Supporting Information

FibrilGen: a Python package for atomistic modelling of peptide β-sheet nanostructures

Chao-Yu Yang[a], Aline F. Miller[b], Alberto Saiani[c], Richard A. Bryce[d,*]

[a] *Department of Materials, Manchester Institute of Biotechnology, School of Natural Sciences, Faculty of Science and Engineering, The University of Manchester, M13 9PL Manchester, U.K.*

[b] *Department of Chemical Engineering, Manchester Institute of Biotechnology, School of Engineering, Faculty of Science and Engineering, The University of Manchester, Manchester M13 9PL, U.K.*

[c] *Division of Pharmacy and Optometry, Manchester Institute of Biotechnology, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester M13 9PL, U.K.*

[d] *Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, M13 9PL Manchester, U.K.*

Table S1. FibrilGen parameters for HP8, AL1 and Aβ$_{42}$ structures and the composition of systems for MD simulation. The 2 x 2 unit of a combination of parallel (p) and antiparallel (a) β-sheet, the FibrilGen models, and the FibrilGen structures parameterized by θ_s (deg), M, r_y (Å), θ_y (deg), θ_z (deg), N, and **K**. Each FibrilGen/ cryo-EM structure was put in the center of simulation box with a minimum distance of 10 Å from the box edges.

Peptides	2 x 2 unit	FibrilGen model	**geometrical parameters**							**number of 2 x 2 units**	**simulation system**
			θ_s	M	r_y	θ_y	θ_z	N	**K**		
HP8	$\beta_p a \beta_a$	stacked ribbon (tube)	72	5	35.0	7.20	30.0	5	-	25	16400 atoms (100 peptides), 64968 TIP3P atoms, 1 Na^+ ion, 101 Cl^- ions
AL1	$\beta_p a \beta_p$	stacked rod	-	-	32.3	1.77	8.8	8	[[1,1],[1,1],[1,1]]	48	38016 atoms (192 peptides), 81723 TIP3P atoms, 140 Na^+ ions, 140 Cl^- ions
Aβ_{42}	pdb 5OQV ($\beta_p a \beta_p$)	a rod	-	-	15.7	1.47	5.0	11	-	11	27588 atoms (44 peptides), 85815 TIP3P atoms, 246 Na^+ ions, 66 Cl^- ions
HP8 (net positively charged)	$\beta_p a \beta_p$	stacked rod	-	-	30.9	2.85	10.0	8	[[1,1],[1,1],[1,1]]	48	31680 atoms (192 peptides), 99339 TIP3P atoms, 1 Na^+ ions, 385 Cl^- ions
HP8 (net neutral)	$\beta_p a \beta_p$	stacked rod	-	-	30.1	2.33	8.0	8	[[1,1],[1,1],[1,1]]	48	31296 atoms (192 peptides), 92115 TIP3P atoms, 7 Na^+ ions, 7 Cl^- ions
AL1	$\beta_p a \beta_p$	2*flat sheet as a tube	-	-	-	-	-	12	-	24	19008 atoms (96 peptides), 45114 TIP3P atoms, 19 Na^+ ions, 19 Cl^- ions
HP8	pdb 7LQI									25[a]	16400 atoms (100 peptides), 65511 TIP3P atoms, 1 Na^+ ion, 101 Cl^- ions

a present but not used in construction

Table S2 Time averaged structural properties of cross-β nanostructures from microsecond MD simulations (final 400 ns), for FibrilGen-constructed HP8 tube, AL1 rod and Aβ_{42} rod, and for cryo-EM structure of HP8 tube, for replicates MD1 and MD2 (italics).

Cross-β structures	**Peptide register t_x (Å)**	**Peptide twist θ_d (deg)**	**Fibril radius R_f (Å)**
HP8 tube (FibrilGen)	7.0 ± 0.3 *7.1 ± 0.2*	14.7 ± 2.2 *13.9 ± 1.8*	28.4 ± 0.2 *28.2 ± 0.1*
HP8 tube (cryo-EM structure 7LQI)	6.9 ± 0.3 *7.1 ± 0.4*	17.4 ± 2.6 *20.8 ± 3.1*	29.8 ± 0.5 *30.2 ± 0.4*
AL1 rod (FibrilGen)	-0.0 ± 0.2 *0.1 ± 0.3*	3.9 ± 0.7 *3.5 ± 1.2*	26.6 ± 0.0 *26.6 ± 0.0*
Aβ_{42} rod (FibrilGen)	-0.1 ± 0.2 *-0.0 ± 0.3*	7.4 ± 3.6 *6.6 ± 4.1*	15.6 ± 0.0 *15.8 ± 0.0*

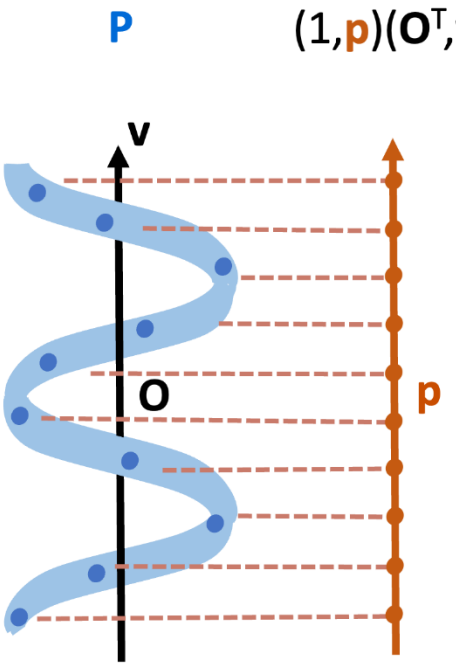

Figure S1. An estimated long axis (black arrow) along the fibril (blue ribbon). A fibril axis that spaces peptides by their numbering in the β-sheet (orange dots) is represented as $(1,\mathbf{p})(\mathbf{O}^T,\mathbf{v}^T)^T$ and then it is least squares fitted to the peptide center of mass **P** (blue dots).

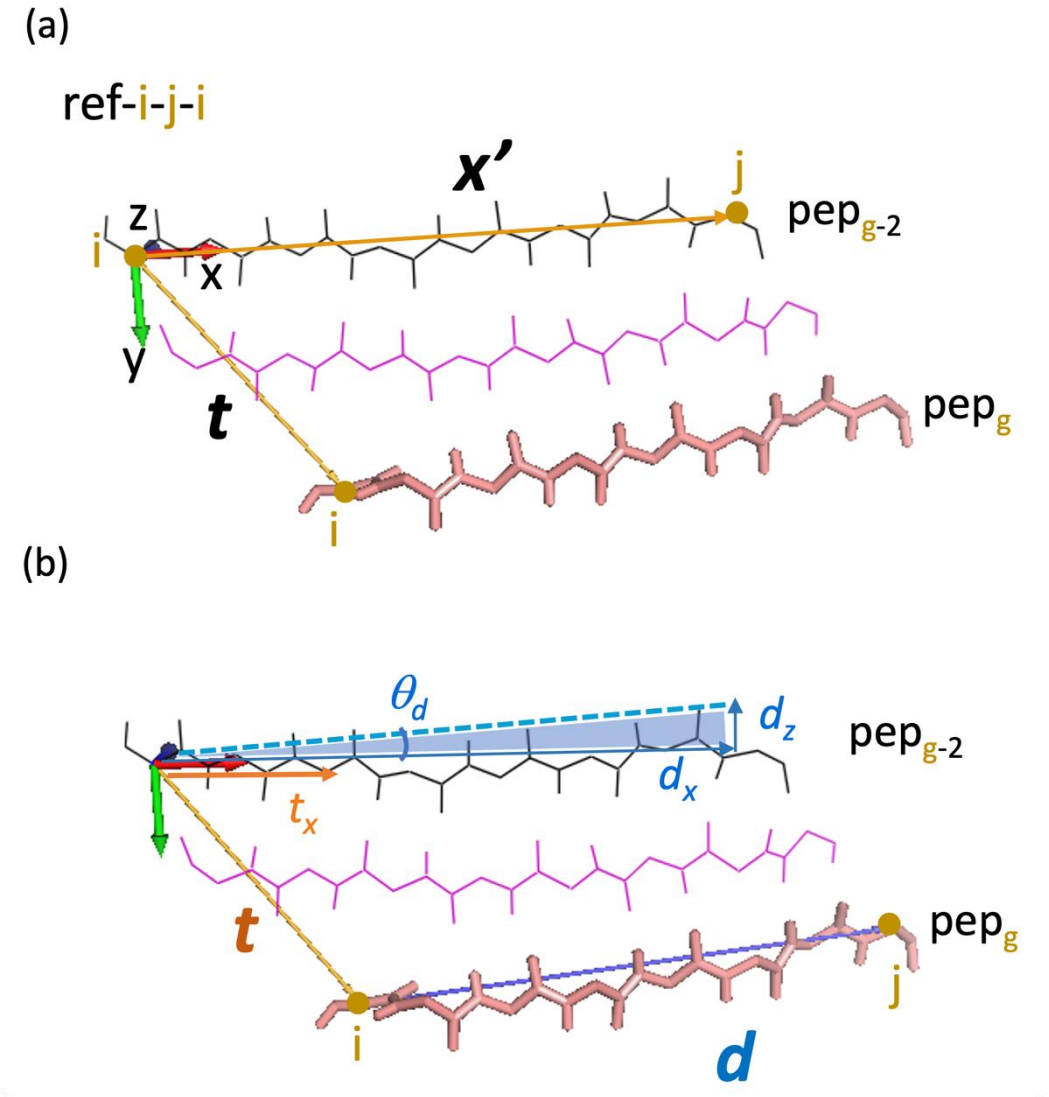

Figure S2. Backbone alignment between two selected peptides, namely peptide *g*-2 (black line) and peptide *g* (pink stick). (a) Reference coordinate system ref-i-j-i is defined by a vector $\boldsymbol{x}'$ from atom $C_{\alpha,i}$ to $C_{\alpha,j}$ of peptide *g*-2, and a vector $\boldsymbol{t}$ from atom $C_{\alpha,i}$ of peptide *g*-2 to atom $C_{\alpha,i}$ of peptide *g*. The coordinate system is constructed by the unit vector of $\boldsymbol{x}', (\boldsymbol{x}' \times \boldsymbol{t}) \times \boldsymbol{x}', \boldsymbol{x}' \times \boldsymbol{t}$. (b) Definition of backbone register of *g* and *g*-2 peptides in the same β-sheet, indicated by the x component of $\boldsymbol{t}$ (orange); and backbone twist $\theta_d = \tan^{-1}(d_z / d_x)$ is defined by vector $\boldsymbol{d}$ (blue) from atom $C_{\alpha,i}$ of peptide *g* to its $C_{\alpha,j}$.

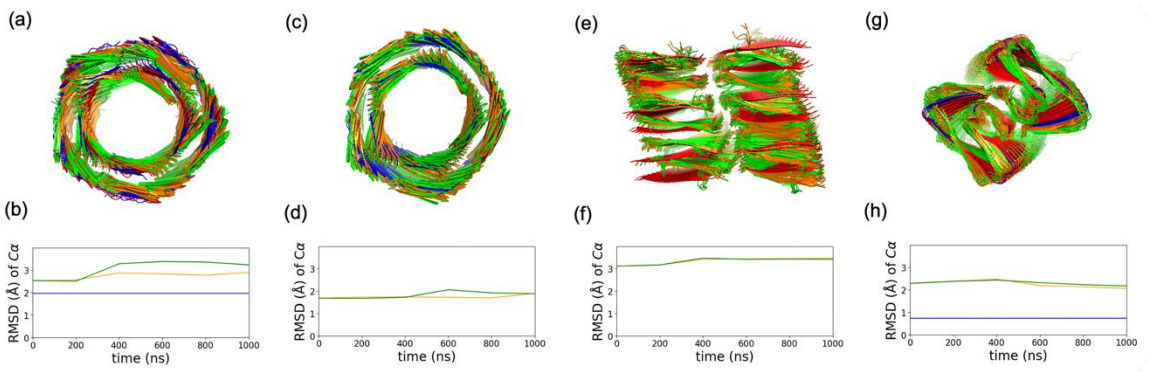

Figure S3. Comparison of replicate 1 μs MD production simulations at 300 K of peptide nanostructures (a-d) HP8, (e,f) AL1 and (g,h) Aβ_{42}, with the initial pre-equilibration structures: (a,e,g) superposition of FibrilGen structures (red) with equispaced MD snapshots (green, orange) and experimental structure (blue); and (c) superposition of cryo-EM structure 7LQI (blue) with equispaced MD snapshots (green, orange). (b,d,f,h) RMSD value over all C_α atoms for each structural alignment (the aligned RMSD is labelled in green, orange for MD snapshots, and blue for experimental structure).

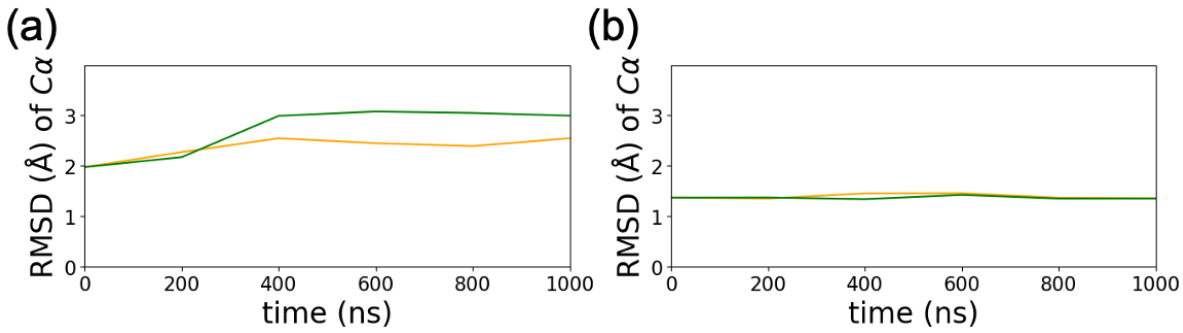

Figure S4. RMSD of C_α atoms over the two MD simulations (green, orange) for (a) a FibrilGen HP8 tube compared to HP8 cryo-EM structure 7LQI; and (b) a 2x2 peptide unit from the FibrilGen Aβ_{42} rod compared to the equivalent unit from Aβ_{42} cryo-EM structure 5OQV.

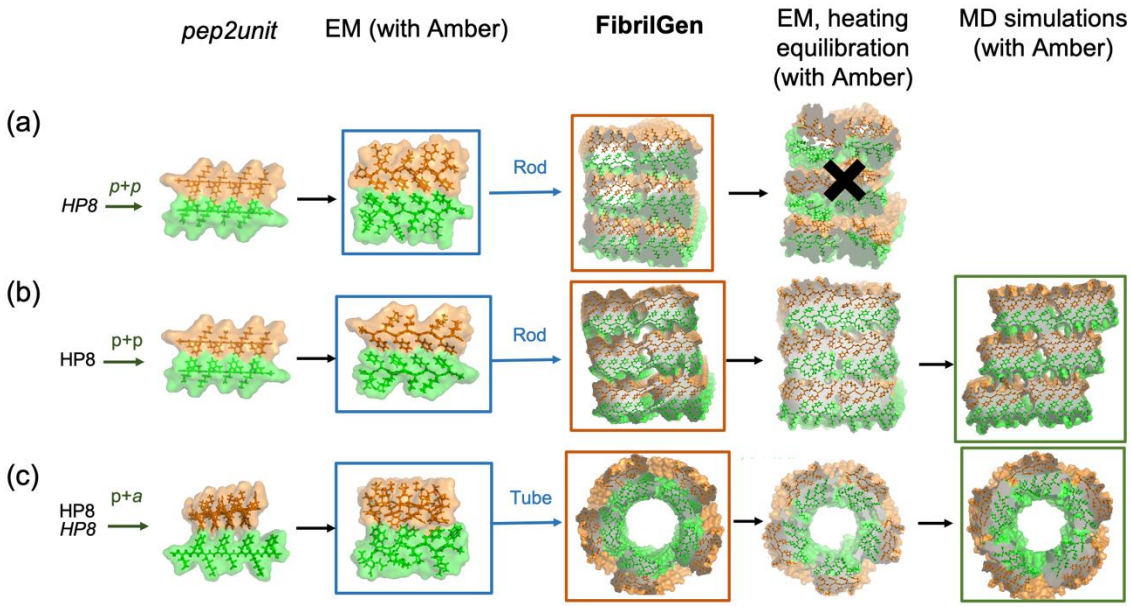

Figure S5. Screening HP8 assemblies for (a) a positively charged HP8 rod; (b) a neutral HP8 rod; and (c) a positively charged HP8 tube. Structures that fail to hold the overall assembly are rejected (labelled X) from the modeling pipeline.

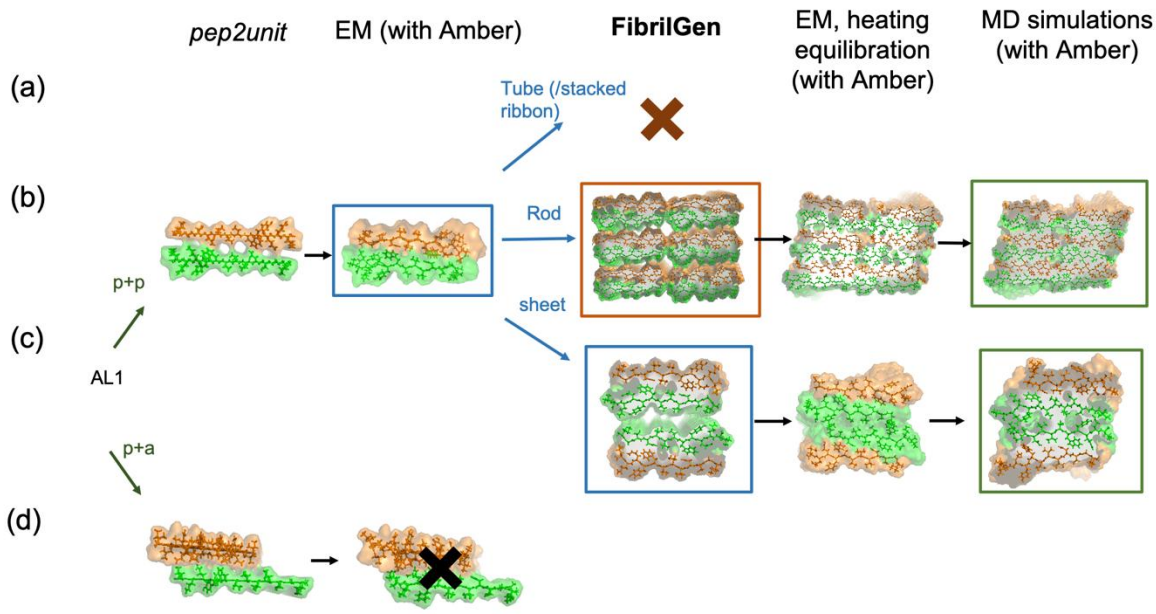

Figure S6. Screening AL1 assemblies for (a) a AL1 tube; (b) a AL1 rod; (c) a face-to-face aligned AL1 sheet; and (d) a AL1 tube. Structures that fail to be assembled (brown X) or fall apart (black X) are rejected from the modeling pipeline.

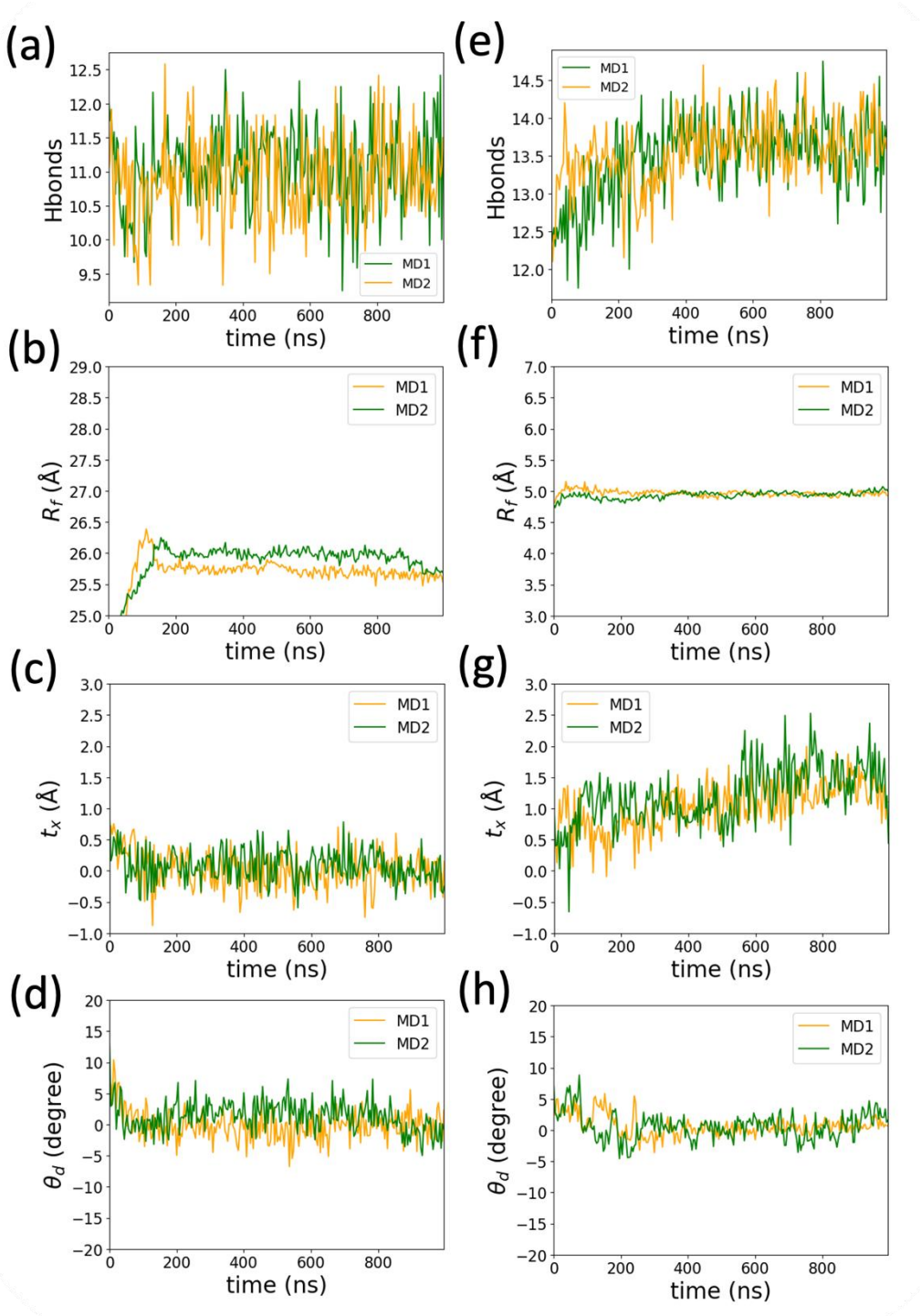

Figure S7. Time series of backbone hydrogen bonds, fibril radius R_f, relative peptide alignment (x-register t_x, y-twist angle θ_d) over replicate microsecond MD simulations at 300 K (green, orange) for a (a-d) neutral HP8 rod, and a (e-h) AL1 sheet.

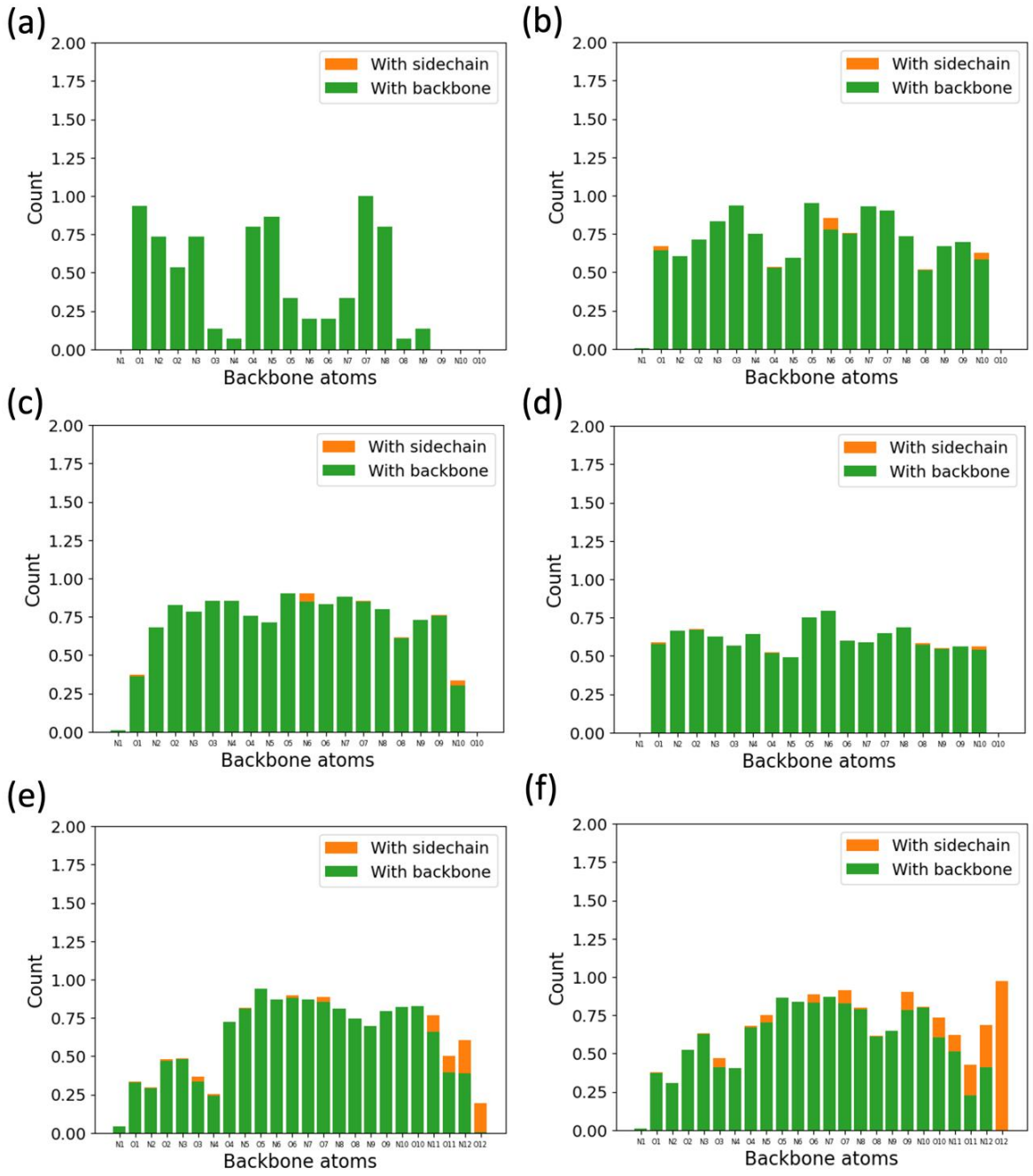

Figure S8. Hydrogen bonds of backbone atoms averaged over β-strands, for (a) antiparallel aligned HP8 in cryo-EM structure 7LQI; (b) antiparallel aligned HP8 after MD simulation of the structure 7LQI; (c) antiparallel aligned HP8 after MD simulation of the HP8 tube; (d) parallel aligned HP8 after MD simulation of the neutral HP8 rod; (e) parallel aligned AL1 after MD simulation of the AL1 rod; (f) parallel aligned AL1 after MD simulation of the AL1 sheet. The nitrogen and oxygen atoms are indexed (started at 1) from the N-terminus to the C-

terminus, counted by the number of atoms they bond to. Backbone-backbone hydrogen bonds are colored in green and backbone-sidechain hydrogen bonds are colored in orange.

FibrilGen functions for fibril generation

FibrilGen has 6 functions under the class "fibril". Each function takes a set of input parameters described below.

1. fibril.build_a_flat_sheet(*N*)

   The function "build_a_flat_sheet" takes only one parameter *N* to specify the repeat of the 2x2 unit along the β-sheet axis (Figure 2a).

2. fibril.build_a_stacked_sheet(**K**, *N*)

   The function "build_a_stacked_sheet" takes input parameter *N* and **K** to stack the 2x2 unit along the fibril long axis (Figure 2a) and on the fibril cross-section (Figure 2b), respectively.

3. fibril.build_a_rod(θ_z, *N*, the sign of θ_y)

   The function "build_a_rod" takes input parameter *N* to stack the 2x2 unit along the β-sheet axis (Figure 2a). An initial helical twist is assigned with a tilt angle θ_z and the direction (assigned to 1 or -1) of the twist angle θ_y (Figure 2c).

4. fibril.build_a_stacked_rod(θ_z, **K**, *N*, the sign of θ_y)

   The function "build_a_stacked_rod" takes input parameter *N* and **K** for a linear stacking of the 2x2 unit along the fibril long axis (Figure 2a) and on the fibril cross-section (Figure 2b). An initial helical twist is assigned with a tilt angle θ_z and the direction (assigned as 1 or -1) of the twist angle θ_y (Figure 2c).

5. fibril.build_a_ribbon(θ_z, r_y, *N*, the sign of θ_y)

   The function "build_a_ribbon" takes input parameter *N* to stack the 2x2 unit along the fibril long axis (Figure 2a). An initial helical twist is assigned with a tilt angle θ_z, a radius r_y and the direction (assigned as 1 or -1) of the twist angle θ_y (Figure 2c).

6. fibril.build_a_stacked_ribbon(θ_z, r_y, θ_s, *M*, *N*, the sign of θ_y)

The function “build_a_stacked_ribbon” takes input parameter *N* to stack the 2x2 unit along the fibril long axis (Figure 2a). The rotational stacking on the fibril cross-section with an incremental rotation angle θ_s repeated for *M* times is assigned (Figure 2b). An initial helical twist is assigned with a tilt angle θ_z, a radius r_y and the direction (assigned as 1 or -1) of the twist angle θ_y (Figure 2c). Here tube is a special case that $\theta_S \cdot M = 360°$.

Python script of templating fibril assemblies

Suppose that we want to assemble an energy minimized 2 x 2 HP8 unit into a 6-sheet ribbon of left-handed chirality; a 3-fold periodic stacking around the fibril long-axis by an increasing angle of 90°; a repeat stacking of 20 units along the fibril long-axis; and in a fibril morphology of a tilt angle of 10° and a radius of 30 Å. We can use the FibrilGen function *build_a_stacked_ribbon* to generate a compact 6-sheet ribbon atomistic structure, where the helical parameters can be automatically refined to yield a tilt angle of 14°, a radius of 26 Å and a twist angle 2.6° (Figure S9).

(a)

```
# Import FibrilGen library
pymol.cmd.run('builder.py')
# Load a bilayer structure
pymol.cmd.load('input/examples/capF8.pdb')
# Select four peptides p1,p2,p3,p4 to construct a 2x2 assembly unit
pymol.cmd.select('p1','resi 21-30')
pymol.cmd.select('p2','resi 31-40')
pymol.cmd.select('p3','resi 111-120')
pymol.cmd.select('p4','resi 121-130')
# Select three atoms po1,po2,po3 to construct a reference coordinate
system
pymol.cmd.select('po1','resi 22 and name ca')
pymol.cmd.select('po2','resi 29 and name ca')
pymol.cmd.select('po3','resi 62 and name ca')
# Create a periodic unit
unit = create_sheet_unit('p1','p2','p3','p4','po1','po2','po3')
# Create a fibril object
fibril = create_fibril(unit)
# Build a cross-beta structure (e.g., a ribbon of tilt angle, radius, angle
for rotational stacking, number of rotational stacking, number of units
in each beta-sheet, and the direction of twist)
fibril.build_a_stacked_ribbon(10,30,90,3,20,1)
```

(b)

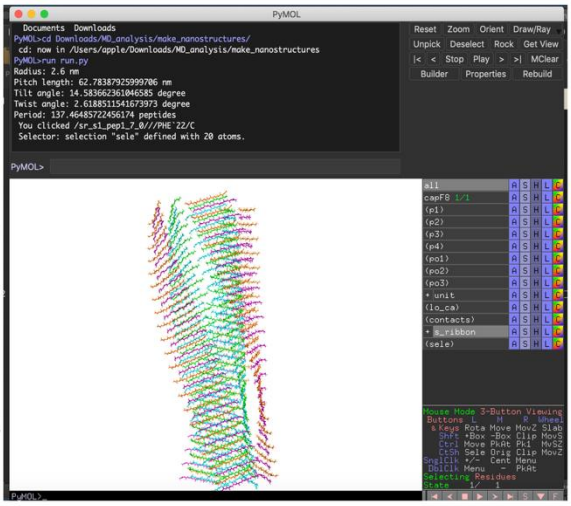

Figure S9. Top-down morphological modelling of a cross-β nanostructure of HP8 peptides with the FibrilGen library. A python script written as (a) selecting four peptides for an assembly unit (2 x 2 peptides) and three C_α from one sheet of the unit to build the reference coordinate system (see Figure S2a). The unit is packed into a ribbon structure by calling the function *build_a_stacked_ribbon* with a set of geometrical parameters: tilt angle θ_z, radius of rotation r_y, stacking angle θ_s, stacking number M, number of units per sheet N, and twisting direction (the sign of θ_y). Then (b) the python script is run at the PyMOL command line to generate and visualize the structure.

Systematic variations of FibrilGen input parameters (Figures S10 – S11)

FibrilGen controls fibril morphologies by variables r_y, θ_z, M, the sign of θ_y, and **K**. In the following, we use peptide HP8 (Figures S10) and AL1 (Figure S11) to demonstrate the effect of input parameters on the overall fibril structures.

Suppose that we want to assemble a HP8 unit into a fibril with a different radius, tilt angle or symmetry of tube. We can observe that a smaller value of r_y assigned to FibrilGen function *build_a_ribbon* makes a tighter ribbon (Figure S10a,b); a smaller value of θ_z assigned to FibrilGen function *build_a_stacked_ribbon* generates a ribbon being less tilted from the fibril long axis (Figure S10c,d); M and θ_s = 360/M assigned to FibrilGen function *build_a_stacked_ribbon* assembles a tube with cyclic symmetry of M (Figure S10e,f).

Suppose that we want to assemble a given AL1 basic 2 x 2 peptide unit into different rods. Given that the sign of θ_y controls the direction of a fibril rotation, we demonstrate that a value of 1 and -1 assigned to FibrilGen function *build_a _rod* makes a left-handed rod (Figure S11a) and a right-handed rod (Figure S11b), respectively. The dimensions of **K** determine the directional stacking of units along the peptide backbone or perpendicular to the peptide backbone. Using FibrilGen function *build_a_stacked_rod*, we show a stacking of 2x3 units (Figure S11c); and 3x2 units (Figure S11d) on the fibril cross-section.

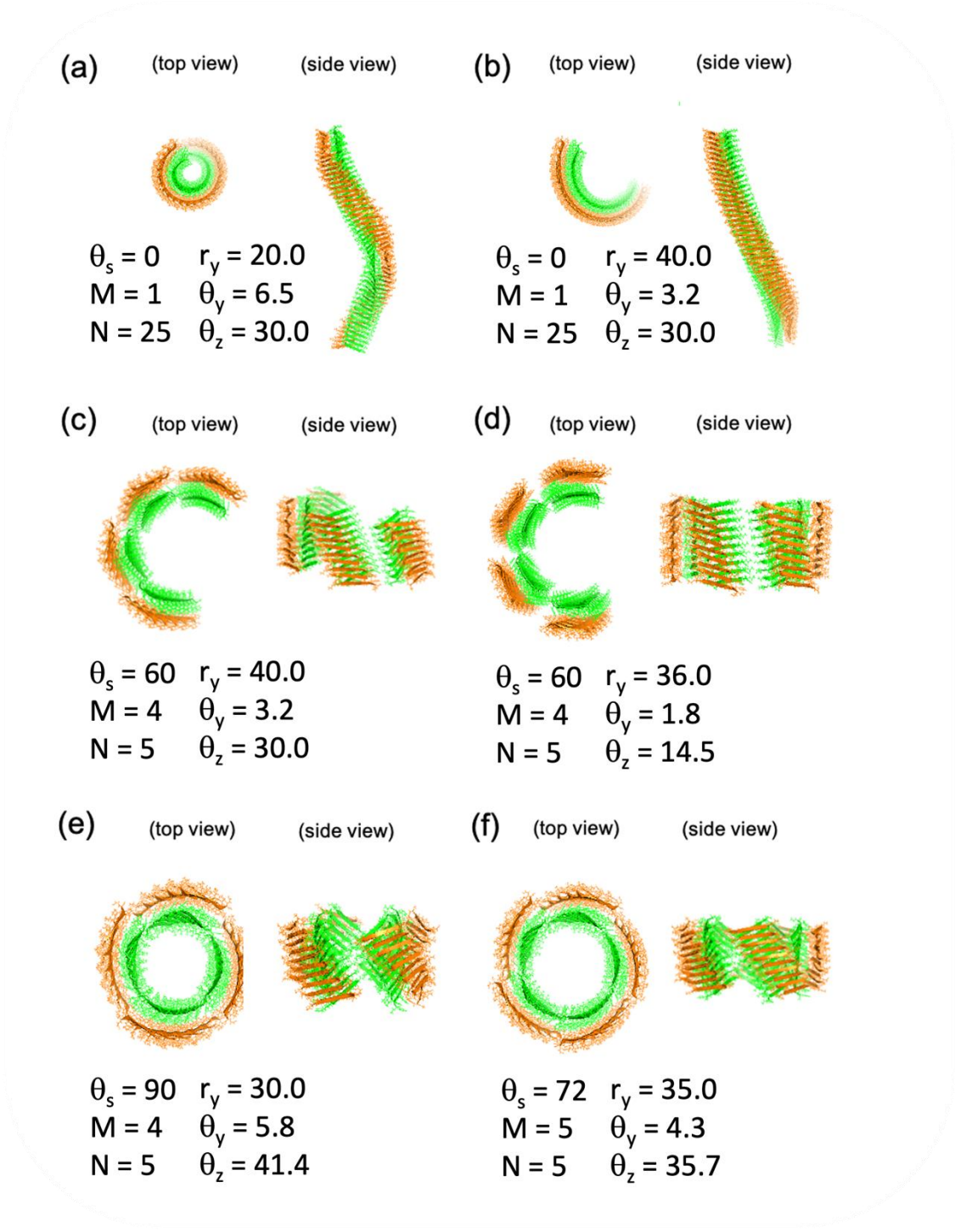

Figure S10. Use of HP8 as the reference peptide to generate a range of ribbon/ tube structures by geometrical parameters θ_s, M, r_y, θ_y, θ_z, and N. A (a) tighter ribbon compared to (b) with a smaller value r_y. (c) A more tilted ribbon compared to (d) by a larger θ_z. (e) A C_4 tube by (M, θ_s) of (4, 90°) (f) A C_5 tube by (M, θ_s) of (5, 72°).

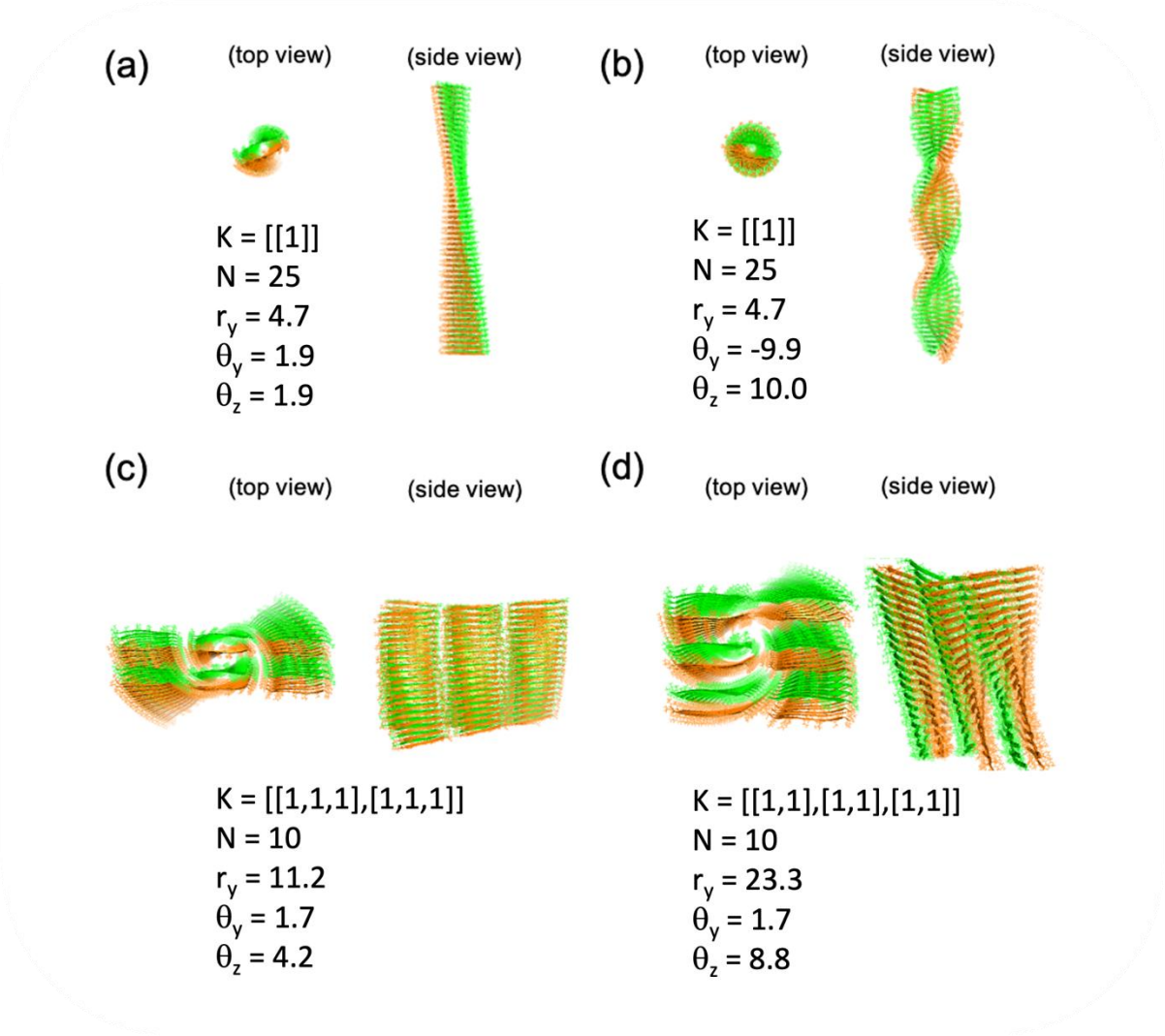

Figure S11. Use of AL1 as the reference peptide to generate a range of rod structures by geometrical parameters r_y, θ_y, θ_z, N, and **K**. (a) A left-handed by a positive θ_y or (b) a right-handed rod by a negative θ_y. (c) A rod that stacks units with $\mathbf{K} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, compared to (d) a rod that stacks units with $\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Hypothesizing fibril structures with FibrilGen

We comment on two cases as follows:

1. Cryo-EM and ssNMR data are available.

The 3D reconstruction of cryo-EM electron density could provide the β-sheet stacking pattern on the fibril cross-section and a rough estimate of helical twist (Figure 1). In this way, we can assign a linear stacking **K** or a rotational stacking (M, θ_s); and a helical twist of the fibril chirality (the sign of θ_y), tilt angle θ_z, and radius r_y. Data from ssNMR may confirm peptide alignments inside the election density that can assist in assigning a 2 x 2 unit (Figure 2a, Figure 3).

2. Experimental data are not available

The modelling scheme in Figure 2 illustrates controllable parameters in FibrilGen, where the choice of the 2 x 2 unit is discrete, parameter space (**K**, M, θ_s) is discrete and parameter space (r_y, θ_y, θ_z) is continuous. In FibrilGen, the 2 x 2 unit and the stacking parameters (**K**, M, θ_s) determine a combinatorial peptide assembly and constrain the feasible helical twists. FibrilGen features feasible geometrical relation by automatically refining user assigned helical twist (r_y, the sign of θ_y, θ_z) to a compact and non-overlapping fibril assembly.

FibrilGen/MD modeling pipeline

FibrilGen provides geometrically feasible peptide assemblies with pre-assembled backbone hydrogen bonding and sidechain packing. Whether such an assembly is energetically favorable can be assessed by energy minimization, heating equilibration and MD simulation via a suitable software package. A modelling pipeline combining FibrilGen with MD simulation to accept or reject hypothetical structures (Figure S12) can provide a systematic investigation of fibril assembly in various protonation states and solvent conditions.

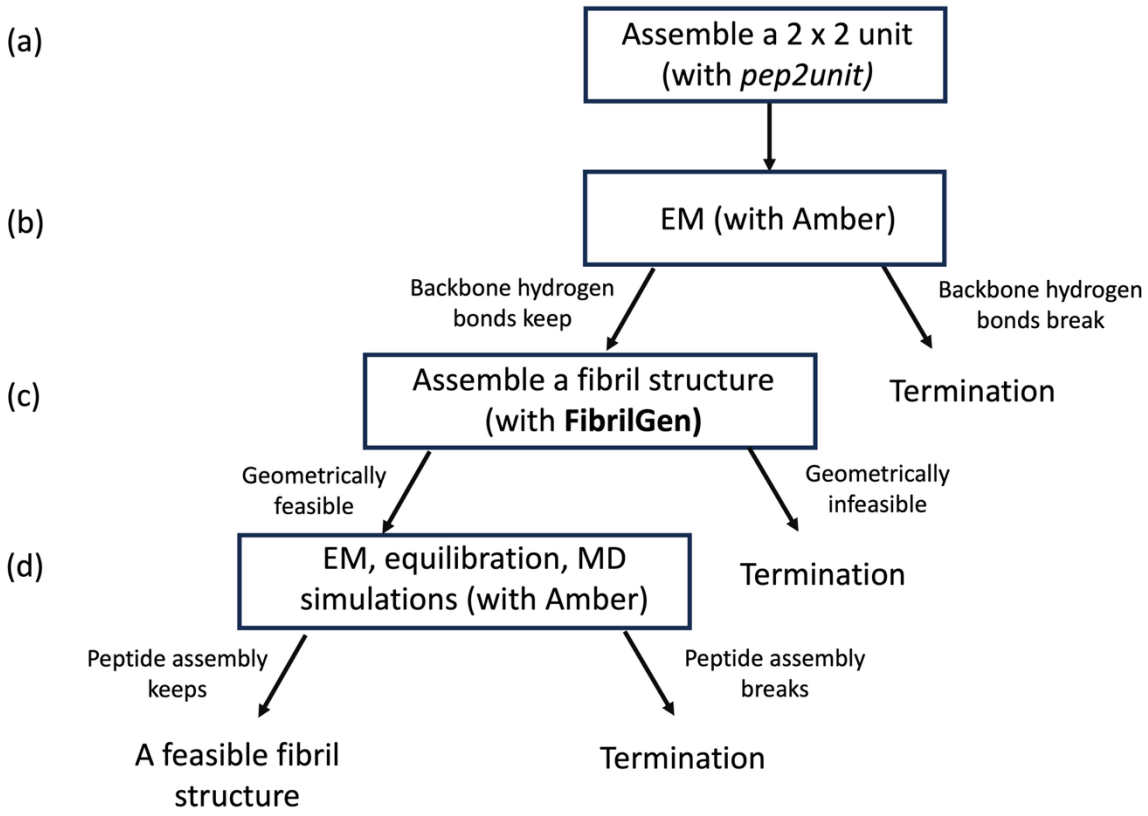

Figure S12. FibrilGen/MD workflow for hypothesizing fibril structures (a) Initialize a 2 x 2 unit (b) Assess whether the 2 x 2 unit can be energy minimized (EM) without losing backbone hydrogen bonds (c) Assemble an energy minimized 2 x 2 unit into a fibril structure (d) Assess whether the fibril structure can be energy minimized and equilibrated without losing the combined assembly, and then assess whether the overall structure is stable in MD simulations.

Helical parameters extracted from cryo-EM reconstructed models

For the HP8 tube cryo-EM structure,[1] the tilt angle θ_z of the inner walls from the fibril axis is 30°, and the inner walls form a pore a radius of 30 Å. Then by Equation 1, the twist angle θ_y is given as:

$$\cos^{-1}\left(1-\frac{\left(\frac{4.8\sin 30^\circ}{30}\right)^2}{2}\right)=4.5^\circ \quad (1)$$

For the AL1 rod cryo-EM structure,[2] the repeat distance along the fibril axis is reported as 4.69 Å; we estimate the tilt angle θ_z as arccos(4.69/4.8) = 12.2°. The given pitch distance is 1150 Å with a peptide-peptide distance 4.69 Å along the fibril axis. This implies 1150/4.69 = 246 peptides to rotate around the fibril long axis in a complete turn, where the average twist angle θ_y is 360/246° = 1.4°.

For the $A\beta_{42}$ rod cryo-EM structure,[3] the pitch distance is reported as 1160 Å with a peptide-peptide distance of 4.67 Å along the fibril axis; we estimate an average of twist angle θ_y of 360/(1160/4.67) = 1.4°. The unit cell for the experimental structure has a width of 31.4 Å, giving 15.7 Å as the radius of rotation. Then by Equation (1), the tilt angle θ_z is $\sin^{-1}(15.7\cdot\sqrt{2-2\cos 1.4^\circ}/4.8)=4.5^\circ$.

References

(1) Wang, F.; Gnewou, O.; Wang, S.; Osinski, T.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Deterministic chaos in the self-assembly of beta sheet nanotubes from an amphipathic oligopeptide. *Matter* **2021**, *4*, 3217-3231.
(2) Close, W.; Neumann, M.; Schmidt, A.; Hora, M.; Annamalai, K.; Schmidt, M.; Reif, B.; Schmidt, V.; Grigorieff, N.; Fandrich, M. Physical basis of amyloid fibril polymorphism. *Nat Commun* **2018**, *9*, 699.
(3) Gremer, L.; Scholzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; et al. Fibril structure of amyloid-beta(1-42) by cryo-electron microscopy. *Science* **2017**, *358*, 116-119.