Comprehensive Characterization of Polyproline Tri-Helix Macrocyclic Scaffolds for Predictive Ligand Positioning

Chia-Lung Tsai,[1] Je-Wei Chang,[2] Kum-Yi Cheng,[3] Yu-Jing Lan,[1] Yi-Cheng Hsu,[1] Qun-Da Lin,[1] Tzu-Yuan Chen,[1] Orion Shih,[2] Chih-Hsun Lin,[3] Po-Hsun Chiang,[1] Mantas Simenas,[4] Vidmantas Kalendra,[4] Yun-Wei Chiang,[1,] Chun-hsien Chen,[3,*] U-Ser Jeng[2,5,6,*] and Sheng-Kai Wang[1,7,*]*

[1] Department of Chemistry, National Tsing Hua University, Hsinchu 300044, Taiwan
[2] National Synchrotron Radiation Research Center, Hsinchu 300092, Taiwan
[3] Department of Chemistry and Centre for Emerging Materials and Advanced Devices, National Taiwan University, Taipei 10617, Taiwan
[4] Faculty of physics, Vilnius University, Sauletekio 3, LT-10257 Vilnius, Lithuania
[5] Department of Chemical Engineering, National Tsing Hua University, Hsinchu 300044, Taiwan
[6] College of Semiconductor Research, National Tsing Hua University, Hsinchu, 300044, Taiwan
[7] Frontier Research Center on Fundamental and Applied Sciences of Matters, National Tsing Hua University, Hsinchu 300044, Taiwan

Emails: skwang@mx.nthu.edu.tw; usjeng@nsrrc.org.tw; chhchen@ntu.edu.tw; ywchiang@mx.nthu.edu.tw

Sheng-Kai Wang, ORCID: 0000-0002-3827-7983
U-Ser Jeng, ORCID: 0000-0002-2247-5061
Chun-hsien Chen, ORCID: 0000-0001-5507-3248
Yun-Wei Chiang, ORCID: 0000-0002-2101-8918

Table of Contents

Analytical Data of polyproline tri-helix macrocycles

General methods for peptides analysis

All peptide products were purified by the analytical HPLC (Agilent Technology, 1260 Infinity or Agilent Technology, 1200 Infinity) with a Vydac C18 column (218TP54 4.6 mm x 250 mm) or a Agilent C8 column (Zorbax Rx-C8 4.6 mm x 250 mm). 0.1% TFA in water (solvent A), Acetonitrile (solvent B). After purification, the peptide products were confirmed by MALDI-TOF mass spectrometry (Bruker Daltonics, Autoflex III smartbeam LRF200-CID). Aviv Model 410 spectropolarimeter (Aviv Associates, Lakewood, NJ) was used for CD measurements. The peptides were dissolved in water or 1-propanol and incubated for more than 96 hour at room temperature before measurement. The solutions were measured in the cuvette with a path length of 1.0 mm and contained peptide solution (300 μL) at 25°C.

Data of peptide monomers: peptide alkynes

Peptide **9**, **10** and **20** were prepared according to the methods for solid phase peptide synthesis and polyproline C-terminus alkyne modification. Yields are based on quantitative Fmoc test and the isolated weight after lyophilization.

Peptide **9**

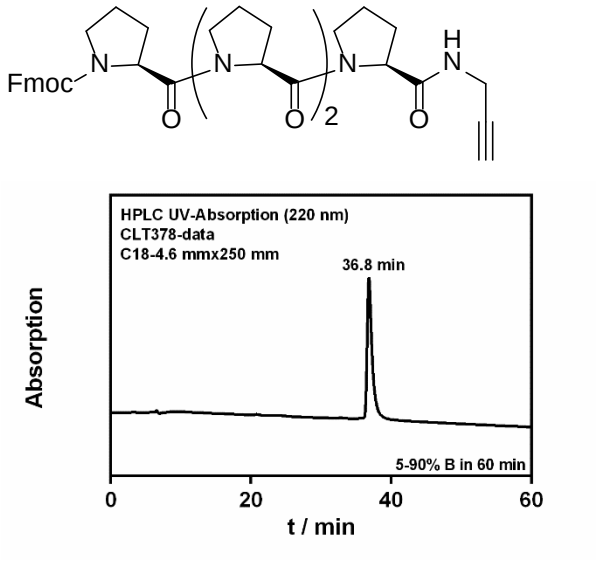


Figure S1. HPLC chromatogram of **9**

Data of peptide **9**

Yield: 121.2 mg, 59% (including SPPS)

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 36.8 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{38}H_{43}N_5NaO_6$: 688.311, found: 688.193

Peptide **10**

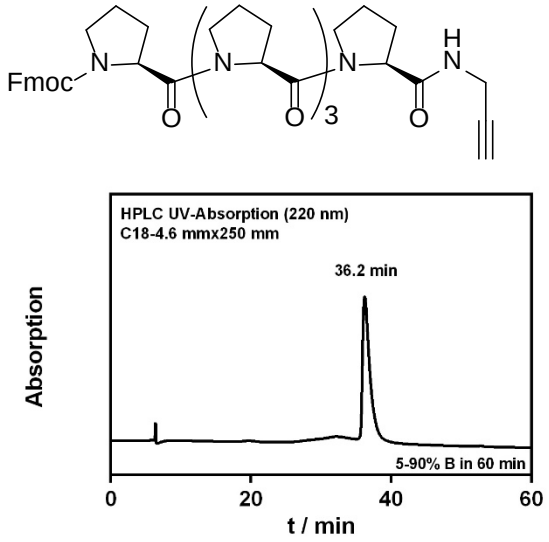


Figure S2. HPLC chromatogram of **10**

Data of peptide **10**

Yield: 11.3 mg, 36% (including SPPS)

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 36.2 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{43}H_{50}N_6NaO_7$: 785.363, found: 785.285

Peptide **20**

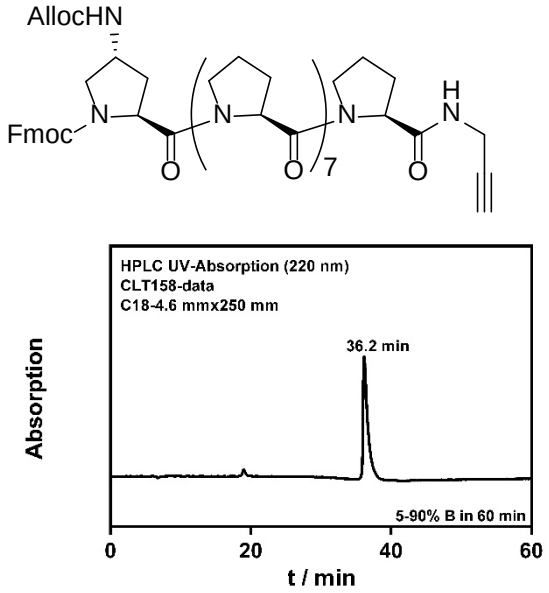


Figure S3. HPLC chromatogram of **20**

Yield: 261.66 mg, 89% (including SPPS)

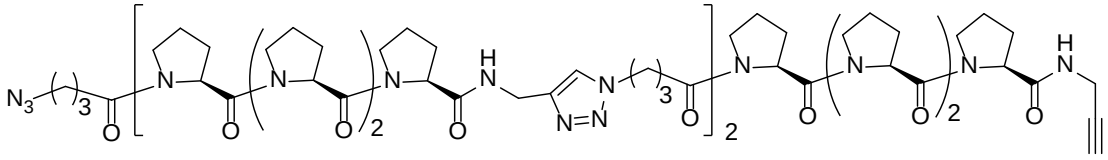
Analytical HPLC: 5% to 90% B over 60 min, 0.5 mL/min; t_R = 36.2 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{67}H_{83}N_{11}NaO_{13}$: 1272.606, found: 1272.319.

Data of linear peptide trimers: peptide alkynes

Peptide **17, 18** and **23** were prepared according to the methods for peptide acid synthesis, polyproline N-terminal azido modification and peptide assembly by CuAAc reaction on resins. Yields are based on quantitative Fmoc test and the isolated weight after lyophilization.

Peptide **17**



17

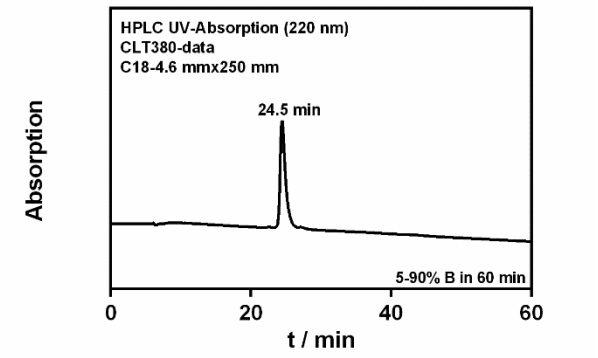


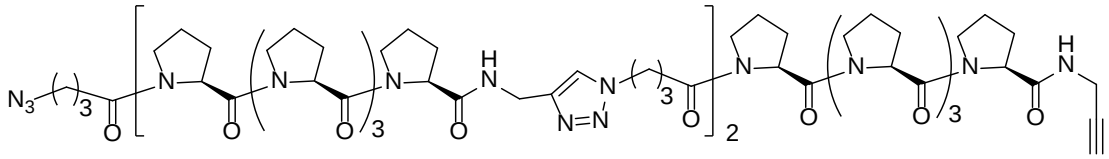
Figure S4. HPLC chromatogram of **17**

Yield: 30.7 mg, 68% (including SPPS)

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 24.5 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{81}H_{114}N_{24}NaO_{15}$: 1685.879, found: 1685.673

Peptide **18**



18

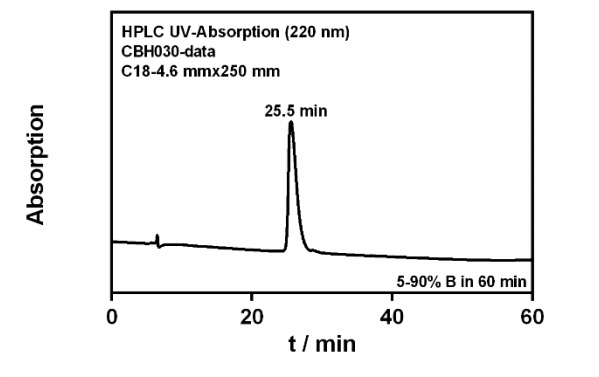


Figure S5. HPLC chromatogram of **18**

Yield: 1.36 mg, 23% (including SPPS)

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 25.5 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{96}H_{135}N_{27}NaO_{18}$: 1977.037, found: 1977.703

Peptide **23**

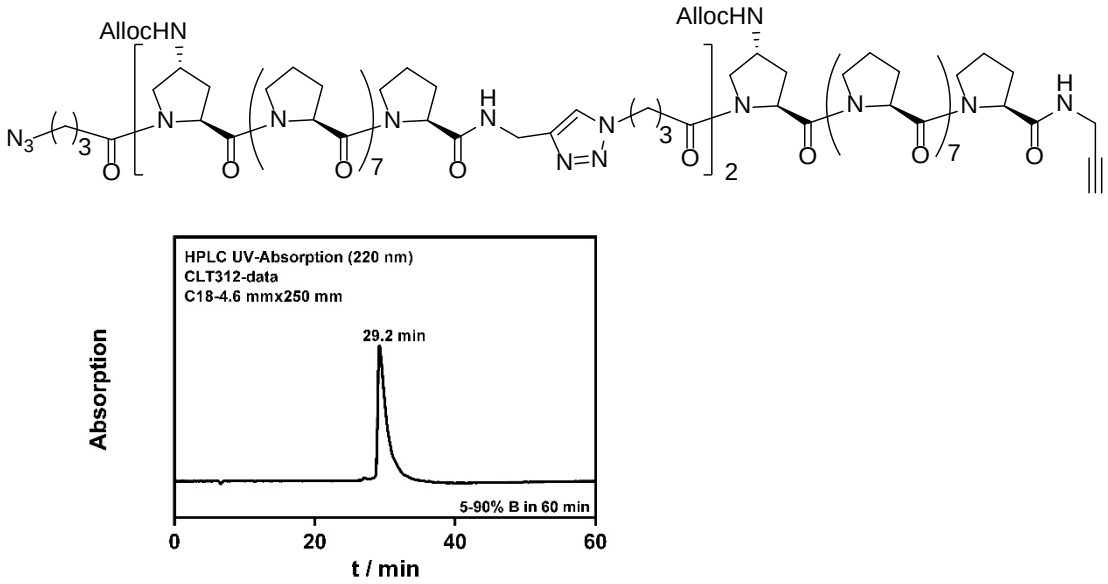


Figure S6. HPLC chromatogram of **23**

Yield: 33.83 mg, 78% (including SPPS)

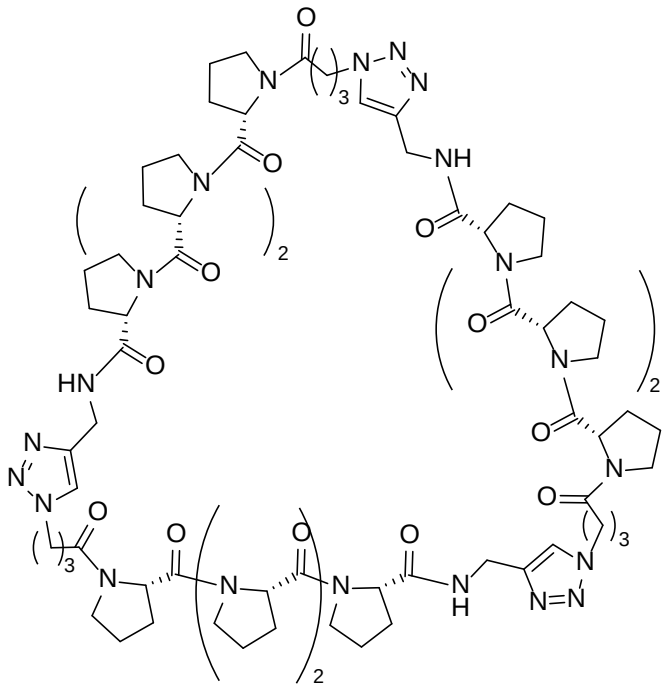
Analytical HPLC: 5% to 90% B over 60 min, 0.5 mL/min; t_R = 29.2 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{168}H_{234}N_{42}NaO_{36}$: 3438.766, found: 3438.366.

Data of polyproline tri-helix macrocyclics

Peptide **1, 2, 24, 26** and **27** were prepared according to the methods for peptide cyclization. Yields are based on the isolated weight after lyophilization.

Peptide **1**



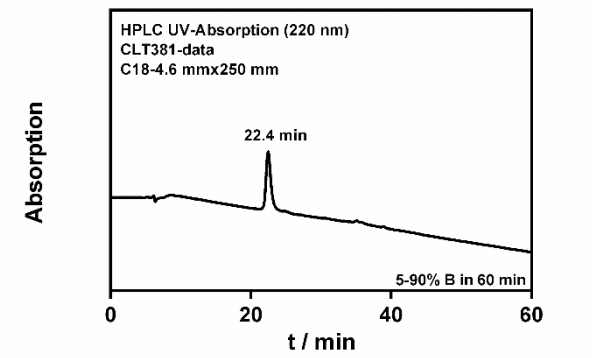


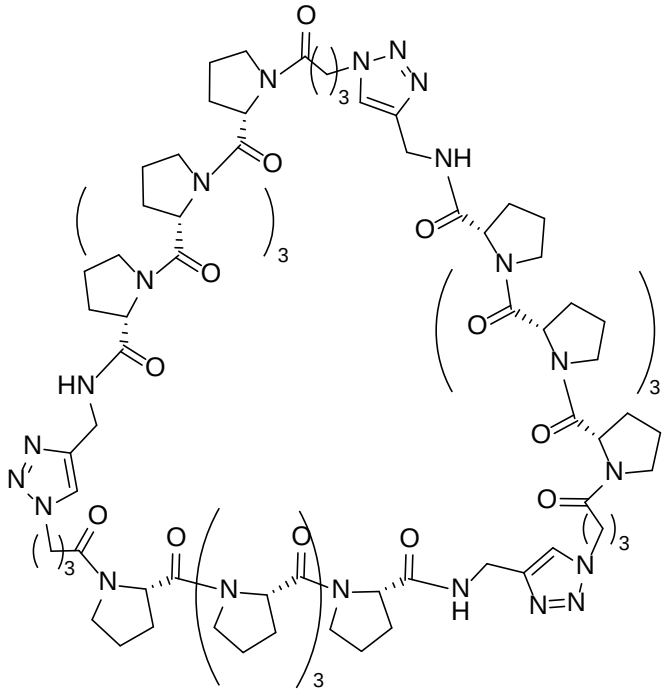
Figure S7. HPLC chromatogram of **1**

Yield: 6.4 mg, 43%

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 22.4 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{81}H_{114}N_{24}NaO_{15}$: 1685.879, found: 1685.856

Peptide **2**



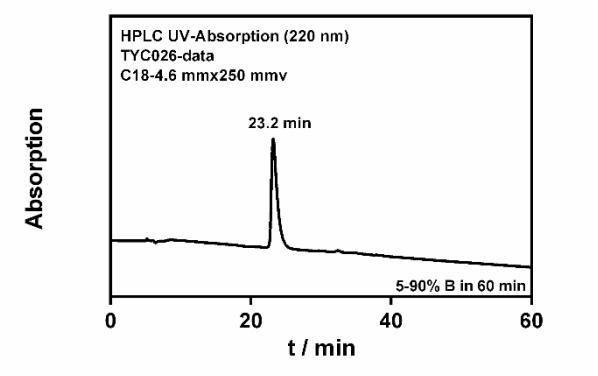


Figure S8. HPLC chromatogram of **2**

Yield: 0.3 mg, 22%

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 23.2 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{96}H_{135}N_{27}NaO_{18}$: 1977.037, found: 1977.376.

Peptide **24**

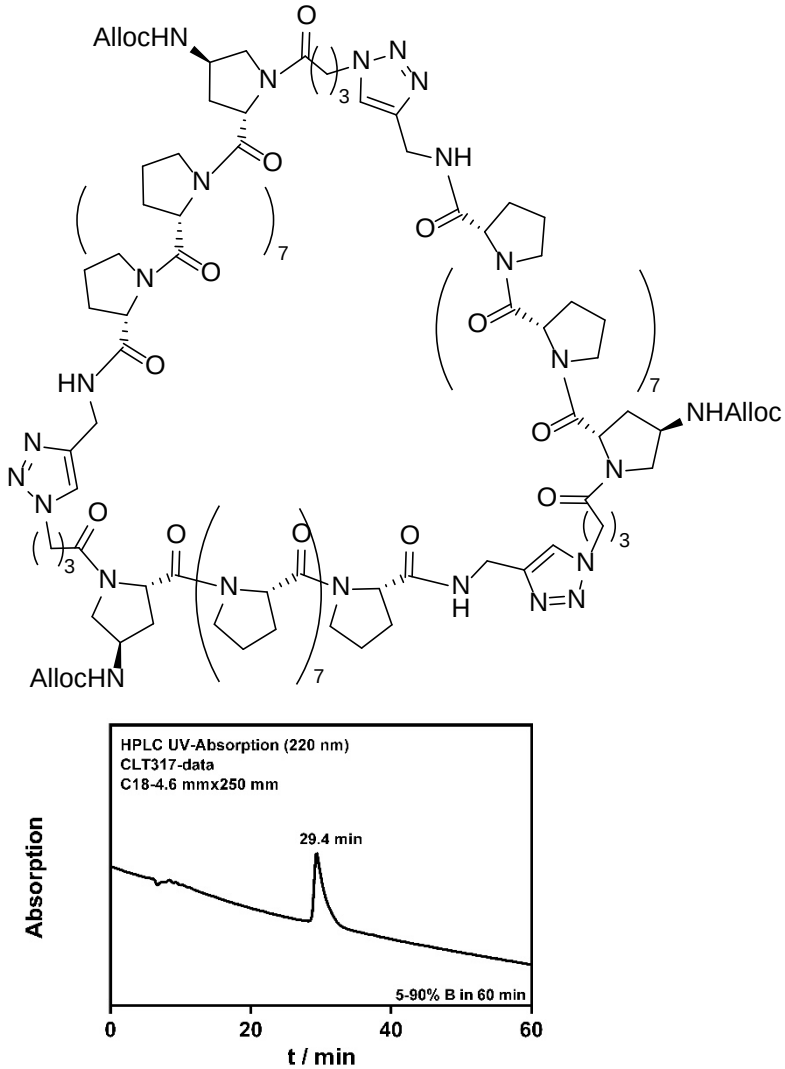


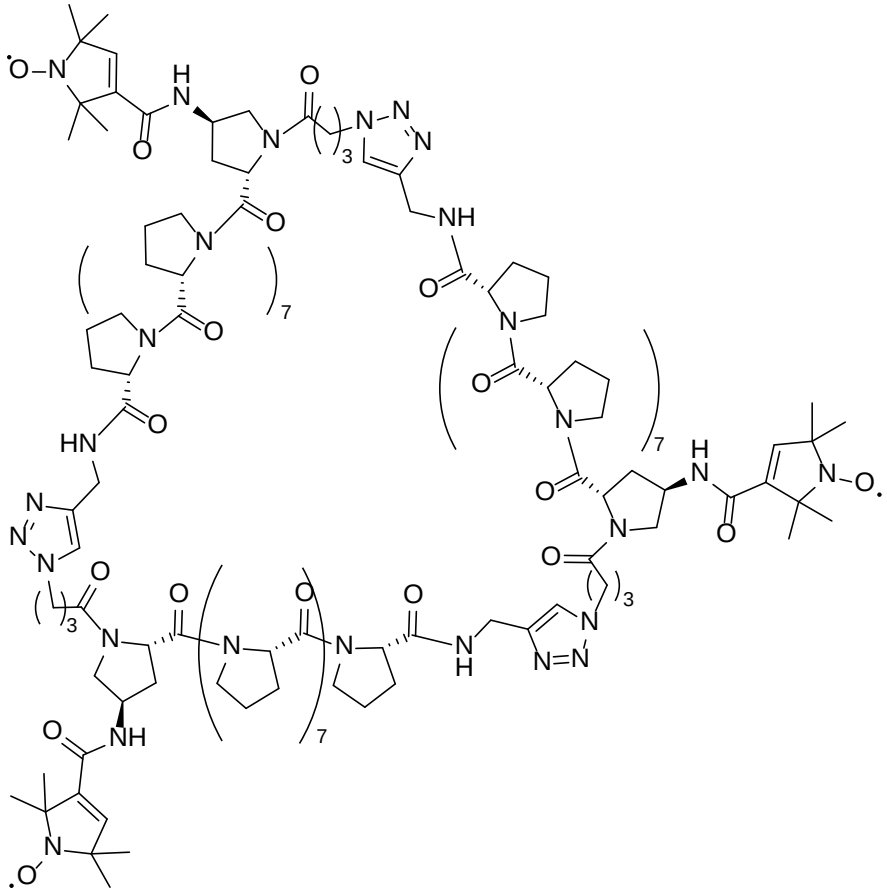
Figure S9. HPLC chromatogram of **24**

Yield: 4.06 mg, 75%

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 29.4 min.

MS(MALDI): $[M+Na]^+$ calcd. For $C_{168}H_{234}N_{42}NaO_{36}$: 3438.766, found: 3439.877.

Peptide **26**



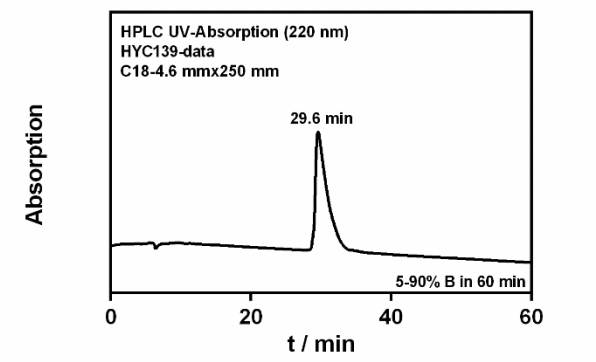


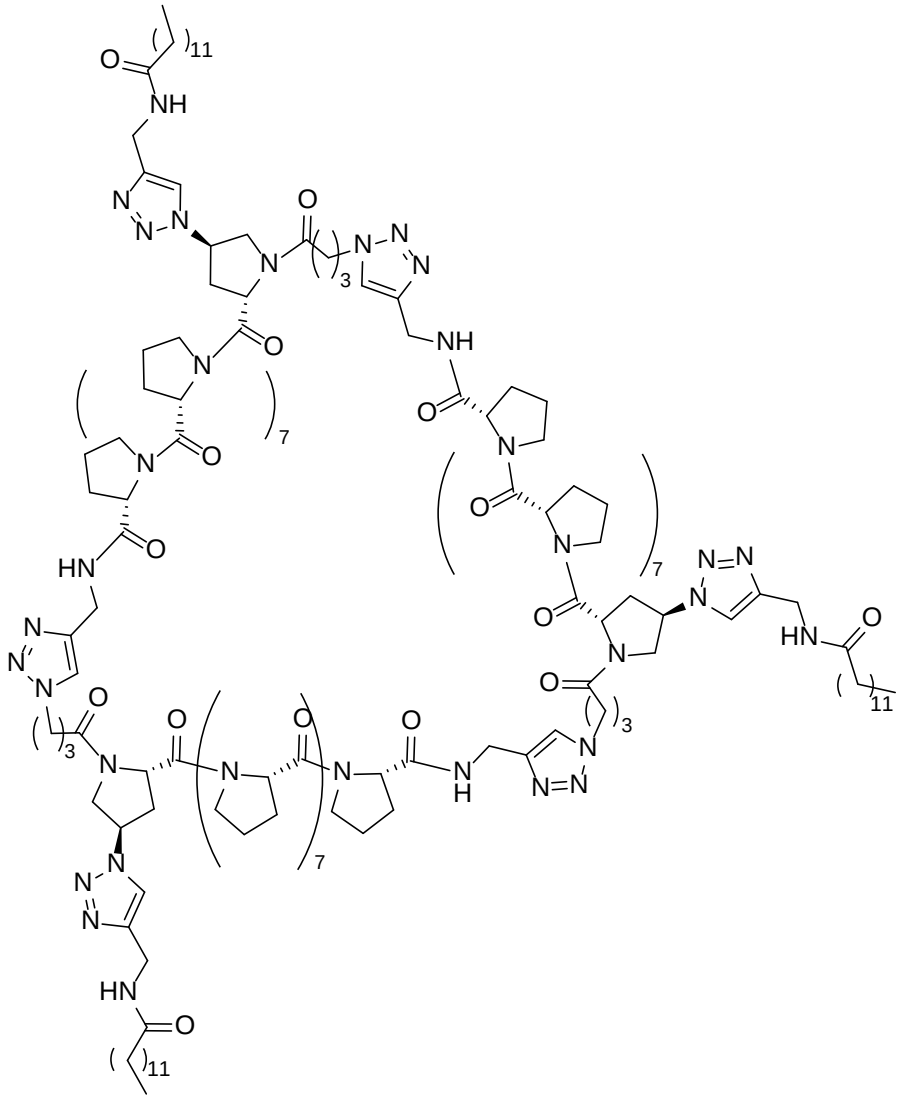
Figure S10. HPLC chromatogram of **26**

Yield: 0.72 mg, 22% (2 steps)

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 29.6 min.

MS(MALDI): $[M+Na]^{3\cdot+}$ calcd. For $C_{183}H_{258}N_{45}NaO_{36}$: 3684.963 found: 3687.107.

Peptide **27**



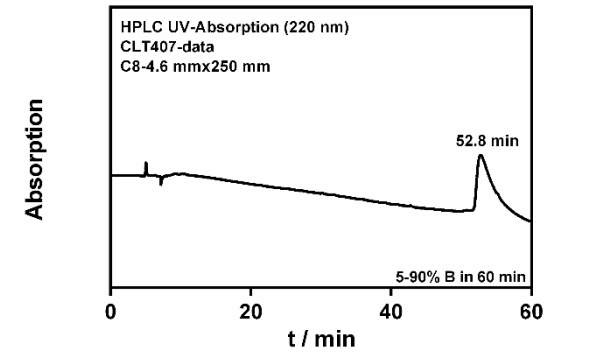


Figure S11. HPLC chromatogram of **27**

Yield: 0.79 mg, 53%

Analytical HPLC: 5% to 90% B (ACN) over 60 min, 0.5 mL/min; t_R = 52.8 min.

MS(MALDI): $[M+H]^+$ calcd. For $C_{204}H_{304}N_{51}O_{33}$: 3996.367, found: 3996.457.

Circular Dichroism Data

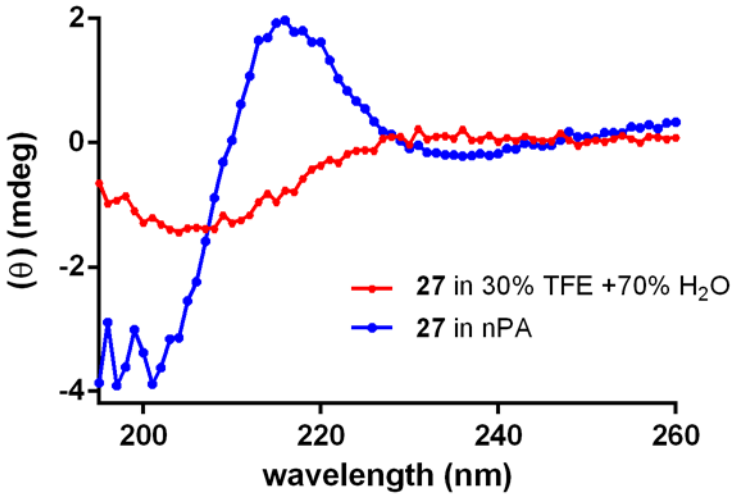


Figure S12. The CD spectra of **27**. measured in 70% water/30% TFE (red, 88 μM) and 1-propanol (blue, 83 μM)

Small molecule CuAAC reaction test

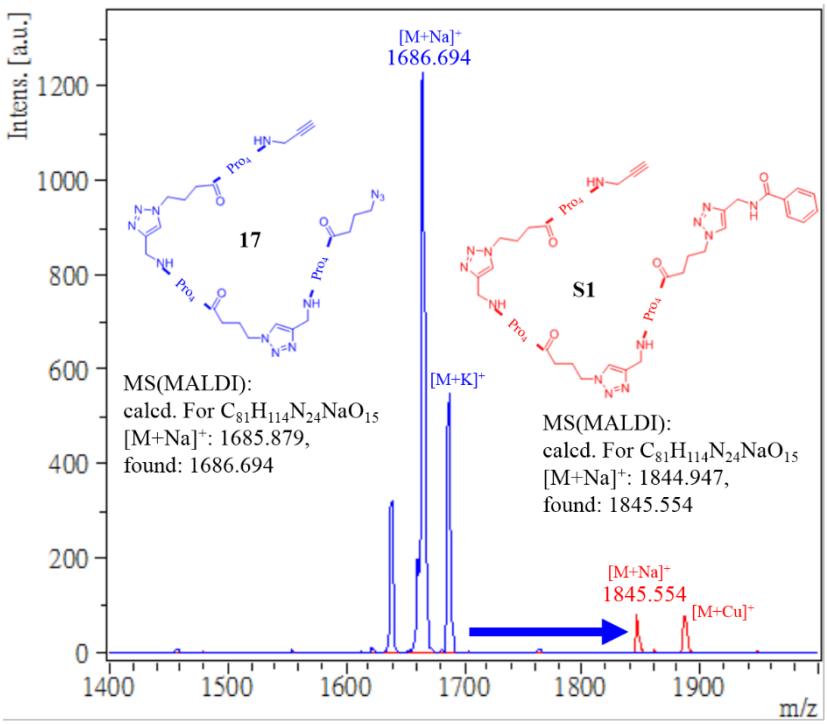


Figure S13. The MALDI mass spectrum of **17** (blue) and after (red) undergo CuAAC reaction with N-(prop-2-yn-1-yl)benzamide. The product **S1** is observed.

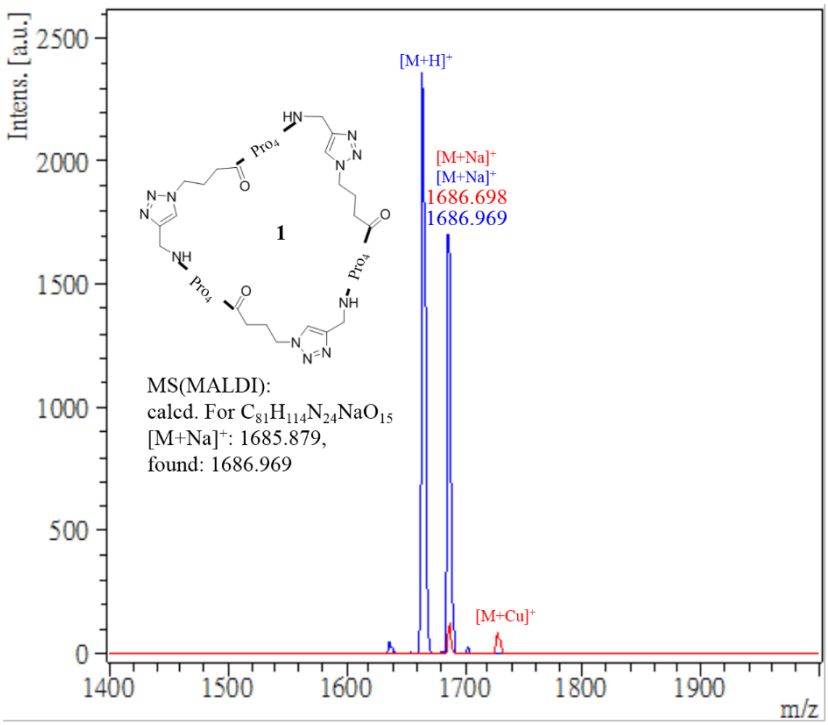


Figure S14. The MALDI spectrum of **1** (Blue) and after undergo CuAAC reaction with N-(prop-2-yn-1-yl)benzamide (Red). Only **1** is observed.

NMR Spectra

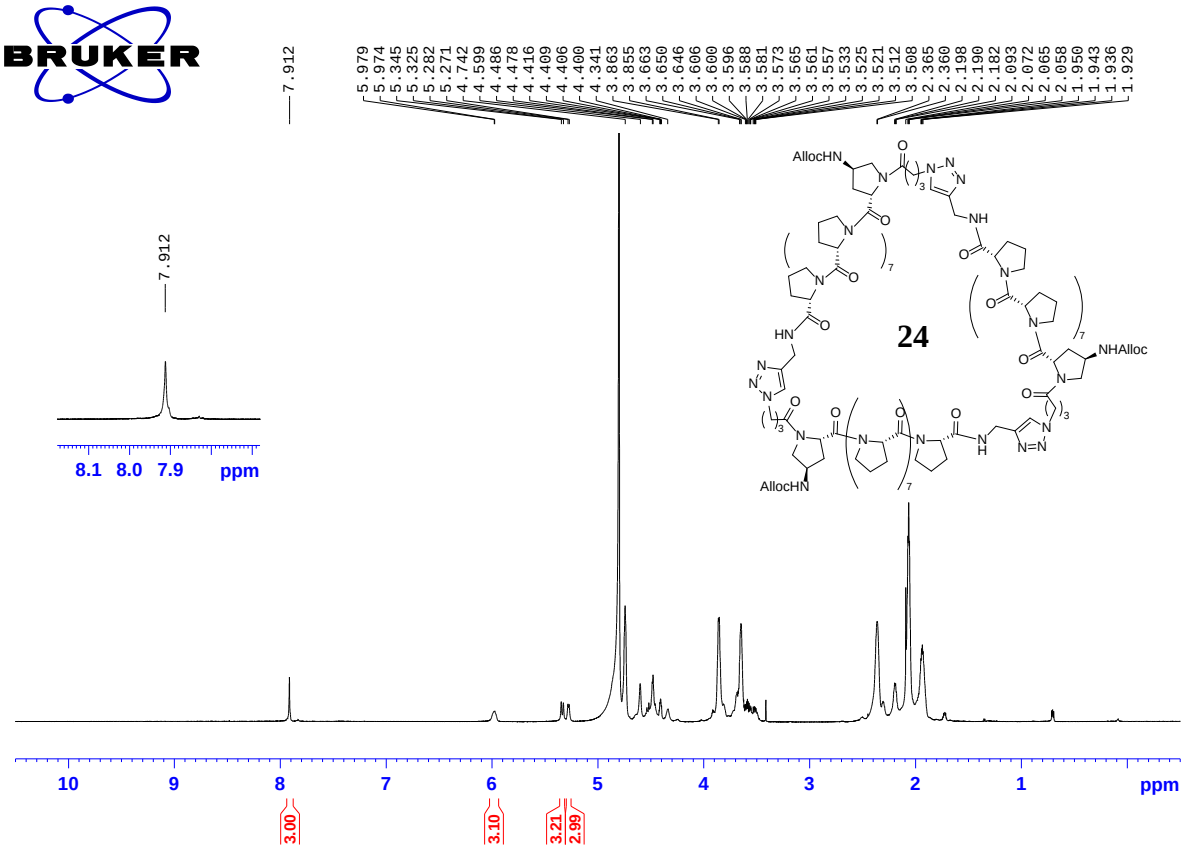


Figure S15. ^{1}H NMR spectrum of **24** (850 MHz, D_2O)

Model fitting of the SWAXS data

The data fitting discrepancy between theoretical $I(q)_{\text{the}}$ and experimental $I(q)_{\text{exp}}$ curves is described by χ^2 (goodness of fit) defined in the following

$$\chi^2 = \frac{1}{N_q}\sum_{i=1}^{Nq,i}\left\{\frac{I_{exp}(q_i)-I_{the}(q_i)}{\sigma(q_i)}\right\}^2$$

Here the symbols N_{q} is the total data points as a function of q_{i}, $\sigma(q_{\text{i}})$ the standard deviation of the i^{th} data point $\text{I}_{\text{exp}}(q_{\text{i}})$.

Table S1. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **6** at different d and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=8 Å	**13.5 ± 0.1 (2.139)**	**13.6 ± 0.1 (2.483)**	**13.7 ± 0.1 (2.971)**	**13.8 ± 0.1 (3.295)**	**NA**[a]	**NA**[a]	**NA**[a]	**13.9 ± 0.1 (3.419)**	**13.7 ± 0.1 (2.905)**	**13.7 ± 0.1 (2.515)**	**13.6 ± 0.1 (2.224)**
d=9 Å	**14.0 ± 0.1 (1.806)**	**14.2 ± 0.1 (2.270)**	**14.3 ± 0.1 (2.944)**	**14.3 ± 0.1 (3.661)**	**14.5 ± 0.2 (3.118)**	**14.6 ± 0.1 (3.136)**	**14.6 ± 0.1 (3.333)**	**14.6 ± 0.1 (3.778)**	**14.5 ± 0.1 (2.964)**	**14.4 ± 0.1 (2.405)**	**14.4 ± 0.1 (1.961)**
d=10 Å	**14.7 ± 0.1 (2.098)**	**14.9 ± 0.1 (2.289)**	**14.9 ± 0.1 (3.023)**	**15.2 ± 0.1 (3.471)**	**15.2 ± 0.1 (3.720)**	**15.3 ± 0.1 (3.869)**	**15.2 ± 0.2 (3.678)**	**15.2 ± 0.1 (3.457)**	**15.2 ± 0.2 (2.649)**	**15.2 ± 0.1 (2.197)**	**15.0 ± 0.1 (2.081)**

[a] Not available as collision occurs between polyproline helices

Table S2. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **5** at different d and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=8 Å	**12.6 ± 0.1 (3.820)**	**12.9 ± 0.1 (3.308)**	**12.8 ± 0.1 (3.118)**	**13.2 ± 0.1 (2.945)**	**13.3 ± 0.1 (2.936)**	**13.3 ± 0.1 (3.102)**	**13.4 ± 0.1 (3.116)**	**13.3 ± 0.1 (3.576)**	**13.1 ± 0.1 (3.701)**	**13.0 ± 0.1 (3.248)**	**12.6 ± 0.1 (3.281)**
d=9 Å	**13.5 ± 0.1 (2.105)**	**13.6 ± 0.1 (2.112)**	**13.6 ± 0.1 (2.354)**	**13.7 ± 0.1 (2.818)**	**13.9 ± 0.2 (2.983)**	**13.9 ± 01 (3.270)**	**13.9 ± 0.1 (3.304)**	**14.0 ± 0.1 (3.013)**	**13.8 ± 0.1 (2.636)**	**13.6 ± 0.1 (2.143)**	**13.4 ± 0.1 (2.272)**
d=10 Å	**14.2 ± 0.1 (2.963)**	**14.3 ± 0.1 (2.828)**	**14.3 ± 0.1 (3.104)**	**14.3 ± 0.1 (3.753)**	**14.5 ± 0.1 (4.041)**	**14.5 ± 0.1 (4.165)**	**14.7 ± 0.1 (3.887)**	**14.7 ± 0.1 (3.669)**	**14.4 ± 0.2 (3.284)**	**14.2 ± 0.1 (3.069)**	**14.2 ±0.1 (3.258)**

Table S3. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **4** at different *d* and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=7 Å	**11.6 ± 0.1 (1.717)**	**11.6 ± 0.1 (1.830)**	**11.9 ± 0.1 (1.872)**	**11.9 ± 0.1 (1.810)**	**12.3 ± 0.1 (1.817)**	**12.2 ± 0.1 (1.823)**	**12.2 ± 0.1 (1.801)**	**12.0 ± 0.1 (1.772)**	**11.9 ± 0.1 (1.792)**	**11.9 ± 0.1 (1.827)**	**11.7 ± 0.1 (1.778)**
d=8 Å	**12.3 ± 0.1 (1.639)**	**12.4 ± 0.1 (1.683)**	**12.7 ± 0.2 (1.752)**	**12.8 ± 0.1 (1.788)**	**12.7 ± 0.1 (1.852)**	**12.8 ± 01 (1.861)**	**12.7 ± 0.1 (1.839)**	**12.5 ± 0.2 (1.787)**	**12.5 ± 0.1 (1.712)**	**12.5 ± 0.1 (1.631)**	**12.4 ± 0.1 (1.620)**
d=9 Å	**12.8 ± 0.1 (1.829)**	**13.1 ± 0.1 (1.875)**	**13.4 ± 0.1 (1.951)**	**13.5 ± 0.2 (1.996)**	**13.4 ± 0.1 (2.032)**	**13.3 ± 0.1 (2.041)**	**13.3 ± 0.1 (2.030)**	**13.2 ± 0.2 (1.977)**	**13.1 ± 0.1 (1.899)**	**13.1 ± 0.1 (1.874)**	**13.0 ±0.1 (1.855)**

Table S4. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **3** at different *d* and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=6 Å	**10.4 ± 0.1 (3.278)**	**10.6 ± 0.1 (3.282)**	**10.6 ± 0.1 (3.315)**	**10.7 ± 0.1 (3.350)**	**10.8 ± 0.1 (3.319)**	**NA**[a]	**11.1 ± 0.1 (3.523)**	**11.0 ± 0.1 (3.431)**	**10.9 ± 0.1 (3.279)**	**10.7 ± 0.1 (3.327)**	**10.7 ± 0.1 (3.495)**
d=7 Å	**11.2 ± 0.1 (3.743)**	**11.3 ± 0.1 (3.447)**	**11.3 ± 0.1 (3.509)**	**11.4 ± 0.1 (3.701)**	**11.5 ± 0.1 (3.862)**	**11.6 ± 01 (3.912)**	**11.7 ± 0.1 (3.919)**	**11.6 ± 0.1 (3.657)**	**11.5 ± 0.1 (3.608)**	**11.4 ± 0.1 (3.350)**	**11.4 ± 0.1 (3.562)**
d=8 Å	**11.9 ± 0.1 (5.951)**	**12.0 ± 0.1 (5.177)**	**12.0 ± 0.1 (5.159)**	**12.1 ± 0.1 (5.381)**	**12.2 ± 0.1 (5.579)**	**12.3 ± 0.1 (5.650)**	**12.3 ± 0.1 (5.424)**	**12.2 ± 0.1 (5.581)**	**12.0 ± 0.1 (5.680)**	**12.2 ± 0.1 (4.827)**	**12.1 ±0.1 (5.681)**

[a] Not available as collision occurs between polyproline helices

Table S5. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **2** at different *d* and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=5 Å	**9.9 ± 0.1 (1.211)**	**9.9 ± 0.1 (1.181)**	**9.6 ± 0.1 (1.195)**	**9.7 ± 0.1 (1.181)**	**9.7 ± 0.1 (1.178)**	**9.7 ± 0.1 (1.177)**	**9.7 ± 0.1 (1.180)**	**9.7 ± 0.1 (1.172)**	**9.5 ± 0.1 (1.195)**	**9.4 ± 0.1 (1.211)**	**9.5 ± 0.1 (1.217)**
d=6 Å	**10.0 ± 0.1 (1.238)**	**10.0 ± 0.1 (1.221)**	**10.1 ± 0.1 (1.196)**	**10.2 ± 0.1 (1.180)**	**10.3 ± 0.1 (1.176)**	**10.2 ± 01 (1.178)**	**10.3 ± 0.1 (1.171)**	**10.3 ± 0.1 (1.178)**	**10.2 ± 0.1 (1.190)**	**10.0 ± 0.1 (1.208)**	**9.8 ± 0.1 (1.258)**
d=7 Å	**10.4 ± 0.1 (1.304)**	**10.6 ± 0.1 (1.270)**	**10.7 ± 0.1 (1.257)**	**10.8 ± 0.1 (1.236)**	**10.9 ± 0.1 (1.226)**	**10.8 ± 0.1 (1.227)**	**10.9 ± 0.1 (1.239)**	**10.9 ± 0.1 (1.228)**	**10.8 ± 0.1 (1.238)**	**10.7 ± 0.1 (1.276)**	**10.6 ±0.1 (1.299)**

Table S6. The calculated radius of gyration (R_g) and χ^2 (in parentheses) of scaffold **1** at different *d* and θ.

	Incline angle θ										
	-25°	-20°	-15°	-10°	-5°	0°	+5°	+10°	+15°	+20°	+25°
d=5 Å	**8.9 ± 0.1 (2.448)**	**8.9 ± 0.1 (2.274)**	**9.0 ± 0.1 (1.933)**	**9.1 ± 0.1 (1.741)**	**9.0 ± 0.1 (1.669)**	**9.1 ± 0.1 (1.689)**	**9.0 ± 0.1 (1.727)**	**9.0 ± 0.1 (2.001)**	**8.9 ± 0.1 (2.205)**	**9.1 ± 0.1 (2.333)**	**8.8 ± 0.1 (2.250)**
d=6 Å	**9.4 ± 0.1 (1.704)**	**9.4 ± 0.1 (1.715)**	**9.5 ± 0.1 (1.713)**	**9.4 ± 0.1 (1.758)**	**9.4 ± 0.1 (1.816)**	**9.3 ± 01 (1.765)**	**9.3 ± 0.1 (1.730)**	**9.4 ± 0.1 (1.659)**	**9.2 ± 0.1 (1.882)**	**9.4 ± 0.1 (1.570)**	**9.3 ± 0.1 (1.562)**
d=7 Å	**9.8 ± 0.1 (1.631)**	**9.9 ± 0.1 (1.646)**	**9.9 ± 0.1 (1.688)**	**9.8 ± 0.1 (1.732)**	**9.8 ± 0.1 (1.757)**	**10.0 ± 0.1 (1.793)**	**9.8 ± 0.1 (1.677)**	**9.7 ± 0.1 (1.527)**	**9.8 ± 0.1 (1.628)**	**9.7 ± 0.1 (1.714)**	**9.8 ±0.1 (1.689)**

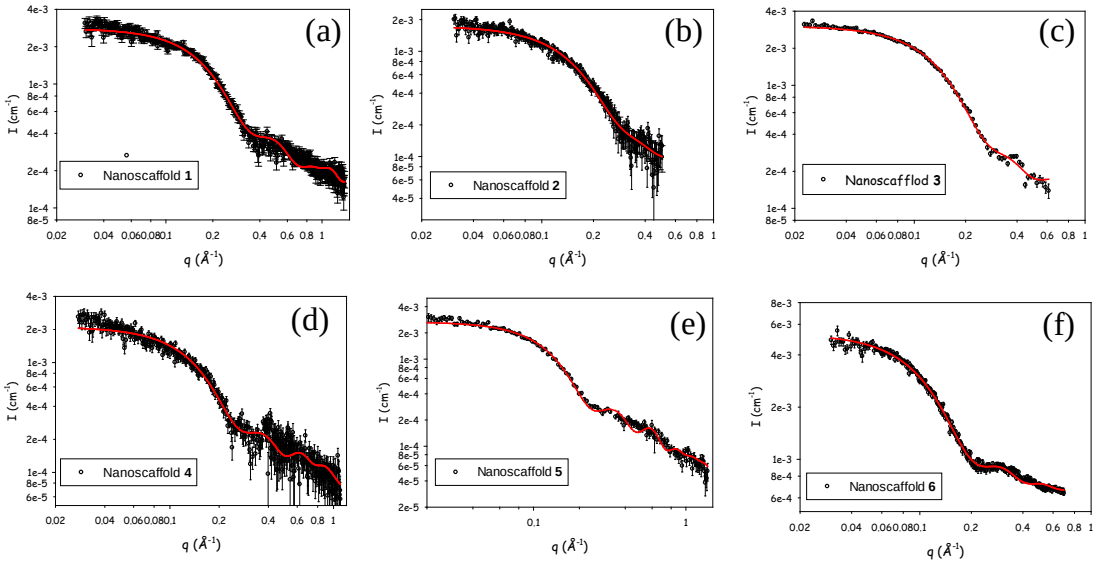


Figure S16. The extended SWAXS data covering a wider q-range of scaffolds **1-6** than that shown Figure 3. The data measured in water could also be fitted well using the models optimized with the data shown in Figure 3. The results further support the optimized models.

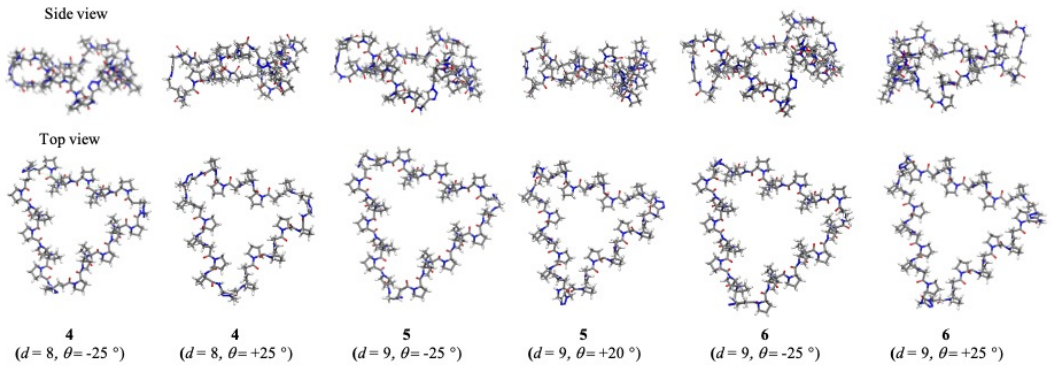


Figure S17. The possible structures of scaffolds **4–6**, with the d and θ values shown below, that can fit well the corresponding SWAXS data. The parameter of the common rotation angle ϕ is set to 0 for all the structures presented.

Additional STM images

Regarding features lack of 3 lobes in Figure 7a, it is likely that there were scaffolds interacting weakly with the substrate due to unfavorably organized shapes and adsorption orientation. Therefore, only one of the lobes is observed. Other possibilities are not ruled out.

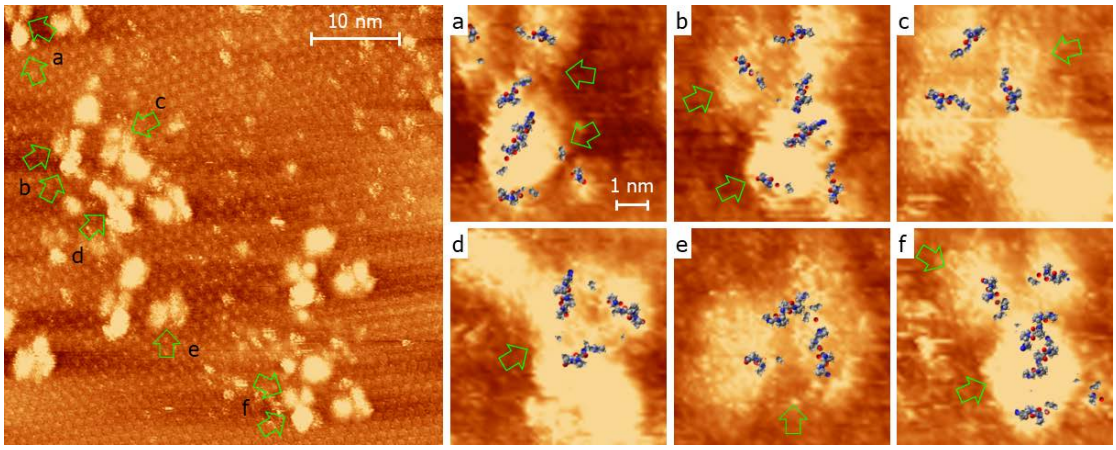


Figure S18. A 50 x 50-nm STM image exhibiting 3-lobe features of Scaffold **27**. (a-f) Magnified 3 x 3-nm images overlaid with models to illustrate the morphology of the 3-lobe features. The upper portion of the models is sliced away to reveal possible segments that contact directly with the substrate and confer electron tunneling more dominantly.[1] Note that the 3-lobe features may arise from two Scaffold **27**. The molecules were trapped on the substrate by the porous TMA (trimesic acid) template.[2] Nonetheless, those Scaffold **27** not affixing perfectly might result in ill-defined features. Imaging conditions: $i_{tunneling}$, 30 pA; V_{bias}, −0.9 V (tip-grounded)

Reference

1. Yang HC, Lin SY, Yang HC, Lin CL, Tsai L, Huang SL, Chen IW, Chen CH, Jin BY, Luh TY. Molecular architecture towards helical double-stranded polymers. *Angew Chem Int Ed.* **2006**, *45*, 726-730.
2. Cheng KY, Lee SL, Kuo TY, Lin CH, Chen YC, Kuo TH, Hsu CC, Chen CH. Template-Assisted Proximity for Oligomerization of Fullerenes. *Langmuir* **2018**, *34*, 5416-5421.