Supporting Information

Hierarchical protein self-assembly into dynamically controlled 2D nanoarrays via host-guest chemistry

Yijia Li,[†1] Linlu Zhao,[†2] Hongwei Chen,[2] Ruizhen Tian,[1]Fei Li,[1] Quan Luo,[1] Jiayun Xu,[1,3] Chunxi Hou,[*1] Junqiu Liu[*1,3]

[1] State Key Laboratory of Supramolecular Structure and Materials, College of Chemistry, Jilin University, 2699 Qianjin Road, Changchun 130012 China

[2] The First Affiliated Hospital of Hainan Medical University, College of Emergency and Trauma, Hainan Medical University, Haikou 571199, China

[3] College of Material, Chemistry and Chemical Engineering, Hangzhou Normal University, Hangzhou 311121, China

Contents

1. General Information

The plasmid pET22b-SP1 provided by Professor Oded Shoseyov (The Hebrew University of Jerusalem, Israel) was used as the template for PCR cloning to obtain the mutant plasmids pET22b-SP1S98C, respectively. The resulting plasmids were confirmed by DNA sequencing and then transformed into Escherichia coli BL21 (DE3) for protein expression.[1-2]

2. Experimental Methods

Computer Simulation

The protein structure model was output by a computer software named PyMOL. This program can produce high-quality 3D images of small molecules and biological macromolecules and can help us observe the distribution of different amino acids on the surface of proteins. After getting the PDB code on NCBI website, we can find the crystal structure of natural SP1 (SP1W). Upon observation of the distributed amino acids and their bond directions on the surface of SP1W, proper site was chosen and mutated in the program. Then a mutant model protein structure was successfully obtained.

Construction, Expression and Purification of SP1S98C Variant

Depending on the structure of SP1W, twelve serine sites with C_6 symmetry at the lateral surface were selected and then mutated into cysteine. The SP1S98C gene was achieved by site-directed mutagenesis using a sense primer with the sequence 5'-CCTACTTTGTGCCAGCGTCTTG-3' and an antisense primer with the sequence 5'-AAGACGCTGGCACAAAGTAGGC-3'. An expression plasmid of pET-22b-SP1S98C was extracted from DH5α *Escherichia coli* strains. The mutant plasmids were confirmed *via* genetic sequencing and subsequently transformed into *E.coli* stains BL21 for protein expression. The *E.coli* strains BL21 were cultured in LB liquid media with 100 μg/ml ampicillin at 37 ℃ until the value of OD_{600} reached 0.7. Then 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce protein expression. The cells were cultured for 4 hours continuously at 28 ℃. The cultures were centrifuged at 8000 rpm for 15 min to harvest the cells and stored at -20 ℃, respectively.

For purification, cells were suspended by Tris-HCl buffer (20 mM, pH=6.3) and then sonicated.

Further centrifugation at 15000 rpm for 30 min and then heating at 85 ℃, then the cultures were centrifuged once more and discarded sediment. The supernatant was further purified by DEAE ion-exchange column with a serious of concentration of NaCl solution to remove the impurity proteins. And then Sephadex G75 column was used to remove nucleic acids.

Supramolecular interactions induced assembly and disassembly of 2D protein nanosheets

The purified SP1S98C variant was dissolved in PBS Buffer (20 mM, pH=7.4) reaching a final concentration of 10^{-5} M. Stirred with 10-fold excess **MMV**$^{+}$ at room temperature overnight and then dialyzed at 4 ℃ for 24 h, SP1S98C-MMV^{+} candidates were obtained.

For assembly, we incubated CB[8] and SP1S98C-MMV^{+} at the ratio of 5:1, 1:1, 1:2 and 1:5 (SP1S98C-MMV^{+}/CB[8]) in PBS buffer (5 mM, pH=7.4) at room temperature overnight (or 4 ℃, 2 days), respectively. For disassembly, we added 10-fold excess FGG guests to the solution of assemblies and stirred for 5 h, respectively.

For reassembly, we dialyzed the solution of disassemblies containing complexes of CB[8] and FGG molecules and dispersed high molecular-weighed SP1S98C-MMV^{+} for 2 days at 4 ℃ to remove FGG_2&CB[8] complexes. Concentrating the solution by differential centrifugation to make the concentration attenuated by dialysis back to the original value. Then we introduced CB[8] hosts to the solution once more according to previous proportion (SP1S98C-MMV^{+}/CB[8]=1:5) and incubated at room temperature overnight (or 4 ℃, 2 days), respectively.

DLS measurement

Dynamic light scattering (DLS) measurement was carried out to monitor the process of SP1S98C-based protein assembly and disassembly with Malvern Instrument Zetasizer Nano ZS instrument at 25 ℃. Samples were dissolved in PBS buffer (10 mM, pH=7.0), mixed at designed corresponding ratios, and then monitored by DLS with different duration of assembly and disassembly, respectively.

Atomic force microscopy characterization

Atomic force microscopy (AFM) characterization by NanoScope Multimode AFM in tapping-mode was used to perform the morphologies of samples. We dropped 10 μL of 0.005 mg/ml

sample solution on a freshly cleaved silicon wafer for 5 min, followed by washing the sample-covered wafer with Milli-Q water and then drying it under air for AFM imaging.

Transmission electron microscopy characterization

The images of samples were captured by a JEM-2100F transmission electron microscopy (TEM) instrument with an accelerating voltage of 200 kV. We dropped 10 μL sample on a carbon-coated copper grid and waited the grid to absorb for 10 min. The excess solution of samples was taken away. Then the samples were negatively stained for 40 s by 1% sodium phosphotungstate solution, and the excess stain was blotted away. Drying the stained-sample-covered grid under air overnight before imaging.

3. SDS-PAGE analysis of protein SP1 variant (S98C)

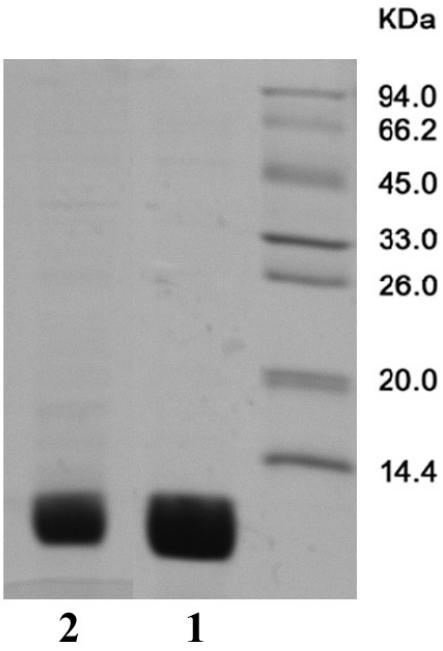


Figure S1. SDS-PAGE analysis of protein SP1 variant (S98C). The acrylamide percentage of the SDS-PAGE gels is 18%. Lane 1: purified SP1W; Lane 2: purified SP1S98C.

4. CD spectral data of protein SP1W and its variant SP1S98C

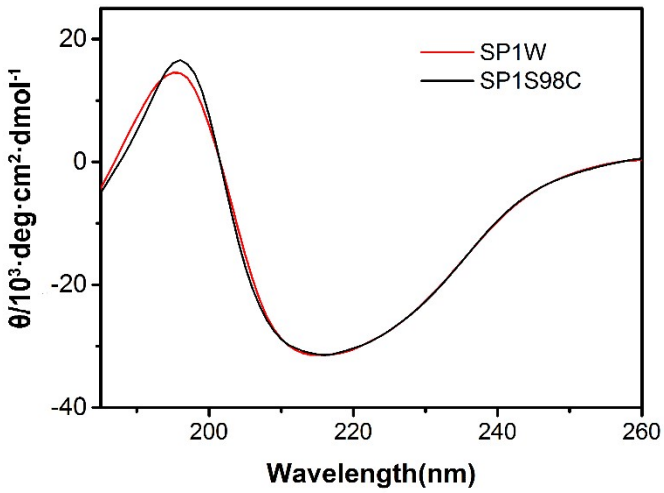


Figure S2. CD spectral data for SP1W and SP1S98C monitored from 185 nm to 260 nm, respectively (Red line: SP1W; Black line: SP1S98C).

5. Chemicals and methods

5.1 Chemicals

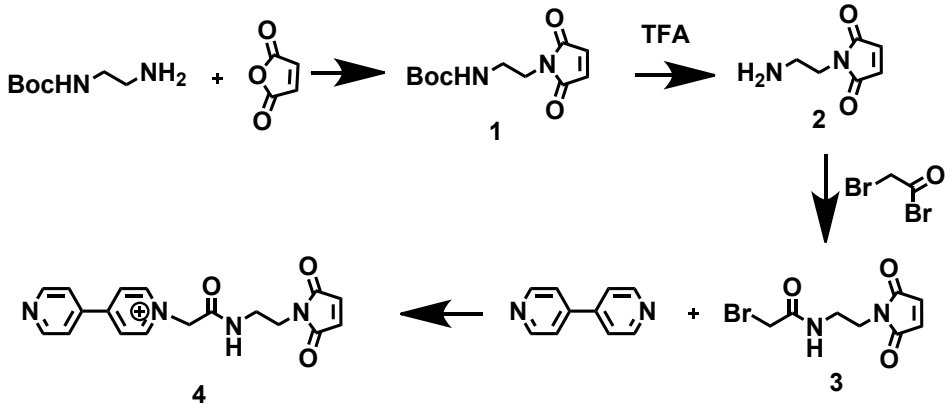
Maleic anhydride, N-Boc-ethylenediamine and bromoacetyl bromide were purchased from Aladdin Co. LLC. CB[8] was purchased from Sigma-Aldrich Co. LLC. FGG was synthesized by peptide synthesis company. Trifluoroacetic acid (TFA), acetic anhydride, acetic acid, acetic acid-sodium salt trihydrate, triethylamine, acetone and dichloromethane were obtained from Beijing Chemical Plant and drying before use. Other chemicals were used without further purification.

5.2 Instruments

Column chromatography was performed with silica gel (200-300 mesh). ^{1}H-NMR spectra was performed on a Bruker AVANCE III 500 apparatus using DMSO or $CDCl_3$ as solvent with the proton signal of tetramethylsilane (TMS) as reference. ESI-MS spectrometric analyses were performed using the Thermo Finnigan LCQ AD System.

6. Synthesis of the maleimide-modified methyl viologen (MMV^+)

6.1 Scheme of synthetic route of MMV^+



Scheme S1. Synthetic route of guest complex 4 (**MMV^+**). The complexes 1-4 following refer to molecules labeled 1-4 in this figure.

6.2 Synthetic methods

Maleic anhydride (2.3726 g) was suspended in 50ml dried dichloromethane. N-Boc-ethylenediamine (3.5 g) was dissolved in 70ml dried dichloromethane and removed into constant pressure dropping funnel. Vacuumized, then injected with nitrogen. N-Boc-ethylenediamine was slowly added to the maleic anhydride under ice-bath. The mixture was incubated at room temperature for 3 h and then evaporated under reduced pressure. The dried product was dissolved by 150 ml acetone, and then added acetic anhydride (2.64 ml), acetic acid-sodium salt trihydrate (1.496 g) and triethylamine (594 μl). The mixture then was heated to 60 ℃ and refluxed overnight protected by nitrogen. The excessive acetone was evaporated under reduced pressure to afford yellow oil. The oil was added into ice water and static settled overnight to obtain yellow solid sediment. The crude product was further purified by silica column chromatography to achieve white solid (**complex 1**).

Complex 1 (1.76 g) was dissolved by 50ml dichloromethane. Then added trifluoroacetic acid (TFA) (3 ml) into the solution and stirred overnight. After the reaction was completed, it was pumped dry with an oil pump and washed twice with ether. The product (**complex 2**) is obtained after thorough drying.

Dried triethylamine (3 μl) was slowly added into 7 ml dried dichloromethane-dissolved complex

2 (200 mg) under ice-bath and reacted for 40 min. Mixed bromoacetyl bromide (103 μl) with dried dichloromethane (7 μl). The mixture was slowly added into the solution of complex 2 and reacted overnight protected by nitrogen. The excessive solvent was evaporated under reduced pressure to afford brown solid. The crude product was further purified by silica column chromatography to achieve **complex 3**.[3]

The complex 3 (0.522 g, 2 mmol) was added into 20 ml dried dichloromethane with 1.562 g 4, 4'-bipyridine and ultrasonically dissolved. The reaction mixture was stirred at 373 K overnight to gain yellow precipitation. After the reaction was completed, the solvents were removed under reduced pressure to afford yellow solid product. The gained product then was dissolved into 50ml distilled water and was filtered to remove insolubles. Then the water was removed under reduced pressure to afford light yellow solid **complex 4**. Placed the product in a vacuum oven to dry overnight and then we gained our target molecule (**MMV**$^{+}$) (0.8 g, 96.4%).

6.3 ^{1}H-NMR spectra and ESI-MS data of complex 4

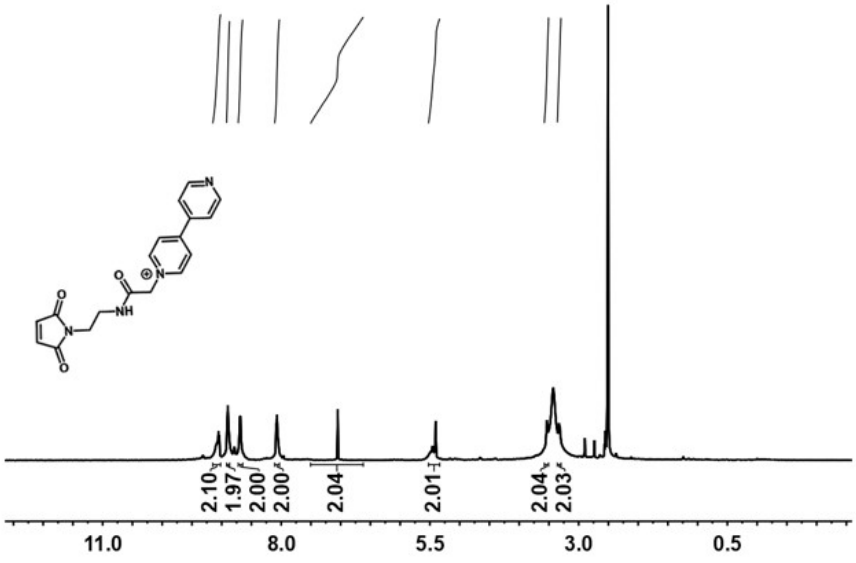


Figure S3. ^{1}H-NMR spectrum of complex 4. ^{1}H-NMR (500 MHz, DMSO) δ 9.15 – 9.02 (m, 2H), 8.90 (s, 2H), 8.69 (d, J = 6.3 Hz, 2H), 8.08 (d, J = 5.2 Hz, 2H), 7.05 (s, 2H), 5.43 (d, J = 27.0 Hz, 2H), 3.58 – 3.51 (m, 2H), 3.32 (d, J = 5.9 Hz, 2H).

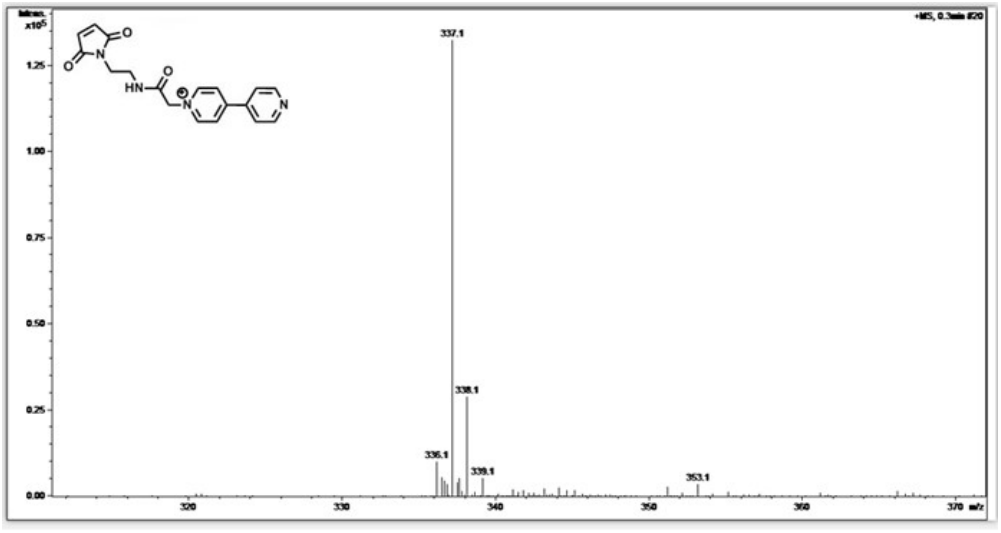


Figure S4. ESI-MS characterization of complex 4.

7. Investigation of the stoichiometry of guest MMV^+ with CB[8] via ^{1}H-NMR spectra

A stock solution of MMV^+ in D_2O was prepared. To 400 μL aliquots of MMV^+ solution were added different equiv of CB[8] solid sample, respectively. After ultrasonic stirring and incubating the corresponding solutions were introduced into the NMR tube. The ^{1}H-NMR spectra (Figure S5) were recorded to investigate the molecular binding behaviors between MMV^+ and CB[8]. Figure S5(a) showed ^{1}H-NMR titration spectra of MMV^+ without CB[8]. When 0.5 equiv of CB[8] solid sample was added, the pyridine characteristic peak disappeared (Figure S5(b), Peak: b, c, d). Then MMV^+ was added until the ratio of MMV^+ : CB[8] reach 3:1 (0.33 equivalent CB[8]), and as Figure S5(c) shown, the pyridine characteristic peak reappeared, which indicated that analogue of methyl viologen as 1^+ form bind to CB in a 2:1 stoichiometry.

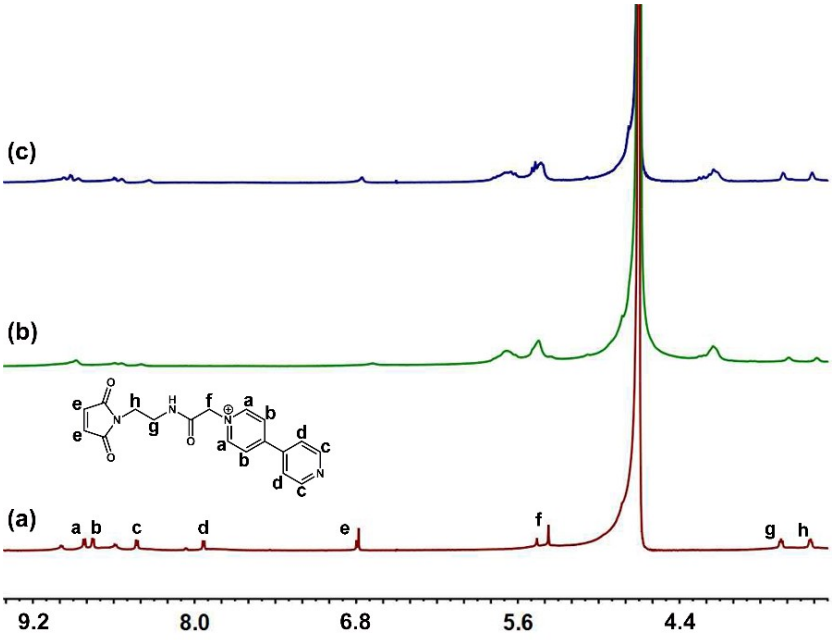


Figure S5. Partial ^{1}H-NMR titration spectra (500 MHz, D_2O, 298 K) of MMV^+ (8 mM) with the ratio of MMV^+ and CB[8] of (a) 0, (b) 2:1, and (c) 3:1.

8. AFM images of host-guest assembly with 5:1 and 1:10 ratio

As illustrated in Figure S6a, when the ratio of SP1S98C-MMV^+/CB[8] was 5:1 with excessively low content of CB[8], assembly hardly happened except for the original dispersed proteins, which was consistent well with the results of Native-PAGE analysis. Moreover, when the ratio changed to 1:10, the result came similarly (Figure S6b), which could be explained by the consequences of the end capping for introducing extremely excess CB[8] that harassed the growth of assemblies and there were only small-scaled nanosheets emerged.

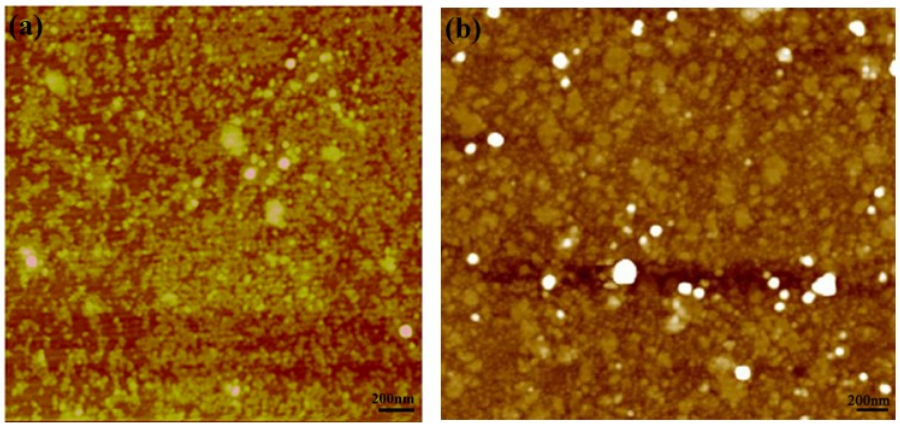


Figure S6. AFM characterizations of host-guest assembly with the ratio of SP1S98C-MMV^+:CB[8] respectively are (a) 5:1, (b) 1:10.

9. AFM images of SP1S98C-MMV$^+$ nanosheets

Fixed the ratio of SP1S98C-MMV$^+$/CB[8] to 1:5, we made the assemblies grew at 4 ℃ for two days and subsequently employed AFM to monitor the results. As expected, the images characterized by AFM illustrated much lager-scaled layer-liked structures in the comparison with the sheet-structures we previously gained and whose diameters really exceeded one micrometer in size (Figure S7a and S7b). Moreover, the 3D images of these nanosheets provided further evidence for the definite large-sized structures (Figure S7c). Similar as before, the obtained sheet-structures showed uniform height (4.3 nm) with the theoretical value of SP1 (Figure S7d). We thus successfully obtained a kind of host-guest interaction-induced, lateral surfaced-oriented, slowly-grown and large-scaled 2D protein nanosheets upon our assumption.

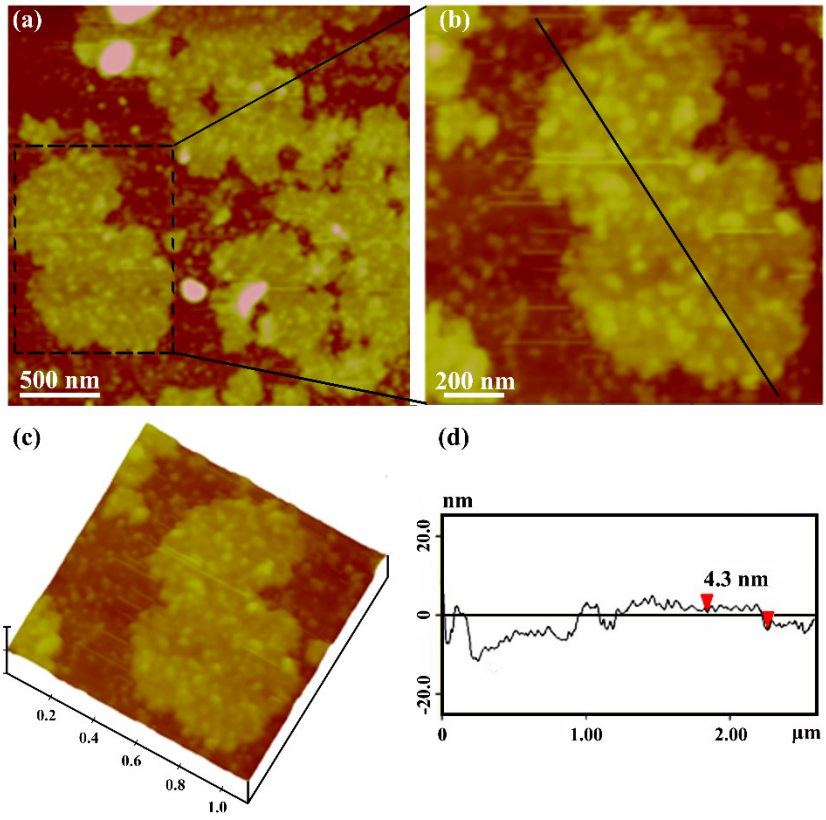


Figure S7. (a) AFM images of SP1S98C-MMV$^+$ nanosheets induced by CB[8] at 4 ℃; (b) An enlarged image of (a); (c) 3D image corresponding to (b); (d) the height profile along the black line in (b).

10. Size exclusion HPLC analysis of the self-assembly process of protein.

As illustrated in Figure S8, as the ratio of SP1S98C-MMV^{+}/CB[8] fixed as 1:5 and incubated for different time, size exclusion high performance liquid chromatography (HPLC) analysis was taken to monitor the self-assembly process of protein. With the assembly process continued, large scaled nanocomposites are gradually formed, which was well consistent with DLS findings.

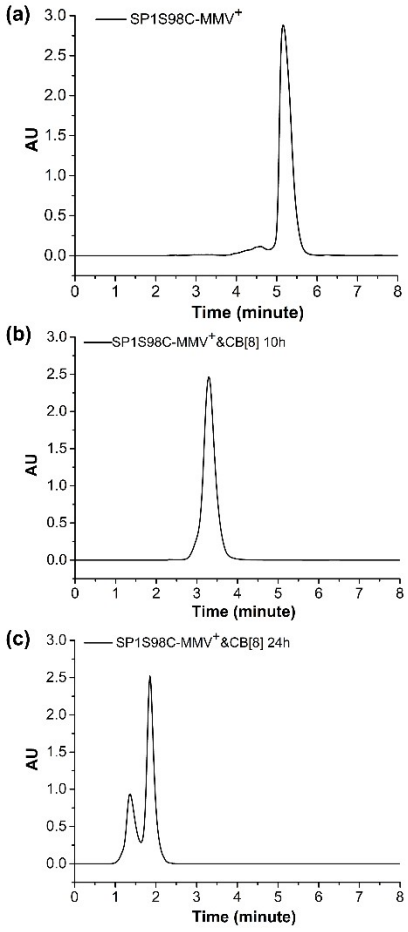


Figure S8. Size exclusion HPLC analysis of the self-assembly process of protein. (a) Free SP1S98C-MMV^{+}; (b) SP1S98C-MMV^{+}&CB[8] assembled for 10 h; (c) SP1S98C-MMV^{+}&CB[8] assembled for 24 h.

11. Schematic representation of the A-D-R process of the 2D protein nanosheets

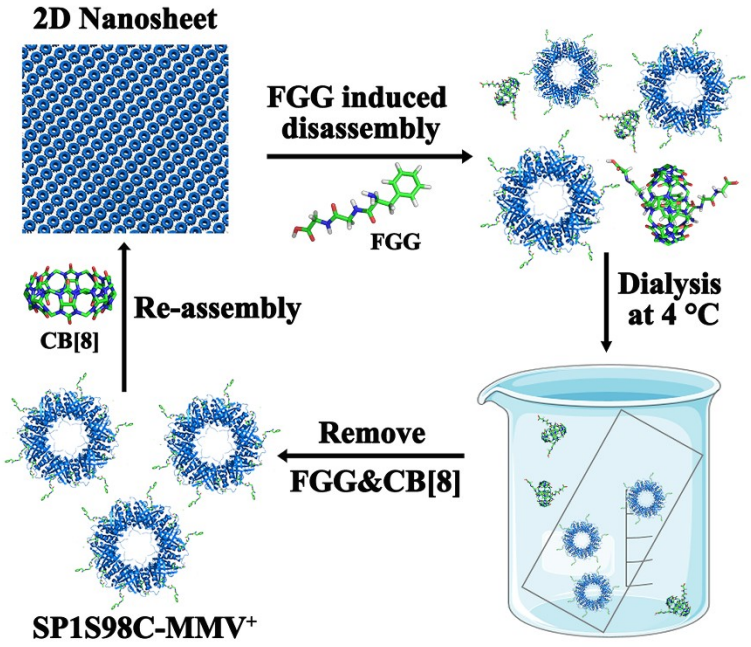


Figure S9. Schematic representation of the A-D-R process of the 2D protein nanosheets.

12. References

1. Zhao, L.; Zou, H.; Zhang, H.; Sun, H.; Wang, T.; Pan, T.; Li, X.; Bai, Y.; Qiao, S.; Luo, Q.; Xu, J.; Hou, C.; Liu, J. Enzyme-Triggered Defined Protein Nanoarrays: Efficient Light-Harvesting Systems to Mimic Chloroplasts. *ACS Nano* **2017**, *11*, 938-945.

2. Li, X.; Qiao, S.; Zhao, L.; Liu, S.; Li, F.; Yang, F.; Luo, Q.; Hou, C.; Xu, J.; Liu, J. Template-Free Construction of Highly Ordered Monolayered Fluorescent Protein Nanosheets: A Bio-Inspired Artificial Light-Harvesting System. *ACS Nano* **2019**, *13*, 1861-1869.

3. Ma, N.; Li, F.; Li, S.; Chu, S.; Han, L.; Liu, S.; Yan, T.; Tian, R.; Luo, Q.; Liu, J. A remote optically controlled hydrolase model based on supramolecular assembly and disassembly of its enzyme-like active site. *Nanoscale* **2019**, *11*, 3521–3526.