model, while in Model C residues Lys4-DPro5-pSer6-Glu7 were superimposed in the turn domain of the β-arch of Lührs' model (Figure 2c in main text). Therefore in each model B and C, the turn domain consists of these four residues and the two β-strands consist of the hydrophobic residues.

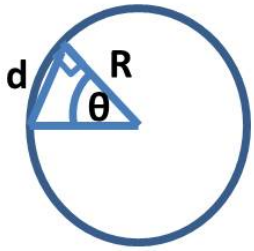
2. Construction of the self-assembly models of L-phosphopeptide and D-phosphopeptide

Since the sizes of the supra-structures of these two phosphopeptides are micrometers (as seen in the TEM), it is beyond the modelling tools to model such huge systems and therefore cannot been performed. We therefore proposed to model the self-assembled fibrilar states, which illustrated in the TEM sizes in the range of 13-17 nm. To model the fibril-like tubular structures for each one of models A, B and C, we considered two tubular conformations: in the first conformation the turn regions (or the pSer) are oriented facing outwards, and in the second conformation the turn regions (or the pSer) are oriented facing inwards.

We therefore, constructed two conformations of model A: A1 and A2. For model B we constructed models B1 and B2 and for model C we constructed models C1 and C2.

To estimate the numbers of the monomeric 'building blocks' for each model A, B and C in each layer of the tubular models, we applied the following geometrical cycle equations:

$$\theta = \arcsin\left(\frac{d}{R}\right)$$



$$\frac{2\pi}{\theta} \approx n$$

While R is the radius of the tubular structure, which is defined by the distance between the edges of N- or the C- termini of the peptides and the centre of the tubular structure. For the two tubular structures of model A the distance *d* is defined as the distance between the edge atom in the C-terminal of peptide one to the edge atom in the C-terminal of the neighboured peptide. For the four tubular structures of models B and C the distance *d* is defined as the distance between edge atom in the C-terminal and the edge atom in the N-terminal of the same peptide. Finally, θ is the rotation angle between the atom in the edge of the C-terminal and the atom in the edge of the N-terminal of the peptide.

Applying these equations, led us to propose that in each layer of the tubular model, 38 monomeric 'building blocks' can be accommodated by model A to form a tubular structure with a diameter of 15 nm (i.e. R=7.5 nm), 17 monomeric 'building blocks' by model B and 21 monomers by model C to form tubular structures of 13 nm (i.e. R=6.5 nm).

To form fibril-like structures, we constructed four layers for models A1 and A2, considering a distance of 10 Å between each two layers, as previously proposed for the β-hairpin MAX1 bilayer amphiphilic peptide. The total number of the monomeric 'building blocks' of four layers in these models is 152. For models B1, B2, C1 and C2, we constructed six layers, considering a distance of 5 Å between each two layers, as previously proposed by experimental models for amyloids, such as Aβ and amylin. The total number of the monomeric 'building blocks' of six layers in model B1/B2 and C1/C2 are 126 and 102, respectively.

3. Molecular dynamics (MD) simulations procedure

MD simulations of the solvated variant models were performed in NPT ensemble using the NAMD program[3] with the CHARMM36 force-field[4-6] with CMAP corrections for 80 ns. The parameters for the phosphoSerine residues were obtained from the MacKerell lab.[7] We integrated these parameters to the force-field potential. The models were explicitly solvated with TIP3P water molecules.[8,9] The Langevin piston method,[3,10,11] with a decay period of 100 fs and a damping time of 50 fs, was used to maintain a constant pressure of 1 atm. The temperature (330 K) was controlled by Langevin thermostat with a damping coefficient of 10 ps^{-1}.[3] The short-range van der Waals (VDW) interactions were calculated using the switching function, with a twin range cutoff of 10.0 and 12.0 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a cut-off of 12.0 Å for all simulations.[12,13] The equations of motion were integrated using the leapfrog integrator with a step of 1 fs. All initial variant models were energy minimized and then solvated in a TIP3P water box with a minimum distance of 15 Å from any edge of the box to any peptide atom. Any water molecule within 2.5 Å of the Aβ was removed. Counterions (Na^+) were added at random locations to neutralize the self-assembled peptides' charge.

The solvated systems were energy minimized for 2000 conjugated gradient steps, where the distance between the β-sheets in the peptides is fixed in the range 2.2 – 2.5 Å. The counterions and water molecules were allowed to move. The hydrogen atoms were constrained to the equilibrium bond using the SHAKE algorithm.[14] The minimized solvated systems were heated at 200 K, where all atoms were allowed to move. Then, the systems were heated from 150 K to 200 K for 300 ps and equilibrated at 310 K for 300 ps. All simulations ran for 80 ns and structures were saved every 10 ps for analysis. These conditions (310 K and 80 ns of simulation time) are applied to test the stabilities of all variant models. A recent study with similar size of self-assembled peptide system had been investigated by MD simulations with 310 K and 40 ns.[15] In this work we extended the timescale to 80 ns.

References:

1. Xiao, J.; Weisblum, B.; Wipf, P. Trisubstituted (E)-Alkene Dipeptide Isosteres as β-Turn Promoters in the Gramicidin S Cyclodecapeptide Scaffold. *Org. Lett.* 2006, 8, 4731.

2. Haines-Butterick, L.; Rajagopal, K.; Branco, M.; Salick, D.; Rughani, R.; Pilarz, M.; Lamm, M. S.; Pochan, D. J.; Schneider, J. P. Controlling Hydrogelation Kinetics by Peptide Design for Three-Dimensional Encapsulation and Injectable Delivery of Cells. *Proc. Natl. Acad. Sci. U.S.A.* 2007, 104, 7791.

3. Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* 1999, 151, 283-312.

4. Best, R. B.; Buchete, N. V.; Hummer, G. Are Current Molecular Dynamics Force Fields Too Helical? *Biophys. J.* 2008, 95, L07-9.

5. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* 2004, 126, 698-9.

6. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.*, et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 1998, 102, 3586-3616.

7. Feng, M.-H.; Philippopoulos, M.; MacKerell, A. D.; Lim, C. Structural Characterization of the Phosphotyrosine Binding Region of a High-Affinity SH2 Domain–Phosphopeptide Complex by Molecular Dynamics Simulation and Chemical Shift Calculations. *J. Am. Chem. Soc.* 1996, 118, 11265-11277.

8. Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* 2000, 112, 8910-8922.

9. 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, 926-935.

10. Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant-Pressure Molecular-Dynamics Algorithms. *J. Chem. Phys.* 1994, 101, 4177-4189.

11. Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *J. Chem. Phys.* 1995, 103, 4613-4621.

12. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. *The Journal of Chemical Physics* 1993, 98, 10089-10092.

13. 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, 8577-8593.

14. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* 1977, 23, 327-341.

15. Lee, O.-S.; Stupp, S. I.; Schatz, G. C. Atomistic Molecular Dynamics Simulations of Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers. *J. Am. Chem. Soc.* 2011, 133, 3677-3683.