Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2021

Supporting Information

for

Reversible structure transformation between protein nanocages and nanorods controlled by small molecules

Xiaorong Zhang, Tuo Zhang, Yingjie Wang, Yu Liu, Jiachen Zang, Guanghua Zhao*

College of Food Science & Nutritional Engineering, China Agricultural University, Beijing Key Laboratory of Functional Food from Plant Resources, Beijing 100083, China

*Corresponding author.
E-mail address: gzhao@cau.edu.cn.

Experimental Section

1. Protein expression and purification

The gene encoding full-length amino acid sequence of *Thermotoga maritima* ferritin (TmFtn) were synthesized by Synbio Technologies, which have been inserted into the plasmid pET-3a. TmFtn was purified as previously described.[1] Briefly, the plasmid corresponding to TmFtn was transformed into BL21 (DE3) *E. coli* cells and then cultured at 37 °C in 1-liter of LB media containing 100 μg/mL ampicillin. After the cell density reached an absorbance of 0.6 at 600 nm, protein expression was induced with 200 μM IPTG for 10 h at 37 °C. The cells were harvested by centrifugation and the precipitate was resuspended in 50 mM Tris-HCl (pH 8.0), followed by disruption by sonication. The supernatant was collected from lysed cell samples and subjected to heating at 90 ℃ for 10 min. The supernatant was collected by centrifugation and fractionated by 60% saturation of ammonium sulfate. The pellet was resuspended in 50 mM Tris-HCl, pH 8.0, and dialyzed against the same buffer. Finally, the protein solution was applied to an ion-exchange column (DEAE sepharose Fast Flow, GE Healthcare), followed by gradient elution with 0-1.0 M NaCl. The purified protein was then dialyzed against 50 mM Tris–HCl (pH 8.0) at 4 ℃ to exclude NaCl from the solution, and protein concentrations were determined according to the Lowry method with bovine serum albumin (BSA) as standard.

2. Construction of nanocages and nanorods

Typically, the purified TmFtn molecules remain as dimers in solution in the absence of salts. For the formation of protein nanocages, 50 mM Tris-HCl (pH 8.0) was used as buffer and 50 mM $MgCl_2$ or 50 μM PLL_9 was required to induce TmFtn nanocage self-assembly. Nanorods were prepared by adding 50 μM PLL_{15} to TmFtn solution (12.0 μM) at room temperature. After being stirred for several minutes, the resulting mixture was incubated at 4 °C overnight. All polypeptides used in this study are synthesized in Scilight Biotechnology, and are linear polymers where each amino acid residue participates in two peptide bonds and is linked to its neighbours in a head-to-tail fashion rather than forming branched chains. The purity of polypeptide is 95% or higher.

3. Transformation of protein nanocages and nanorods

To convert protein nanocages into nanorods, 100 mM sodium citrate was added into a solution of Mg^{2+}-induced protein nanocages. Subsequently, the above protein solution was dialyzed against 50 mM Tris-HCl (pH 8.0) for three times to remove Mg^{2+} and sodium citrate. Finally, 50 μM PLL_{15} was added to the resulting solution, followed by stirring for several minutes, and incubation at 4 °C overnight. As for the conversion from nanorods to protein nanocages, just added 50 mM $MgCl_2$ into the solution containing PLL_{15} induced nanorods.

4. High-Resolution Gel Filtration Chromatography Analyses.

High-resolution gel filtration chromatography analyses were performed using an ÄKTA pure system coupled to a Superdex 200 increase 10/300 column (GE Healthcare) in buffer (50 mM Tris, 100 mM NaCl, pH = 8.0) with a flow rate of 0.5 mL/min.

5. Dynamic light scattering (DLS) analysis

DLS experiments were performed at 25 °C by using a dynamic light scattering instrument (Viscotek, Europe, Viscotek model 802). The OmniSIZE 2.0 software was used to calculate the size distribution of samples, and Origin was used to present the collected data. The concentration of protein was 1.0 μM.

6. Polyacrylamide gel electrophoresis (PAGE)

The purity and molecular weight of protein samples was estimated by polyacrylamide gel electrophoresis. Gel electrophoresis under denaturing conditions was carried out using a 15% polyacrylamide-SDS gel as reported by Laemmli.[2] Native PAGE was performed by using a 4–20% polyacrylamide gradient gel and run at 180V for 1.5 h at 4 ℃. Gels were stained with Coomassie brilliant blue R250.

7. Transmission electron microscopy (TEM) imaging

10 μL diluted protein samples were placed on carbon-coated copper grids and excess solution was removed with filter paper after a 5-min incubation. Then protein samples were stained using 2% uranyl acetate (Beijing Zhongxingbairui Technology Co., Ltd.) for 3 min. Transmission electron micrographs were obtained at 80 kV through a Hitachi H-7650 transmission electron

microscope.

8. Crystallization, data collection and structure determination of T4FY assemblies

The purified protein was concentrated to 6 mg/mL in a buffer consisting of 20 mM Tris-HCl at pH 8.0, and crystals were obtained using the hanging drop vapor diffusion method. TmFtn-cage crystals were crystallized in a buffer containing 2 mM PLL_9, 100 mM Tris-HCl (pH 8.0) at 20 °C. TmFtn-nanorod crystals were crystallized in a buffer containing 1.2 mM PLL_{15}, 100 mM Tris-HCl (pH 8.5) at 4 °C. Cubic-shaped and Rod-shaped crystals appeared and grew to a full size within one weeks.

Diffraction data of the cubic and rod-shaped crystals were collected to resolutions of 2.2 Å and 2.3 Å, respectively. Data were indexed, integrated and scaled with the HKL-3000 (HKL Research).[3] Data statistics are shown in Table S1. The structure of TmFtn assemblies was determined by molecular replacement using the Molrep program in CCP4 using the structure of *Thermotoga maritima* ferritin (TmFtn) (PDB code 1VLG) as a search model. Structure refinement was conducted using the Refmac5 program and PHENIX software.[4] The structure was rebuilt using COOT,[5] which made the model manually adjusted. Figures of protein structures were generated by using the PyMOL[6] program and UCSF Chimera package.[7]

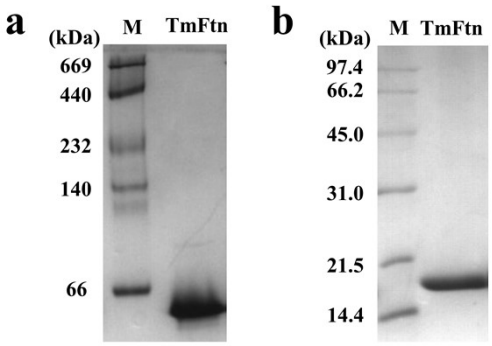

Fig. S1 Purification and characterization of TmFtn. Native PAGE (a) and SDS-PAGE (b) analyses of purified native TmFtn. Lane M, protein markers and their corresponding molecular masses.

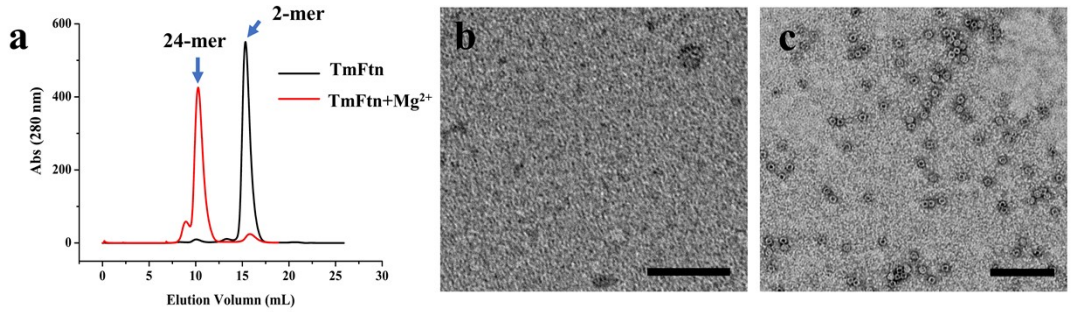

Fig. S2 characterization of the self-assembly property of TmFtn induced by Mg^{2+}. (a) High-resolution gel filtration chromatography analyses revealed that dimeric TmFtn assemble into hollow protein nanocage in the presence of Mg^{2+}. (b, c) TEM images of TmFtn in the absence (b) and presence of 50 mM Mg^{2+} (c). Conditions: [TmFtn] = 12.0 μM in 50 mM Tris-HCl, pH 8.0. (b, c) Scale bars represent 100 nm.

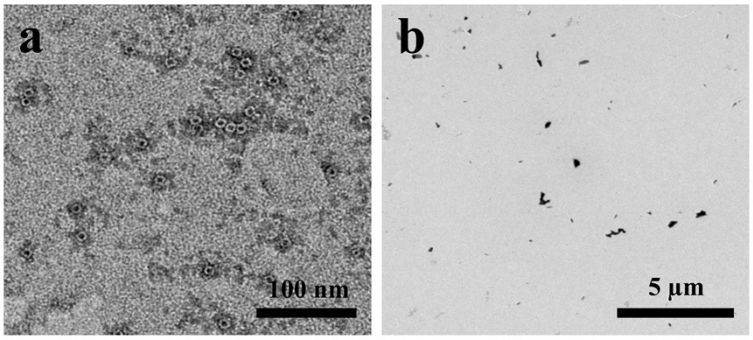

Fig. S3 (a) TEM image of TmFtn upon treated with 50 μM PLL_9. (b) TEM image of TmFtn upon treated with 0.5 mM PLL_9.

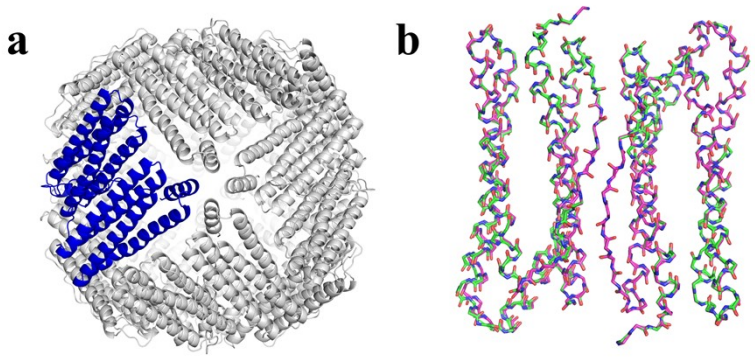

Fig. S4 (a) The crystal structure of TmFtn upon treated with PLL_9, which exhibits a shell-like structure, similar to TmFtn protein nanocage induced by salts. (b) Structure alignment of the main chain of dimeric molecules between PLL_9 treated TmFtn and wild type TmFtn. Magentas: Wild type TmFtn; green: PLL_9 treated TmFtn.

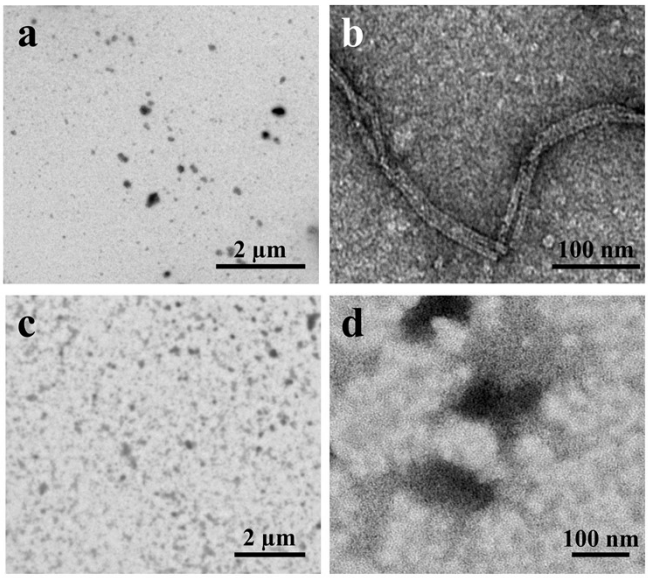

Fig. S5 TEM images of TmFtn upon treatment with 50 μM PLL_{30} (a, b) or PLL_{50} (c, d).

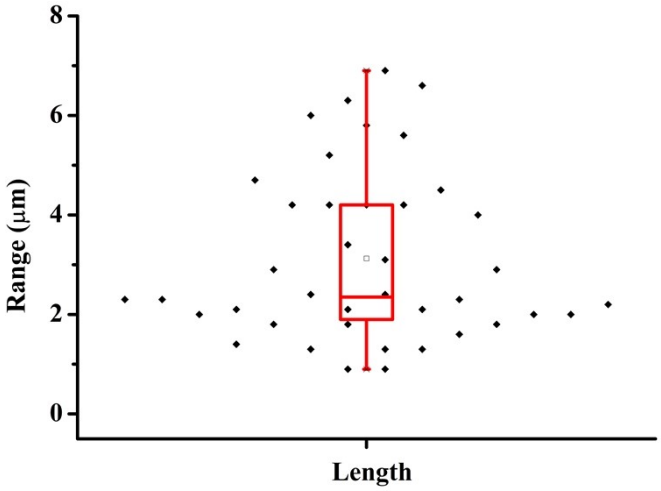

Fig. S6 Length distribution of the nanorods. Statistic for random 40 nanorods.

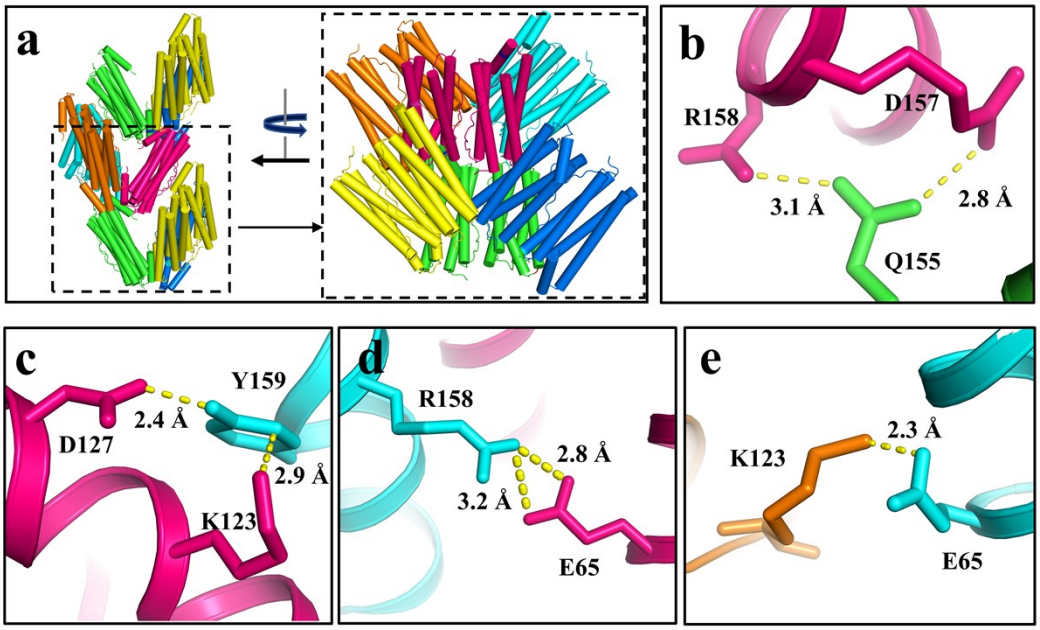

Fig. S7 Structural analyses of the noncovalent interactions between dimers in the crystal structure of nanorods. (a) The crystal structure revealed that a dimer is surrounded by five other dimers, then producing a basic repeating unit. (b-e) The noncovalent interactions involved in such unit consist mainly of salt bridges and hydrogen bonds. Additionally, there are cation-π interactions between K123 and Y159 and anion-π interactions between D127 and Y159. The identical dimer is shown in the same color in (a-e).

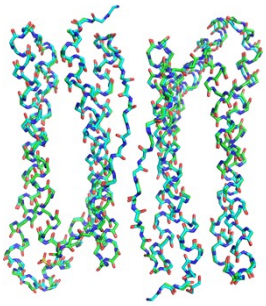

Fig. S8 Structure alignment of the main chain of dimeric molecules between PLL_{15} treated TmFtn and wild type TmFtn. Cyans: Wild type TmFtn; green: PLL_{15} treated TmFtn.

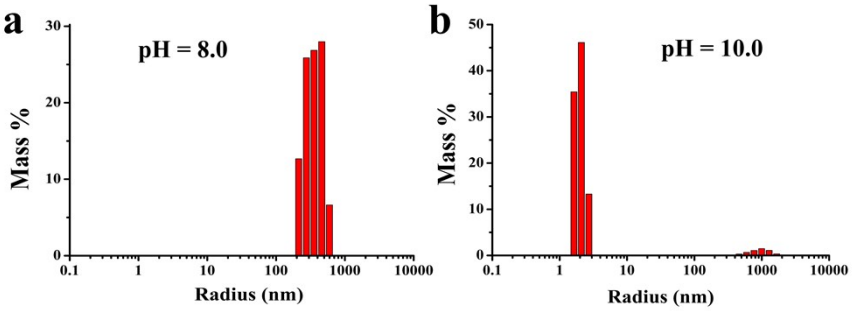

Fig. S9 (a) Dynamic light scattering (DLS) characterization of the nanorods formed upon treatment of 12.0 μM of TmFtn with 50.0 μM of PLL_{15} at pH 8.0. (b) The formed nanorods disassociated when solution pH was adjusted from 8.0 to 10.0.

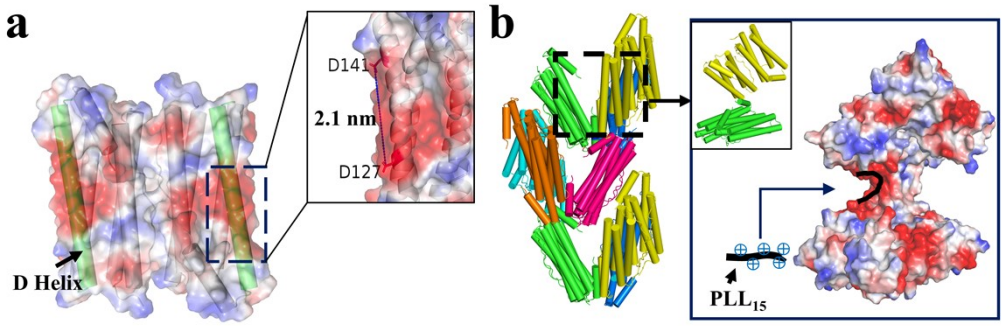

Fig. S10 (a) Electrostatic characterization of dimer inner surface. The negative charges on D helix occupied a length of 2.1 nm. Red: negative charge. Blue: positive charge. (b) The possible binding site of PLL_{15} to dimers in the rod-shaped crystal.

Table S1 Crystallographic properties and data collection and model refinement statistics.

Parameters	Rod-shaped crystal	TmFtn-cage induced by PLL_9
Wavelength (Å)	0.97919	0.97919
Space group	P212121	R32:H
Resolution range (Å)	33.58-2.3 (2.382-2.3)	38.56-2.2 (2.279-2.2)
Unit cell	84.808, 94.495, 108.121 90, 90, 90	176.193, 176.193, 357.508 90, 90, 120
Unique reflections	38772 (3428)	107941 (10690)
Completeness (%)	98.69	99.91
Mean I/sigma (I)	1.4	3.32
Wilson B-factor	43.57	32.81
CC1/2	0.996	0.957
Reflections used in refinement	38758 (3428)	107909 (10690)
Reflections used for R-free	1998 (176)	2000 (198)
R-work	0.2016 (0.2767)	0.1803 (0.2080)
R-free	0.2579 (0.3680)	0.2274 (0.2808)
Number of non-hydrogen atoms	5557	11749
macromolecules	5428	10928
ligands	8	16
Protein residues	651	1311
RMS (bonds)	0.014	0.013
RMS (angles)	1.29	1.12
Ramachandran favored (%)	97.20	98.38
Ramachandran allowed (%)	2.64	1.62
Ramachandran outliers (%)	0.16	0.00
Rotamer outliers (%)	0.18	0.00
Clashscore	9.01	3.96
Average B-factor	52.15	39.49
macromolecules	52.17	39.26
ligands	61.27	46.01
solvent	50.44	42.58

References

1. S. Chakraborti, A. Korpi, M. Kumar, P. Stępien, M. A. Kostiainen and J. G. Heddle, *Nano Lett*., 2019, **19**, 3918–3924.

2. U. K. Laemmli, *Nature*, 1970, **227**, 680-685.

3. Z. Otwinowski and W. Minor, *Methods Enzymol*., 1997, **276**, 307-326.

4. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr D.*, 2010, **66**, 213-221.

5. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr D.,* 2010, **66**, 486-501.

6. W. L. DeLano, *Proteins: Struct., Funct., Bioinf.*, 2002, **30**, 442-454.

7. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem*., 2004, **25**, 1605-1612.