Supporting Information

Molecular Insight into the β-Sheet Twist and Related Morphology of Self-Assembled Peptide Amphiphile Ribbons

Qinsi Xiong[1], Samuel I. Stupp[1], George C. Schatz[1*]

[1] Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3113, USA.

* Correspondence: g-schatz@northwestern.edu

COMPUTATIONAL METHOD

Models and Simulation Setup

We attached peptide amphiphiles consisting of sequences varying from PA1 to PA6 (Figure 1a and Figure S1) to a 16-carbon palmitoyl tail at the N-terminus; all amino acids were protonated to simulate an acidic environment. Then we constructed the initial bilayer ribbon based on our previous work[1] by replicating a dimer building block along the x and y axis respectively, with a distance of 4 Å in the x direction and 8 Å in the y direction, respectively (Figure 1c). We also constructed a cylindrical micelle structure for PA3, which is based on our previous work[2] by radially placing ten PA chains per layer and stacking 16 layers with a distance of 3.5 Å. To monitor distortion of the ribbon, periodic boundary conditions with a finite bilayer were used. We used the TIP3P water model[3] and added 0.1 M NaCl. We applied the CHARMM36 force field[4] to simulate the self-assembly of PAs using the GROMACS 2020.5 package[5] to perform the AAMD simulations. All the systems were subjected to a 5000-step energy minimization followed by a 1ns NVT ensemble with the V-rescale thermostat. PAs were positionally restrained during the energy minimization, and a 2ns NPT ensemble calculation was performed. After that, we slowly removed the constraints and conducted another 5 ns simulation at 300 K to further relax the system. The particle mesh Ewald method[6] was employed to calculate electrostatic interactions with a short-range cutoff of 1.2 nm. The boundary conditions were periodic in all three directions, the simulation time step was 2 fs. Next, we performed a production run for 100 ns at 300K and 1.0 atm. The V-rescale method and the Parrinello−Rahman method were used to maintain the simulation temperature and pressure, respectively.

Calculation Methods

Secondary structure

The β-sheets secondary structure information for the PA nanostructures was calculated using the DSSP algorithm implemented in GROMACS.

Calculation of non-covalent interactions

Since our resulting ribbons were separated from their periodic images under periodic boundary conditions (PBC), PAs on the ends of the ribbon will be exposed to solvent, which would not happen for a normal ribbon that is microns long. This will bias our calculation of the non-covalent interactions. To avoid this, we set a cut-off, that is, we only consider non-covalent interactions of PAs within the cut-off distance from the geometric center of the aggregate. The cut-off we used in this work was 3nm, 4nm, and 4nm for 8*10PA, 8*16PA, and 11*16PA systems, respectively.

Backbone's hydrogen bond (HB) 2D contact map

To generate contact maps for two neighboring monomers, a pair of NH (or OH)---CO groups (Figure S19a) was considered to have a HB if the distance between donor and acceptor atoms was <0.35 nm and the donor-hydrogen-acceptor angle was >120°, where both donor and acceptor come from backbones.
Here we take PA1 as an example, as shown in Figure S19a, with monomer1 and monomer2 being two neighboring monomers. Each PA1 has 4 amino acids but there are 5 pairs of NH (or OH)---CO groups. The contact map is therefore a 5*5 matrix, as shown on Figure S19b. The color bar is the averaged counting number of HB.

Conformationally constrained simulations

In conformationally constrained simulations, additional harmonic restraining forces were applied to keep PA4s in β-sheet conformational states:

$$E_{\text{cons}} = \sum_N \sum_{i\in\{1-7\}} \left(K_{i,N}\left(\phi_{i,N} - \phi^{\circ}_{i,N}\right)^2 + K_{\psi_{i,N}}\left(\psi_{i,N} - \psi^{\circ}_{i,N}\right)^2 \right), \text{[S1]}$$

where $K_{\phi\ i,N}$ and $\phi^{\circ}_{i,N}$ are the force constant and the equilibrium value of the harmonic force applied to the backbone dihedral $\phi_{i,N}$ of the i_{th} PA4s residue in the N_{th} monomer and, similarly, $K_{\psi_{i,N}}$ and $\psi^{\circ}_{i,N}$ are associated with the force applied to the dihedral $\psi_{i,N}$ of the same residue. To keep the PA4 in the β-sheet conformation, the backbone dihedral angles were restrained in the second quadrant of the Ramachandran plot ($-180° < \phi < 0°, 0° < \psi < 180°$). The specific ϕ and ψ distributions for each residue were obtained from the unconstrained self-assembly simulations of PA4, then we further use these distributions as a target to parameterize the $K_{\phi\ R,M}$ and $\phi^{\circ}_{R,M}$ values of PA4s; the parameterized results are shown in Table S2. Note that we only considered the first 7 amino acids, the last one (aspartic acid)'s conformation was not constrained.

Table S1. Summary of simulations conducted in this study.

PA Sequence	Ribbon type (width layer*length layer)	Numbers of PAs/particles	Box Size [a] (X*Y*Z nm^3)	Constrained Conformational state	Simulation Type (pH, temperature)	Simulation time (ns)[b]
$C_{16}(VE)_2$	8*10	160/102754	8*14.8*8.8	No	conventional MD（low pH, 300K）	RUN1:100
	8*16	256/155130	9*18.8*8.8			RUN1:100; RUN2:100
	11*16	352/207432	12*18.8*8.8			RUN1:100
$C_{16}(VE)_4$	8*10	160/110099	8*14.0*9.2			RUN1:100
	8*16	256/202172	9*18.8*11.4			RUN1:100; RUN2:110
	11*16	352/207402	10*18.8*10.2			RUN1:100
$C_{16}(VE)_6$	8*10	160/143924	8*14*12			RUN1:100
	8*16	256/306758	9*18.8*18			RUN1:100; RUN2:110
	11*16	352/356741	10*18.8*18			RUN1:100
$C_{16}(VE)_2GRGD$	8*10	160/108371	7.2*14*10.2			RUN1:100
	8*16	256/193242	9*18.8*11			RUN1:100; RUN2:100
			9*18.8*11	Yes		RUN1:100
	11*16	352/185453	9*18.8*10.2	No		RUN1:100
$C_{16}GRGD(VE)_2$	8*16	256/193449	9*18.8*11			RUN1:100; RUN2:100
$C_{16}(VE)_6GRGD$	8*16	256/352155	9*18.8*20			RUN1:100; RUN2:100
PA Sequence	Cylinder	Numbers of PAs/particles	Box Size [a] (X*Y*Z nm^3)	Constrained Conformational state	Simulation Type (temperature)	Simulation time (ns)[b]
$C_{16}(VE)_6$	10*16	160/126916	15*15*5.6	No	conventional MD（low pH, 300K）	RUN1:100
PA Sequence	Ribbon type (width layer*length layer)	Numbers of PAs/particles	Box Size [a] (X*Y*Z nm^3)	Constrained Conformational state	Simulation Type (temperature)	Simulation time (ns)[b]
$C_{16}(VE)_6$	8*10	160/142514	8*14*12	No	conventional MD（neutral pH, 300K）	RUN1:100

[a] Due to the different PA sizes, the size of the box may be different. The basic principle is to ensure at least 2.5nm, 3.0nm, 2.0nm distance from the edge of the ribbon to the boundary of the PBC box in the x, y, and z directions, respectively.

[b] Several properties, such total energy, pressure, temperature, density, torsion angle of ribbons and β-sheet content were examined throughout the simulations to make sure the structure was converged. It turns out most of the time 100ns is enough for the nanostructures to converge.

Table S2. Parameters for conformational constraints used to fix PA4 to β-sheet conformational states

	Conformational state
	β-sheet
K_{ϕ_1}**(kJ/mol/rad^2)**	60
ϕ_1° **(deg)**	-115
K_{ψ_1}**(kJ/mol/rad^2)**	300
ψ_1° **(deg)**	125
K_{ϕ_2}**(kJ/mol/rad^2)**	300
ϕ_2° **(deg)**	-125
K_{ψ_2} **(kJ/mol/rad^2)**	60
ψ_2° **(deg)**	135
K_3 **(kJ/mol/rad^2)**	70
ϕ_3° **(deg)**	-125
K_{ψ_3}**(kJ/mol/rad^2)**	70
ψ_3° **(deg)**	125
K_{ϕ_4}**(kJ/mol/rad^2)**	60
ϕ_4° **(deg)**	-125
K_{ψ_4}**(kJ/mol/rad^2)**	60
ψ_4° **(deg)**	125
K_{ϕ_5}**(kJ/mol/rad^2)**	70
ϕ_5° **(deg)**	-135
K_{ψ_5} **(kJ/mol/rad^2)**	70
ψ_5° **(deg)**	115
K_6 **(kJ/mol/rad^2)**	300
ϕ_6° **(deg)**	-125
K_{ψ_6}**(kJ/mol/rad^2)**	300
ψ_6° **(deg)**	125
K_7 **(kJ/mol/rad^2)**	300
ϕ_7° **(deg)**	-125
K_7**(kJ/mol/rad^2)**	300
ψ_7° **(deg)**	125

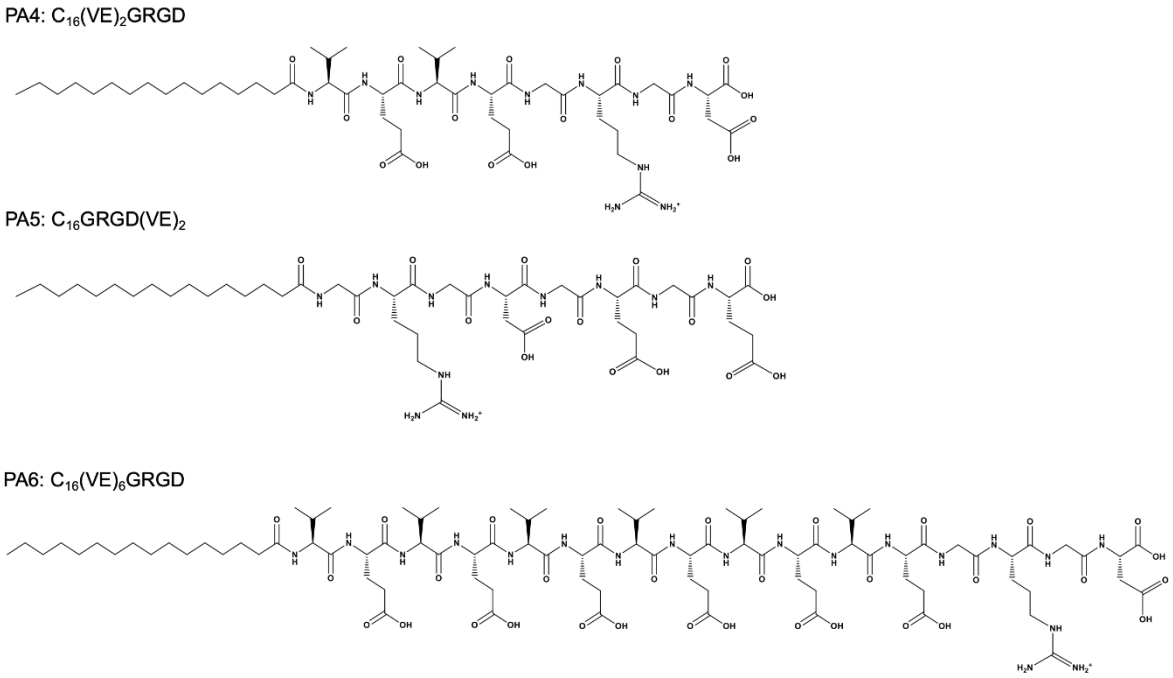

Figure S1. Chemical structures of $C_{16}(VE)_2GRGD$, $C_{16}GRGD(VE)_2$ and $C_{16}(VE)_6GRGD$ PAs.

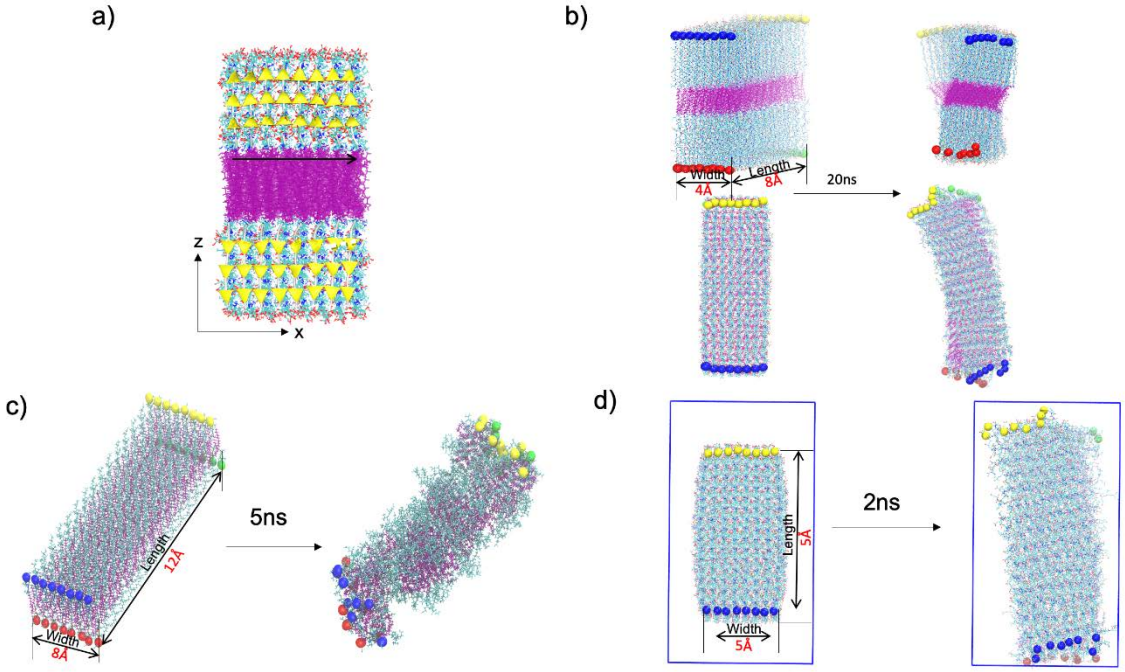

Figure S2. a) Bilayer model with hydrogen bond chain elongated in width (x, black arrow) axis. We simulated the double layers with appropriate (b), lower (c) and larger (d) PA densities respectively. We found that the appropriate density can quickly distort PA3, while the bilayer with lower density would quickly lose its stability, and the higher density system would quickly expand and twist. All systems use the 8*16 PA model. Red marked numbers show the interval distance between layers in the width and length directions.

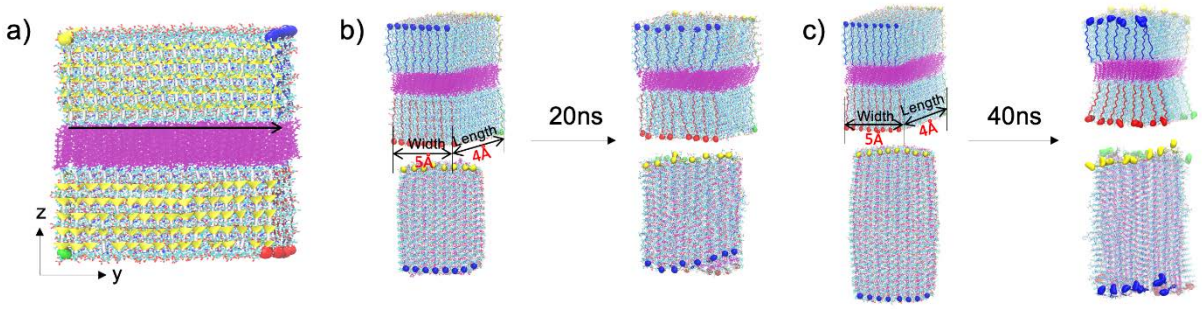

Figure S3. a) Bilayer model with hydrogen bond chain elongated in length (y, black arrow) axis. We simulated the double layers with (b)8*16PA model (8 layers in width and 16 layers in length) and (c)8*24PA model (8 layers in width and 24 layers in length) respectively. All simulations show a flat ribbon (torsion angle within ±5º). Red marked numbers give the interval distance between layers in the width and length direction.

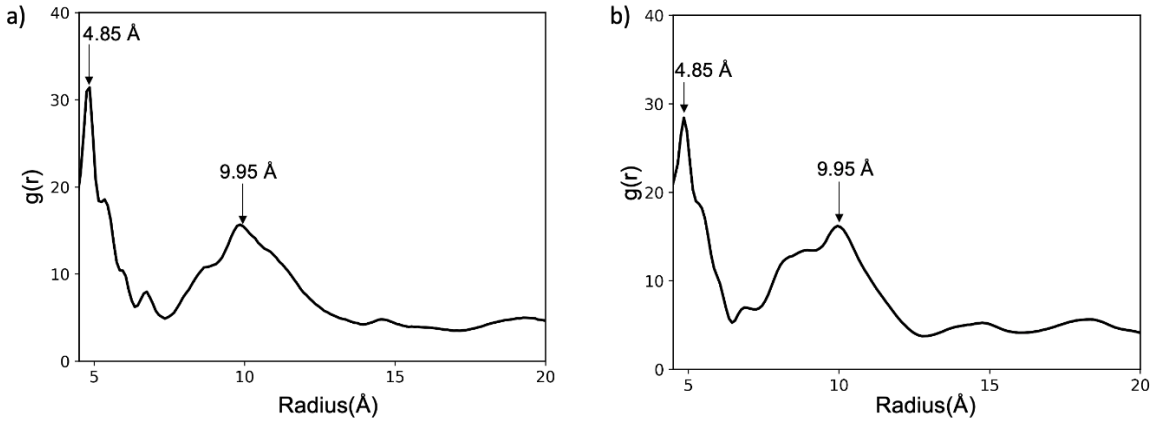

Figure S4. The radial distribution function (RDF) of protonated Glu (GLUP) residue's backbone for 8*16 PA3 system, with H-bonding oriented along the width(a) and length(b) direction calculated separately. Here we show results of GLUP located at the innermost region, GLUP located at other positions share a similar peak. All data were averaged over the last 10 ns of each simulation.

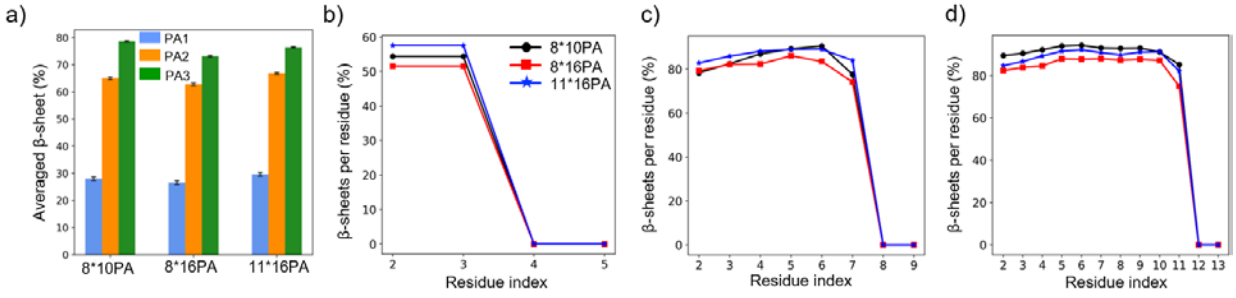

Figure S5. a) Averaged percentage of β-sheets per monomer. Distribution of β-sheets along the PA1(b), PA2(c), PA3(d) sequence chain. All data were averaged over the last 10 ns of each simulation.

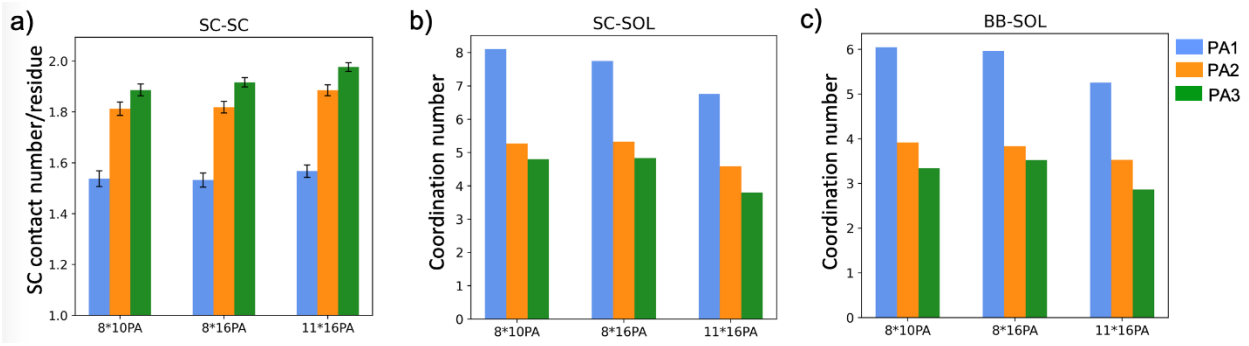

Figure S6. a) Average contacts involved within each PA sidechain. Two side chain neighbors form a contact if their interatomic distance is shorter than 4.5 Å. The coordination number of water molecules within 4.5 Å of the monomer side chains(b) and backbones(c). All data were averaged over the last 10 ns of each simulation.

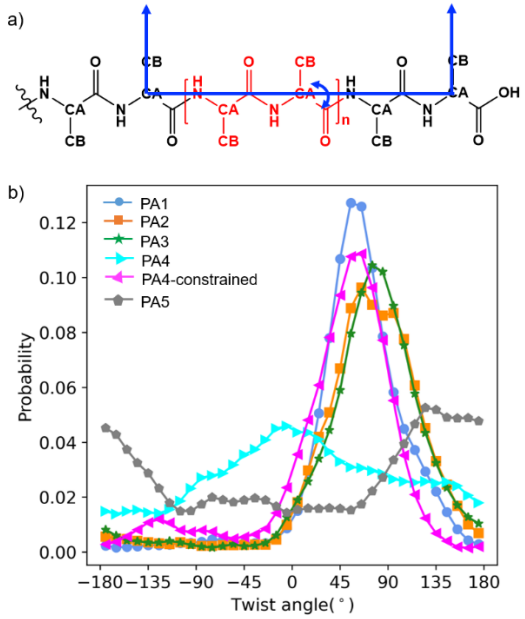

Figure S7. a) Definition of the twist angle within a single strand, defined by the solid blue line, dihedral angle $C\beta_2$-$C\alpha_2$-$C\alpha_m$-$C\beta_m$, where $C\alpha_2$ and $C\beta_2$ are atoms from the second amino acid of the PAs, $C\alpha_m$-$C\beta_m$ are amino acids from the last residue of the PAs, i.e. in PA1, m=4. b) Distribution of the general twist angle within a single strand, only 8*16 PA results are shown. All data were averaged over the last 10 ns of each simulation.

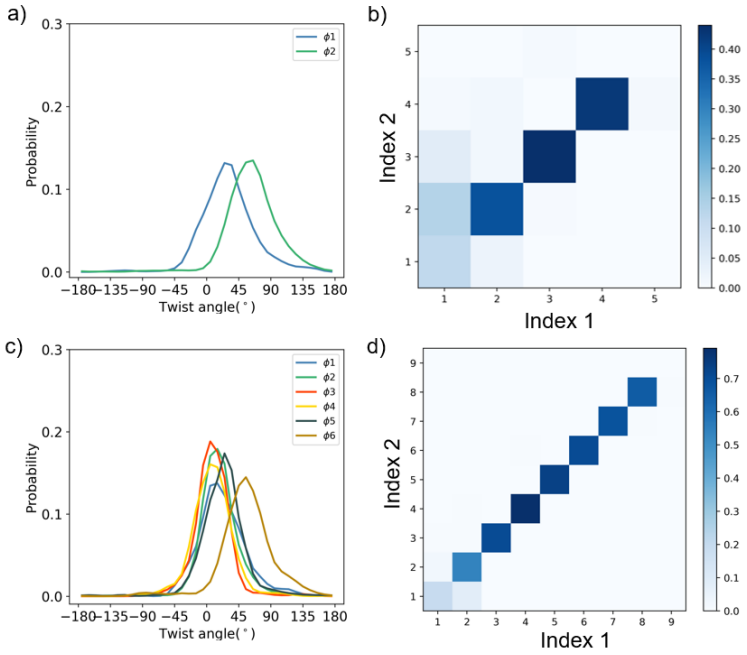

Figure S8. Distributions of the average twist angle (defined in the main text Figure 3a) within PA1 monomers(a) and PA2 monomers(c). Backbone's 2D HB contact map for PA1(b) and PA2(d) ribbon structure, which is calculated between two neighboring peptides. Here we show results for the 8*16 PA system. All data were averaged over the last 10 ns of each simulation.

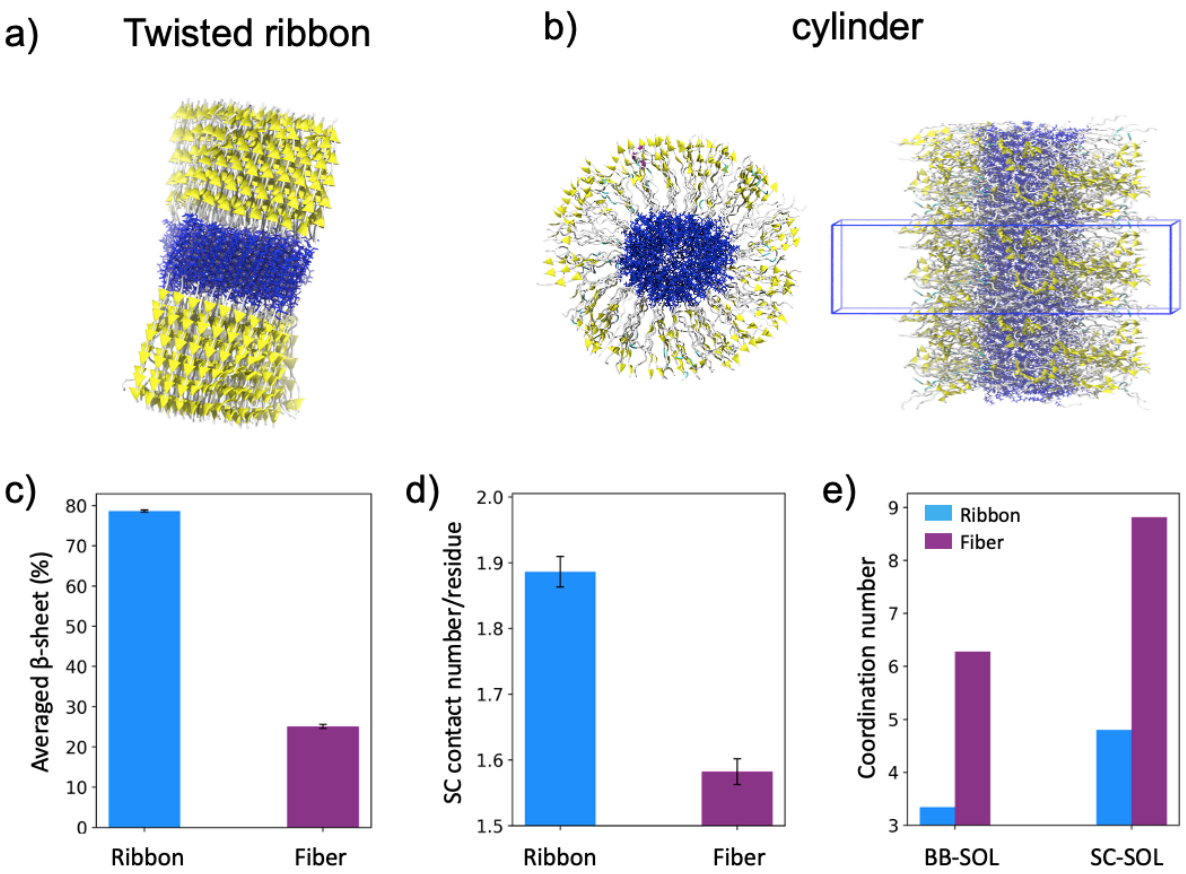

Figure S9. a) Side view of PA3 ribbon at 100ns. b) Top and side views of PA3 nanofiber at 100 ns. The PA assembly structure's core, β-sheets, turns and random coil are shown in blue, yellow, cyan, and gray, respectively. For the peptide portion of the PA, only the peptide backbone is shown. Ions and water are omitted for clarity. Periodic boundaries are shown as a blue square. c) Averaged β-sheet content. d) Average contacts involved within each PA sidechain. e) The coordination number of water molecules within 4.5 Å of the monomer backbone (BB-SOL) and side chain (SC-SOL). All data were averaged over the last 10 ns of each simulation.

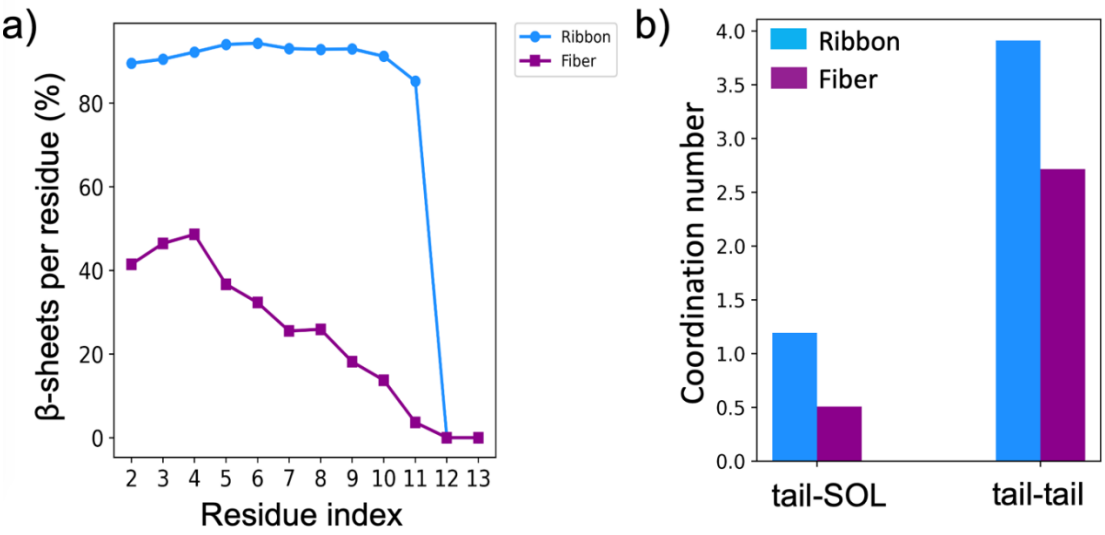

Figure S10. a) Distribution of *β*-sheets along the PA3 chain. b) Coordination number of lipid tail with solvent (tail-SOL) and lipid tail with surrounding lipid tails (tail-tail). Cutoff is 4.5 Å, ribbon and fiber morphology are shown in blue and purple, respectively. All data were averaged over the last 10 ns of each simulation.

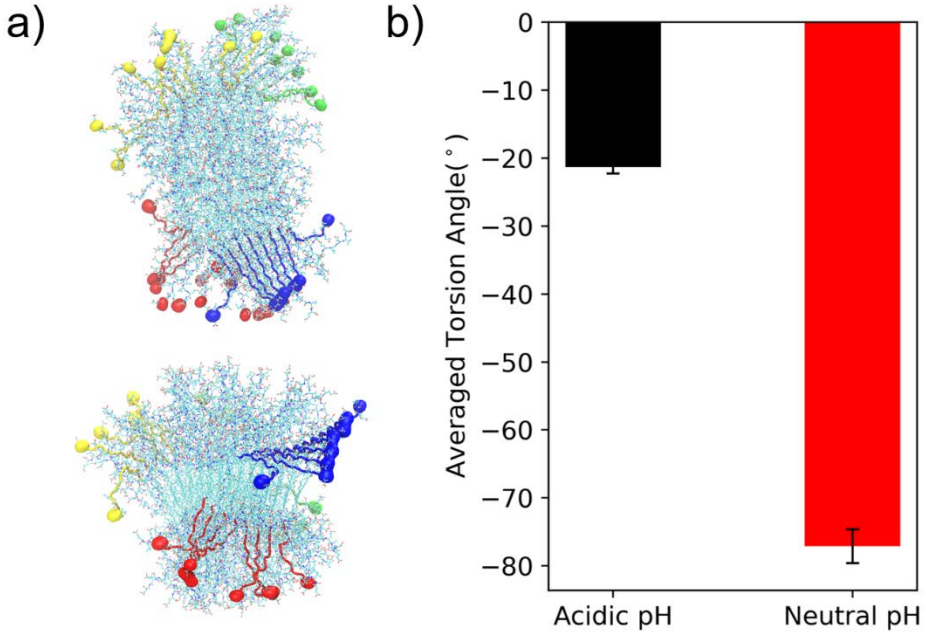

Figure S11. a) Top and side view of PA3 ribbon under neutral pH. b) Averaged torsion angle of PA3 ribbons under acidic pH and neutral pH. All ribbons share the same width/length ratio, that is 8 layers in width and 10 layers in length. All data were averaged over the last 10 ns of each simulation.

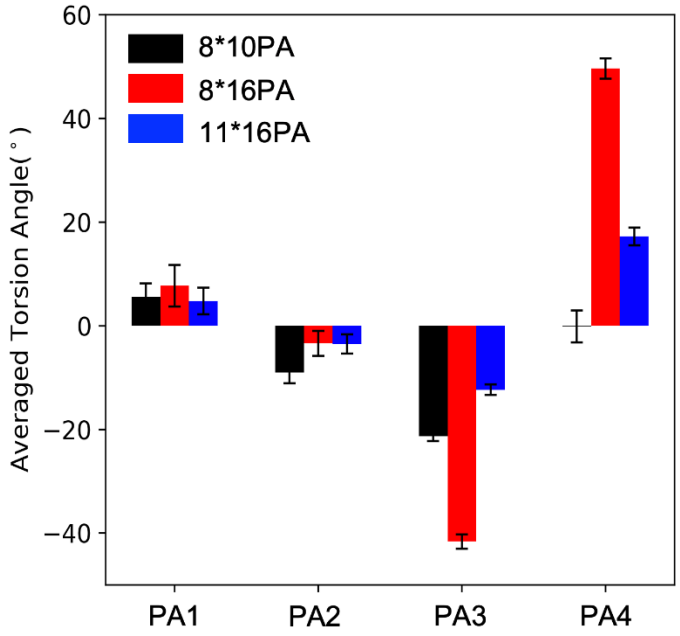

Figure S12. Average torsion angle of ribbon with PA sequence from PA1 to PA4. All data were averaged over the last 10 ns of each simulation.

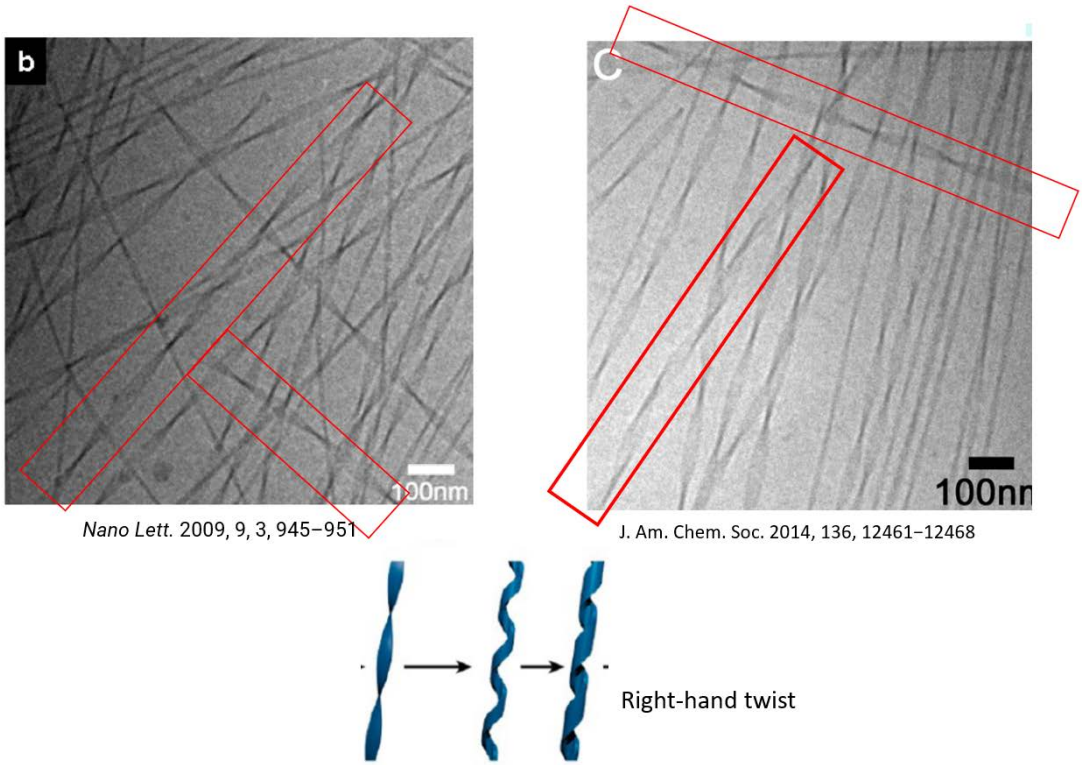

Figure S13. Cryo-TEM images of self-assembled twisted ribbons for PA4. Red frames are examples of right-handed twisted ribbons.

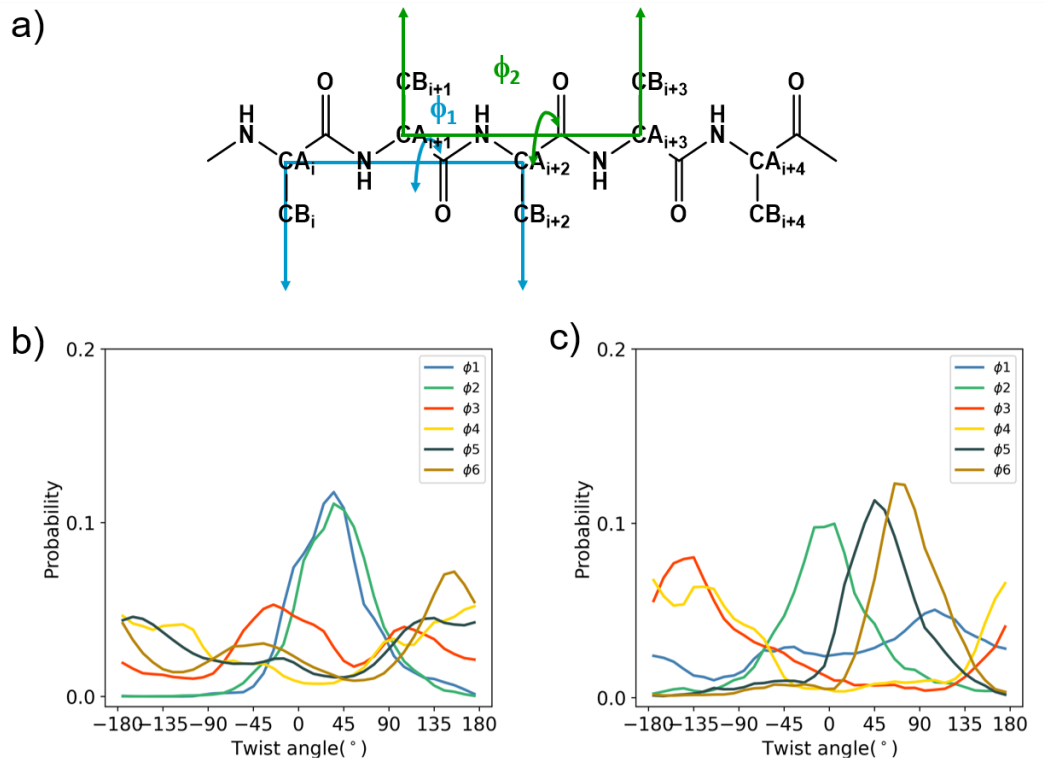

Figure S14 a) Definition of the twist angles ϕ (defined by the solid lines, dihedral angle $C\beta_i$-$C\alpha_i$-$C\alpha_{i+2}$-$C\beta_{i+2}$), where a positive angle represents right-handed twist, and a negative angle represents left-handed twist. Distributions of the averaged twist angle within each PA4(b) and PA5(c) monomer. All data were averaged over the last 10 ns of each simulation.

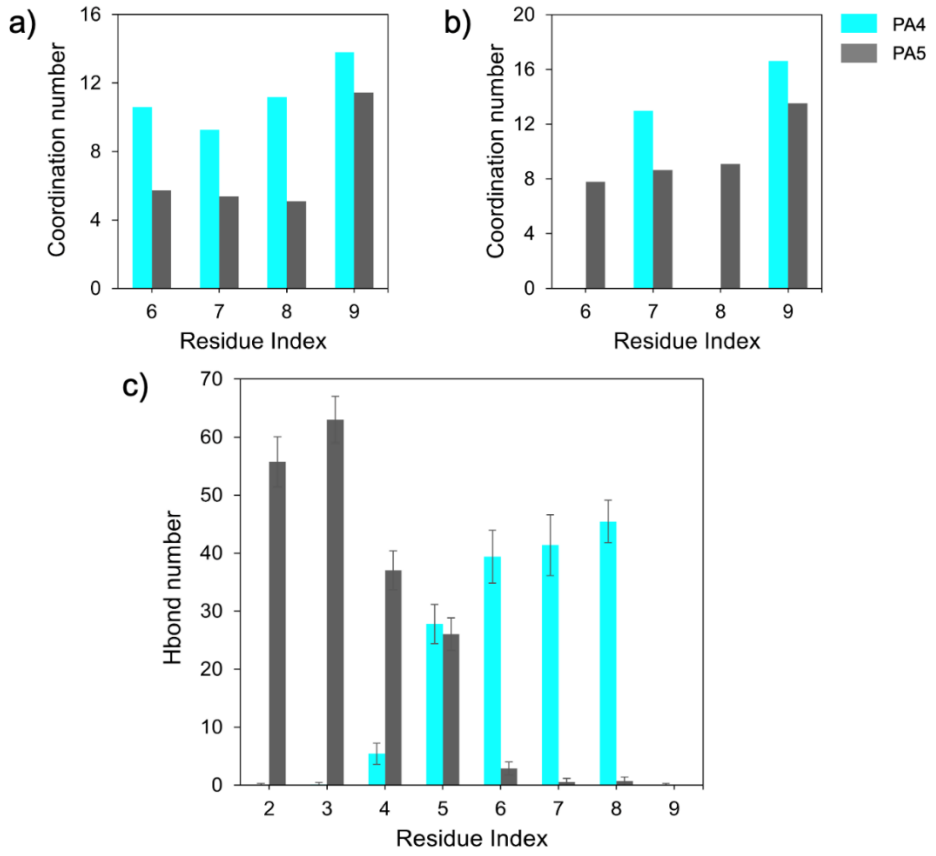

Figure S15. Solvent accessibility of individual residues along the sequence located in the head region of the PAs, with backbone(a) and sidechain(b) calculated separately. c) Number of hydrogen bonds calculated between the Arg's side chain and each residue's backbones along the PA chain. Here the solvent accessibility is evaluated by the coordination number of water molecules within 4.5 Å of the residue's backbone and sidechain. An HB is thought to arise if the distance between donor and acceptor atoms is <0.35 nm and the donor-hydrogen-acceptor angle is >120°. PA4 and PA5 systems are shown in cyan and dark gray, respectively. All data were averaged over the last 10 ns of each simulation.

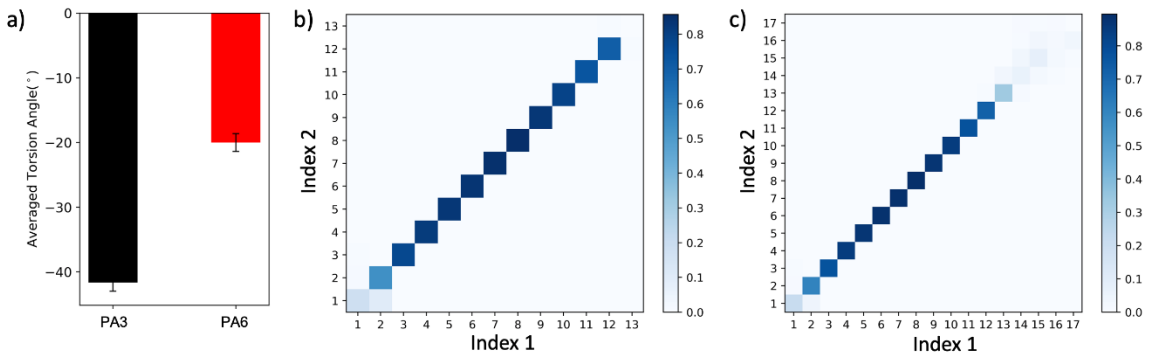

Figure S16. a) Average torsion angle for PA3(black bar) and PA6 ribbons (red bar). Backbone's 2D HB contact map for PA3(b) and PA6(c) ribbons. All data were averaged over the last 10 ns of each simulation.

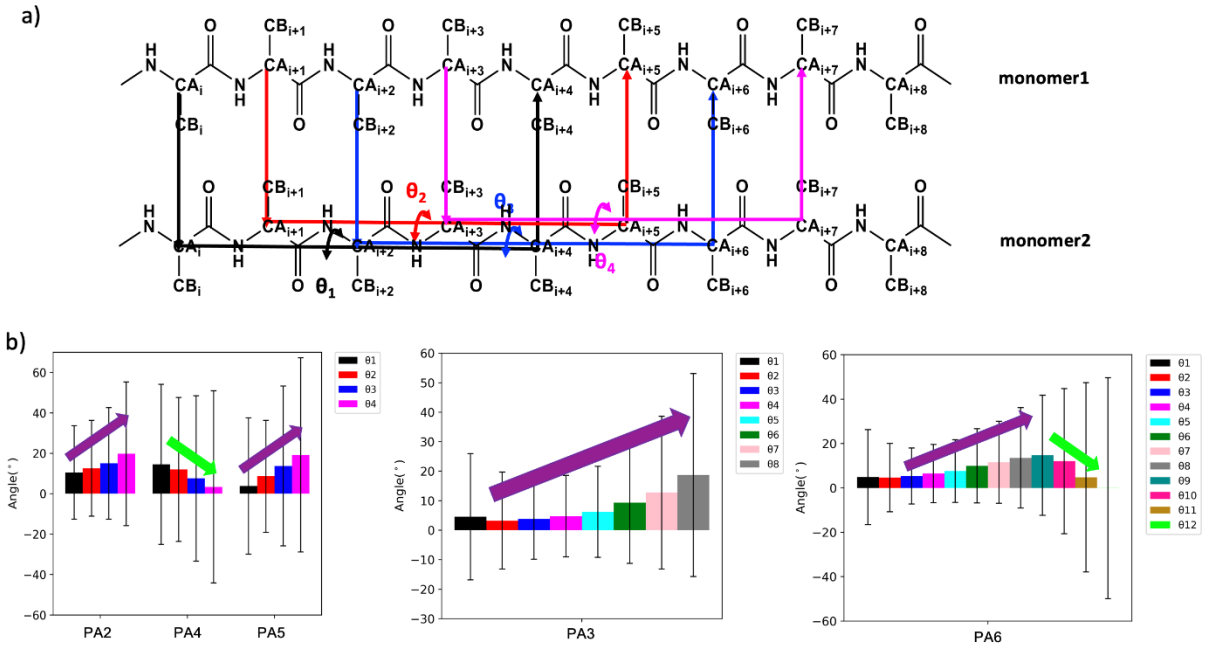

Figure S17. a) Definition of the twist angles θ between neighboring strands (defined by the solid lines, dihedral angle $C\alpha^1_i$-$C\alpha^2_i$-$C\alpha^2_{i+4}$-$C\alpha^1_{i+4}$), which determines the direction and magnitude of the twist between neighboring monomers. b) Averaged twist angles θ between neighboring strands; only 8*16 PA results are shown. All data were averaged over the last 10 ns of each simulation.

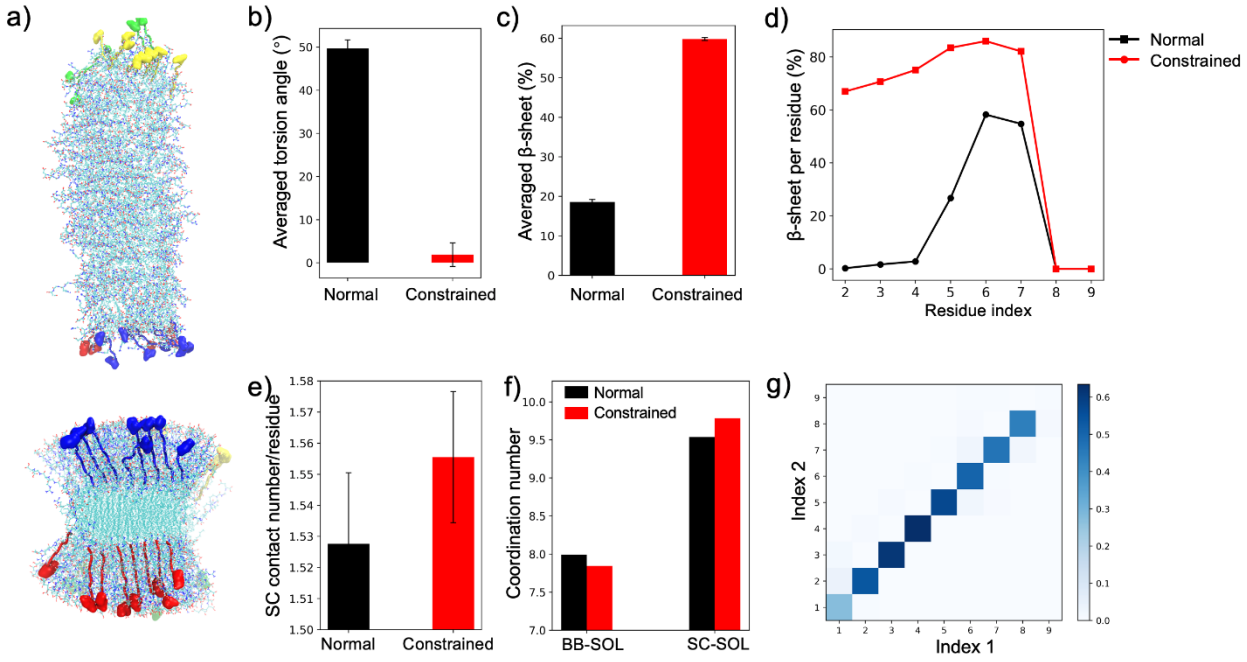

Figure S18. a) Top and side view of final simulation structures of constrained PA4. b) Averaged torsion angle of normal and constrained PA4 ribbons. All ribbons share the same width/length ratio, that is 8 layers in width and 16 layers in length. c) Averaged percentage of *β*-sheets per monomer. d) Distribution of *β*-sheets along the PA4 chain. e) Average contacts involved within each PA sidechain. Two side chain

neighbors form a contact if their interatomic distance is shorter than 4.5 Å. f) The coordination number of water molecules within 4.5 Å of the monomer backbone. g) Backbone's 2D HB contact map for constrained PA4 ribbon. All data were averaged over the last 10 ns of each simulation. In contrast, the rate of side chain accumulation increased by only 2%, and the side chain accumulation is looser compared to PA2. This may be due to the strong hydrophilicity and weak hydrophobic effect of GRGD. All data were averaged over the last 10 ns of each simulation.

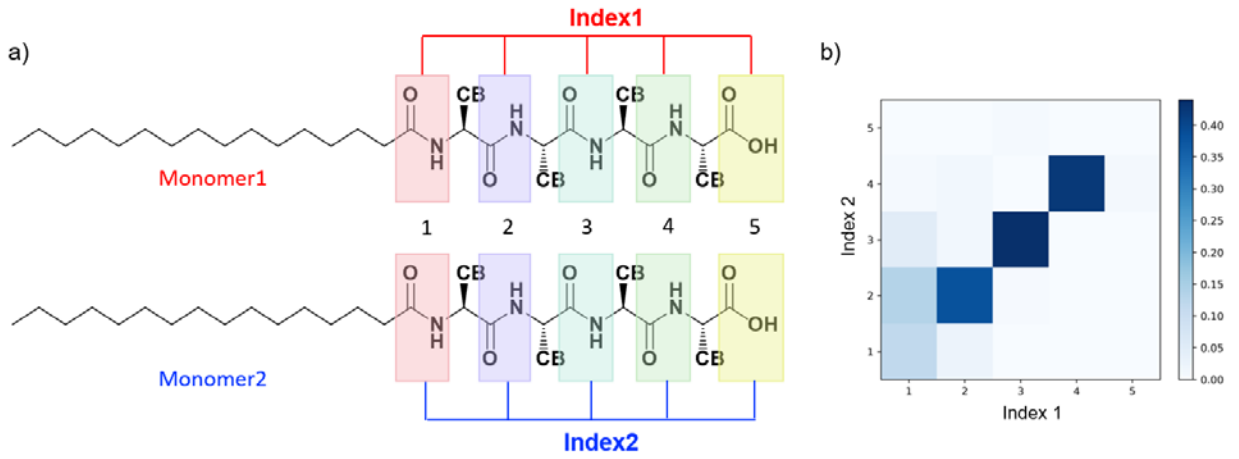

Figure S19. a) Schematic of backbone's hydrogen bond between two neighboring molecules. b) One example of PA1's backbone HB 2D contact map.

REFERENCES

(1) Lai, C. T.; Rosi, N. L.; Schatz, G. C. All-Atom Molecular Dynamics Simulations of Peptide Amphiphile Assemblies That Spontaneously Form Twisted and Helical Ribbon Structures. *J. Phys. Chem. Lett.* **2017**, *8*, 2170–2174.

(2) 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.

(3) 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.

(4) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(5) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(6) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.