Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2017

Supplementary Information

Dynamic Protein Self-Assembly Driven by Host-Guest Chemistry and the Folding-Unfolding Feature of a Mutually Exclusive Protein

Ruidi Wang,[a] Shanpeng Qiao,[a] Linlu Zhao,[a] Chunxi Hou,[a] Xiumei Li,[a] Yao Liu,[a] Quan Luo,[a] Jiayun Xu,[a] Hongbin Li*[b] and Junqiu Liu*[a]

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

[b] Department of Chemistry, University of British Columbia Vancouver, British Columbia, Canada V6T 1Z1.

E-mail: junqiuliu@jlu.edu.cn
Hongbin@chem.ubc.ca

Contents

1. Engineering fusion protein FGIG for constructing redox stimuli-responsive protein self-assemblies

The strategy of engineering expression vectors of fusion protein FGIG was described as Fig. S1. First, we obtained a mutant glutathione S-transferase (GST) that was removed all the Cys *via* PCR site-directed mutagenesis. The open reading frame of varietal GST was amplified with the template of pGEX-5x-2 using the primers GSTP1 (5'-CCATGGTCCCCTATACTAGGTATGTGGAA-3') and GSTP2 (5'-AAGCTTGTCAGTCACGAT
GCGGC-3'). The underlined sequences were the recognition sites of restriction endonuclease Nco I and Hind III, respectively. Digested the resulting GST open reading frame with Nco I/Hind III, and ligated it into NcoI/HindIII treated pET-22b vector, then we got the vector of pET-GST. For the second step, fused a tripeptide Phe-gly-gly (FGG) encoding sequence to the N-terminus of GL5CC-I27w34f (G-I) mutually exclusive protein, resulting to the FGG-G-I. G-I open reading frame was amplified with the template of pUC-19 with the primers FGIP1 (5'-CATATG**TTTGGCGGC**GACACCTACAAACTGATCCTGAACG-3') and FGIP2 (5'-CCATGGG
CCGCCTTCGGTAACCGTGAAGGTTTTG-3'). The underlined sequences were the recognition sites of restriction endonuclease Nde I and Nco I, and the sequence in bold was tripeptide FGG-encoding sequence. The resulting FGG-G-I DNA fragment was digested with Nde I/Nco I and ligated into Nde I/Nco I treated pET-GST, and we got expression vector pET-FGIG finally. The resulting plasmid pET-FGIG was confirmed by agarose gel electrophoresis (AGE) analysis (Fig. S2) and DNA sequencing.

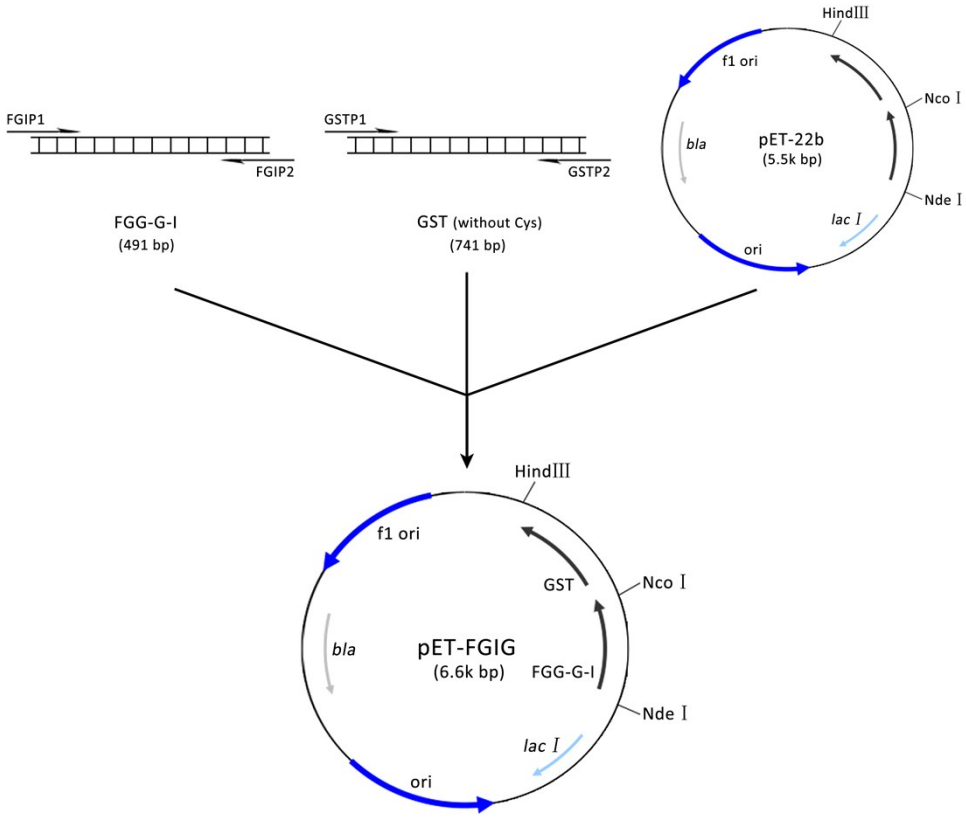

Fig. S1 Engineering expression vector pET-FGIG of fusion protein FGIG.

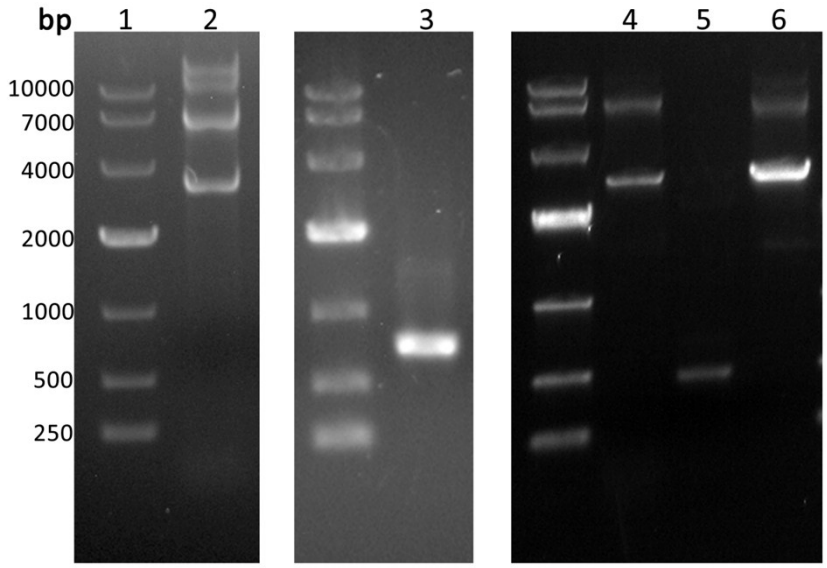

Fig. S2 Agarose gel electrophresis analysis. Lane 1: DL10,000 DNA marker; lane 2: pET-22b plasmids; lane 3: GST (without Cys) DNA amplified fragments; lane 4: pET-GST plasmids; lane 5: FGG-G-I DNA amplified fragments; lane 6: pET-FGIG plasmids. The agarose percentage of the AGE gels was 1%.

2. Over expression and purification of fusion protein FGIG

The expression vector pET-FGIG was transformed into *Escherichia Coli* strain BL21 (ED3), which was incubated in 5mL LB liquid medium including 100 μg/mL Amp overnight. Then cells were transformed into 1 L LB liquid medium with 100 μg/mL Amp and incubated until OD600 is about 0.6-0.8, where upon adding IPTG to induce the expression at a final concentration of 1 mM. After that, incubated the cells for about 4 hours and collected cells by centrifugation at 8000rpm, 4 °C, 15 min. All the incubation was under 37 °C along with rotary shaking.

Over expressing cells were sonicated and separated sediment by centrifugation at 15000 rpm, 30 min, under 4 °C, then the soluble part was purified with Glutathione Sepharose. Excessive glutathione was removed by dialysis over 24 hours and dialyzed solution was lyophilized to dry powder finally. Purified protein was confirmed by SDS-PAGE gel electrophoresis (Fig. S3).

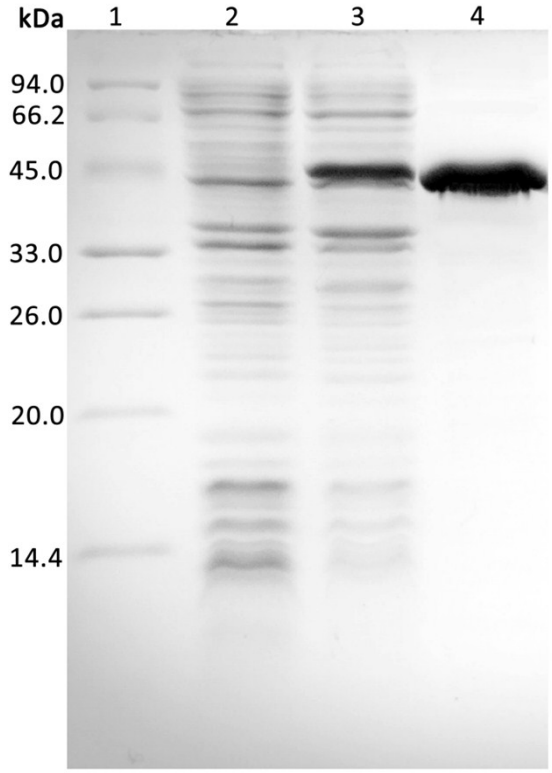

Fig. S3 SDS-PAGE analysis of fusion protein FGIG. Lane 1: protein marker; lane 2: *E.coil* BL21 (DE3) containing pET-FGIG expression vector without IPTG induction; lane 3: *E.coil* BL21 (DE3) containing pET-FGIG expression vector after 4-hour inducing with 1 mM IPTG; lane 4: purified FGIG over Glutathione Sepharose. The acrylamide percentage of the SDS-PAGE gel was 15%.

3. MALDI-TOF mass spectrometry of fusion protein FGIG

MALDI-TOF mass spectrometry was performed using autoflex speed^{TM} TOF/TOF mass spectrometer (Bruker, daltonics Inc., USA) as described[1]. Prior to determination, all the samples were desalted by dialysis against MilliQ water with dialyzing tubes, Slide-A-Lyzer Dialysis Cassettes (Pierce) or G25 filtration. The resulting FGIG protein was lyophilized and resolved at a concentration of 5 µM. For preparation, sinapic acid was chosen as matrix which was saturated in 70% acetonitrile and supplemented with 1‰ trifluoroacetic acid. 1 µL samples and 1 µL matrix were sequentially dropped onto the ground steel and dried in air. The molecular weights of FGIG were determined as shown in Fig. S4.

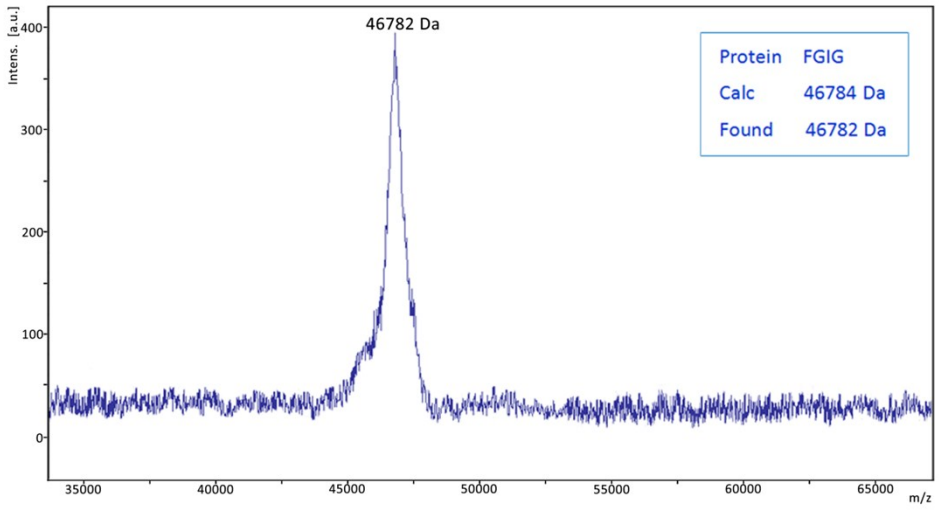

Fig. S4 MALDI-TOF mass spectrometry analysis of FGIG

4. Isothermal titration calorimetry (ITC) for CB[8] induced FGIG self-assembling

Calorimetric experiment were carried out on an MCS ITC (Microcal Inc., Northampton, MA).[2] Prior to the titration, the calorimeter was calibrated by known heat pulse as recommended by the manufacture. Titrations of 50 μM CB[8] to 2.5 μM FGIG were performed in 20 mM sodium phosphate buffer, pH 7.4. The reference cell was filled with MilliQ water; the sample cell was filled with 200 μL buffer or FGIG solution. Samples were degassed for 20 minutes before use. The resulting mixture kept stirring at 400 rpm during titration.

Experimental data were plotted as the amount of heat evolved per second following each injection of CB[8] into the FGIG solution along with time and blank titration of CB[8] into buffer were performed to correct for heat generated by dilution and mixing. The ITC plot in Fig. S5 was consistent with the complex stoichiometry of 2:1 (FGG:CB[8]) with exothermic interaction enthalpies of 3.74×10^4 cal mol^{-1} and the Kd was 4.8×10^{10} M^{-2} at 25 °C.

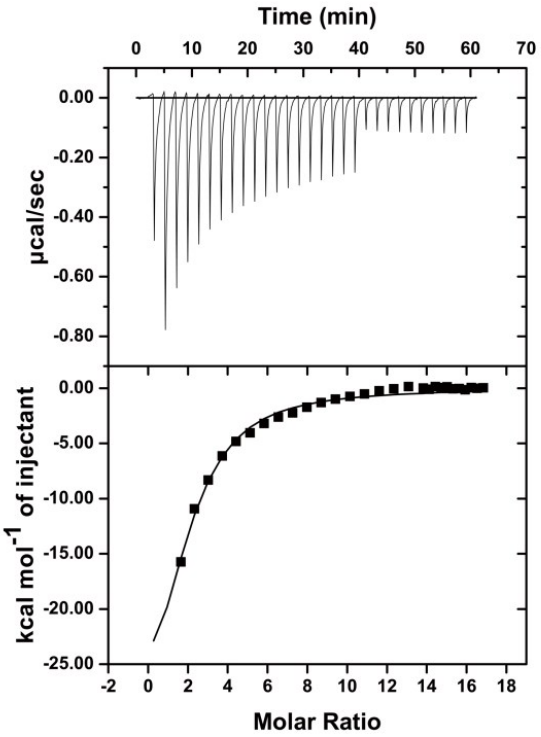

Fig. S5 ITC data of the titration of 50 μM CB[8] solution into 2.5 μM FGIG solution. The raw data for power versus time and the integrated enthalpy values versus the molar ratio of FGIG:CB[8] was at top and bottom, respectively.

5. Engineered mutually exclusive protein G-I and schematic design strategy of redox-responsive FGIG protein self-assembly

Mutually exclusive proteins are a special class of engineered domain-insertion proteins, whose structure is extremely incompatible. In the mutually exclusive protein, a guest domain is inserted into the host protein. However, the distance between the N-C termini of the guest protein is significantly longer than that between the two ends of the surface loop of the host protein where the guest domain is inserted. In a consequence of such topological constrains, only one of the two domains remains originally folded at equilibrium conformation.

In previous work, we engineered a novel redox sensitive mutually exclusive protein G-I. The G-I protein was able to interconvert reversibly between two conformations and this conformation conversion was proved redox dependent *via* an introduced disulfide bond. The G-I composed of the host protein GL5-C41and43(GL5CC) and inserted guest protein I27w34f, which were respectively the disulfide mutant of GB1 with inserted 5-residue-loop and the Trp34Phe mutant of I27. In oxidative condition, G-I formed a disulfide bond between C41 and C43 readily, which drove the equilibrium toward the G(F)-I(U) conformation. In reductive condition, the disulfide bond would be break which resulted to a G(U)-I(F) conformation (Fig. S6A). Therefore G-I could be a sensitive switch and played a crucial role in constructing stimuli-responsive protein self-assembly. When fusing G-I with the mutant GST (without Cys), a natural homodimer at C-terminal of G-I, we acquired a *C*2-symmetric dimer that was redox-controlled. With the help of fused FGG tripeptide at the N-terminal of G-I, newly engineered fusion protein FGG-G-I-GST (FGIG) began to assemble as Fig. S6B after adding cucurbit[8]uril (CB[8]).

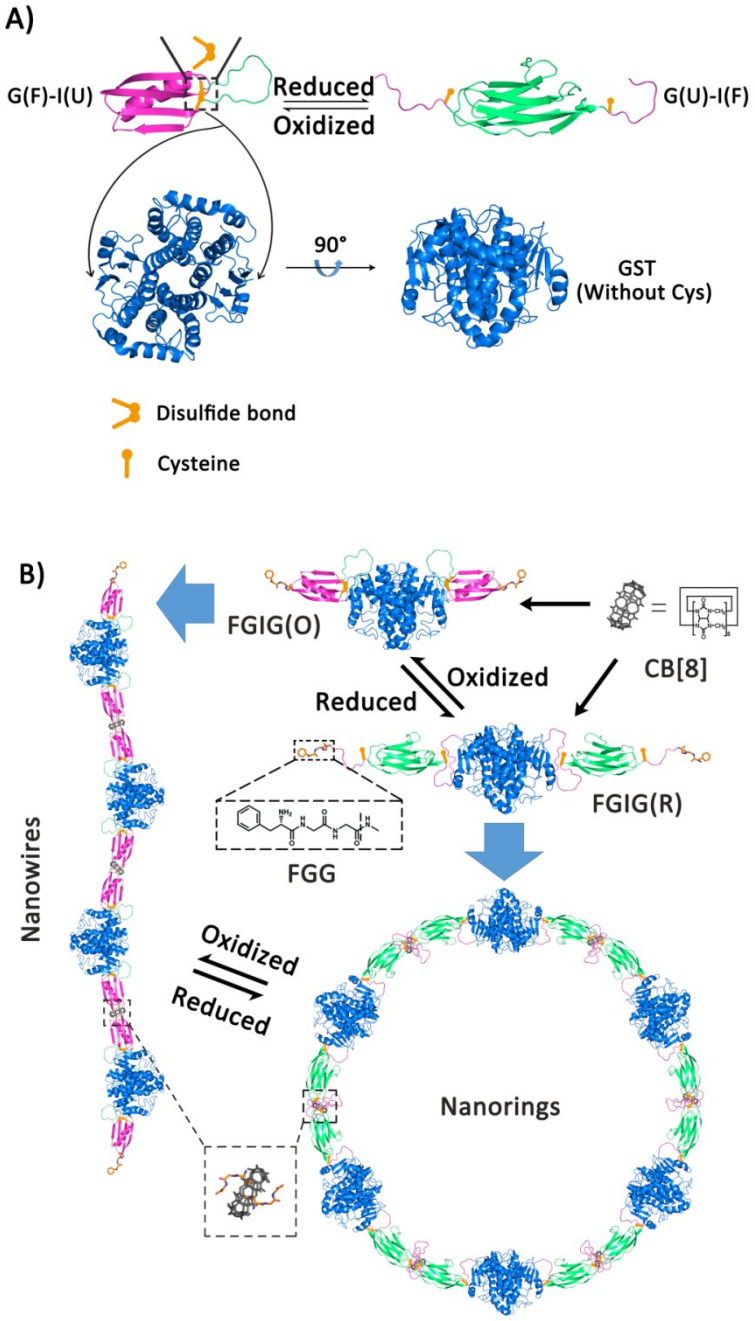

Fig. S6 Scheme presentation of the design strategy of the redox stimuli-responsive protein self-assembly. A) Structure transforming mechanism of redox-responsive domain. B) Protein self-assembly formation induced by the interaction between CB[8] and FGG.

6. Preparation of FGIG protein self-assembly under redox control

Preparation process of FGIG protein self-assembly was illustrated by taking a protein concentration of 5 μM as an example. Lyophilized FGIG was dissolved into MilliQ water at a final concentration of 5 μM with 2 mM PBS (pH 7.4) as well as 0.5 mM β-mercaptoethanol (βMe) which slowly decomposed in aqueous solution and resulted in gradually more oxidizing conditions, standing over 24 hours in order to acquire oxidized protein FGIG(O). Then 2.5 μM CB[8] (2:1 FGIG:CB[8]) was added into this solution with slight stir overnight at 4 °C to assemble into nanowires. Adding 10 μM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) or 1.5 mM βMe to FGIG(O) assembly system to realize morphology transform by making FGIG(O) into reduction state, at last 3 hours nanorings can be observed. Re-oxidizing FGIG(R) took place when reductant decomposed gradually contributing to a oxidizing condition again after exposed in air for over one day. The morphology of obtained architectures was shown in AFM and TEM images.

7. Dynamic light scattering (DLS) analysis of fusion protein FGIG and CB[8] directed self-assemblies in different states

Dynamic Light Scattering (DLS) measurements and analyses were made on a Malvern Instrument Zetasizer Nano ZS instrument (Malvern, U.K.). All the samples were dissolved in MilliQ water containing 2 mM PBS (pH 7.4) at a concentration of 5 μM with 0 μM and 2.5 μM CB[8] (2:1 FGIG:CB[8]). The results were exhibited in Fig. S7, which showed hydrodynamic diameters of FGIG system had a significant increase when CB[8] was added, indicate CB[8] will induced the formation of self-assemblies. All samples were filtered before analysis and all measurements were repeated for three times at 4 °C.

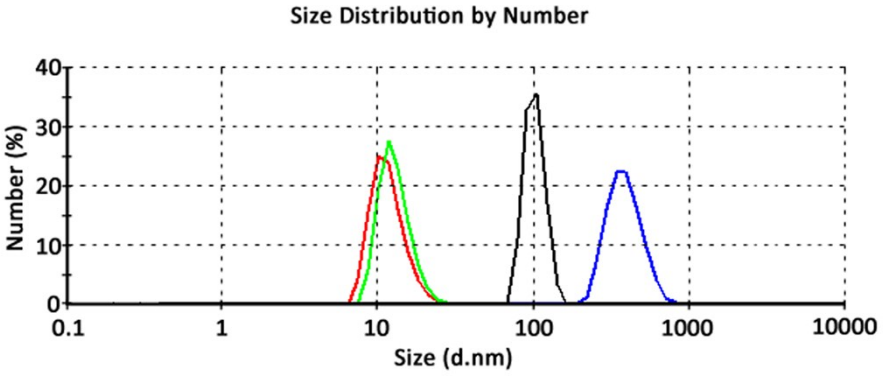

Fig. S7 The dynamic light scattering (DLS) analysis of the hydrodynamic diameters of oxidation-state FGIG before adding CB[8] (red), reduction-state FGIG before adding CB[8] (green), the assemblies after adding CB[8] to system of oxidative state (blue) and reductive state (blank).

8. Circular Dichroism (CD) Spectral Analysis of fusion protein FGIG and CB[8] directed protein self-assemblies

CD spectra were performed using a MOS-450/AF-CD spectropolarimeter (Bio-Logic, France) equipped with a thermostatted cell holder, using a 0.1 cm quartz cell. All sample volume in CD measurements was 350 µL and protein concentration was 5 µM. Spectra were collected by averaging three repetitive scans with using 1 nm intervals, 1 nm bandwidth and 120 nm/min. Spectra were baseline corrected before converting to mean residue ellipticities.

CD results showed that there was significant difference between FGIG(O) and FGIG(R) owing to secondary structure transform. From Fig. S8, reduction-state band had blue shift comparing with oxidation-state band. While CD spectra of FGIG(O) assemblies induced by CB[8] had no obvious transition against FGIG(O) as shown in Fig. S9, suggesting that the formation of FGIG(O) assemblies didn't disarrange the secondary of FGIG(O).

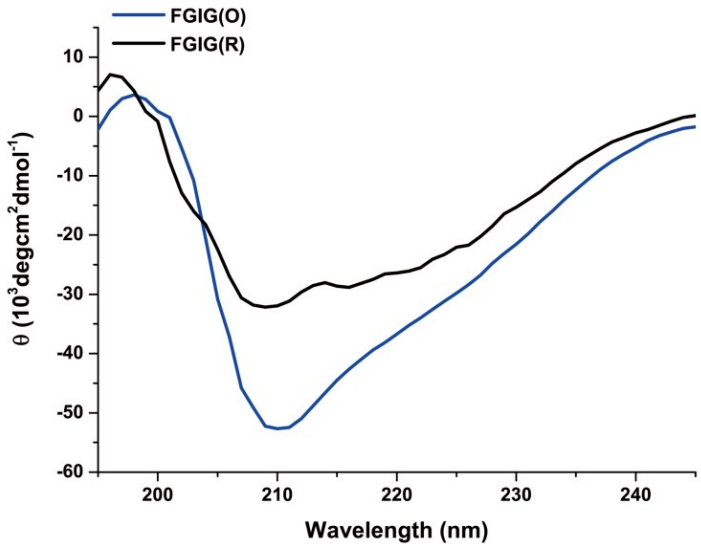

Fig. S8 CD spectra of FGIG(O) and FGIG(R) monitored from 195 nm to 245 nm.

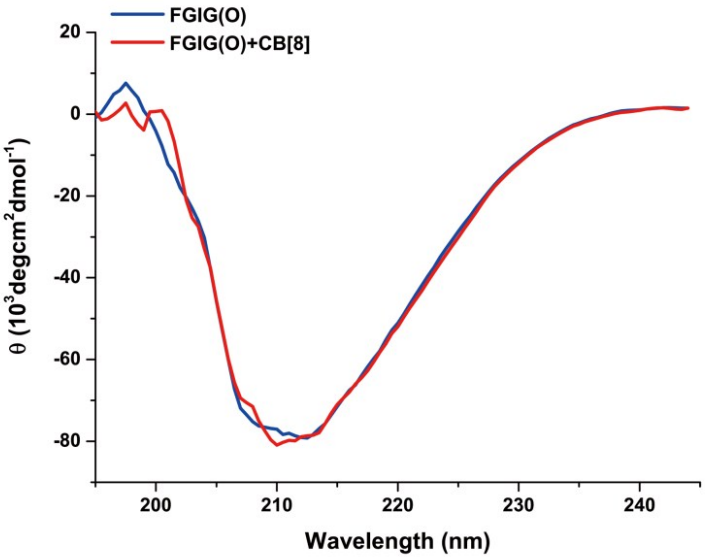

Fig. S9 CD spectral data of FGIG(O) and FGIG(O) assemblies directed by CB[8], monitored from 195 nm to 245 nm.

9. Dynamic light scattering (DLS) monitored the growth of nanowires

Dynamic light scattering (DLS) measurements and analysis were made on a Malvern Instrument Zetasizer Nano ZS instrument (Malvern, U.K.). Hydrodynamic diameters of FGIG(O) assemblies were measured after different time durations. Some typical data is shown in Fig. S10A, the average hydrodynamic diameters of self-assemblies become bigger with time, and up to about 350 nm after 24 hours. The sample was filtered before analyses and all the measurements were repeated for three times at 4 °C. Kinetic curve of the formation of protein assemblies is obtained from the data of DLS studies (Fig. S10B).

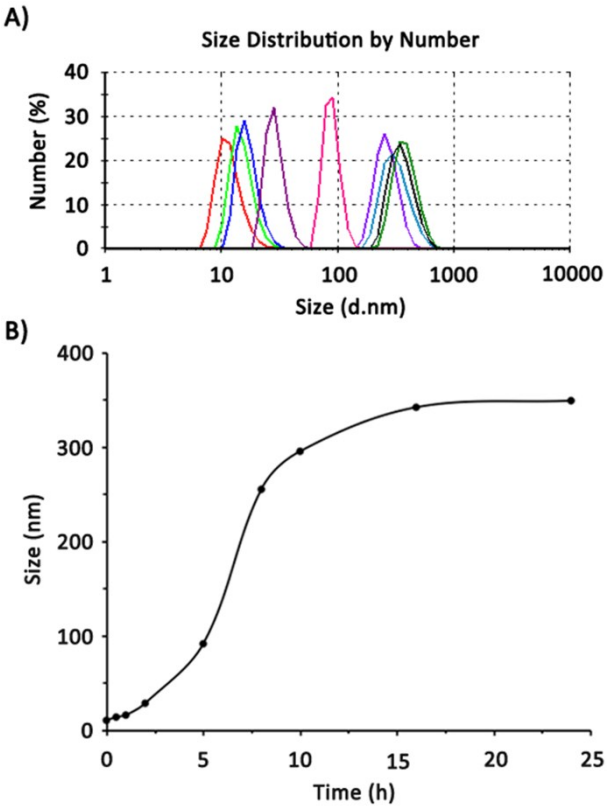

Fig. S10 A) The dynamic light scattering (DLS) analysis of the hydrodynamic diameters of FGIG(O) before the addition of CB[8] (red), and the self-assemblies after the addition of CB[8] at 0.5h (green), 1h (blue), 2h (purple), 5h (pink), 8h (lavender), 10h (marine), 16h (black) and 24h (forest); B) kinetic curve of the formation of FGIG(O) nanowires obtained from DLS studies.

10. Atomic force microscope (AFM) and transmission electron microscopy (TEM) characterization of protein nanowires

The morphology of nanowires constructed by FGIG(O) was first determined by atomic force microscopy (AFM, multimode atomic force microscope with Nanoscope III a controller, Veeco Metrology, Santa Barbara, CA) in tapping mode with silicon probes.[3] The sample height and phase image were recorded and analyzed with the Nanoscope software provided by the AFM manufacturer.

Transmission electron microscopy (TEM) measurements was also used to observe morphology though a JEM-2100F transmission electron microscope with 200 kV accelerating voltage. The sample was deposited on a carbon-caoted copper grid for 10 min, then negative stained by 2% sodium phosphotungstate for 40 s and got dried in air prior to TEM detection.

The morphology was shown in Fig. S11. Results indicated when protein concentration was 0.1 μM dispersed single nanowires showed up (Fig. S11A, C). And when protein concentration increased to 5 μM single nanowires preferred to further assemble into bundles due to the intertwining of unfold loop domains on the side of assemblies (Fig. S11B, D). Through height analysis of several AFM images, we acquired the height distribution histogram of linear nanostructures in protein assembly system with 5 μM protein concentration (Fig. S11E).

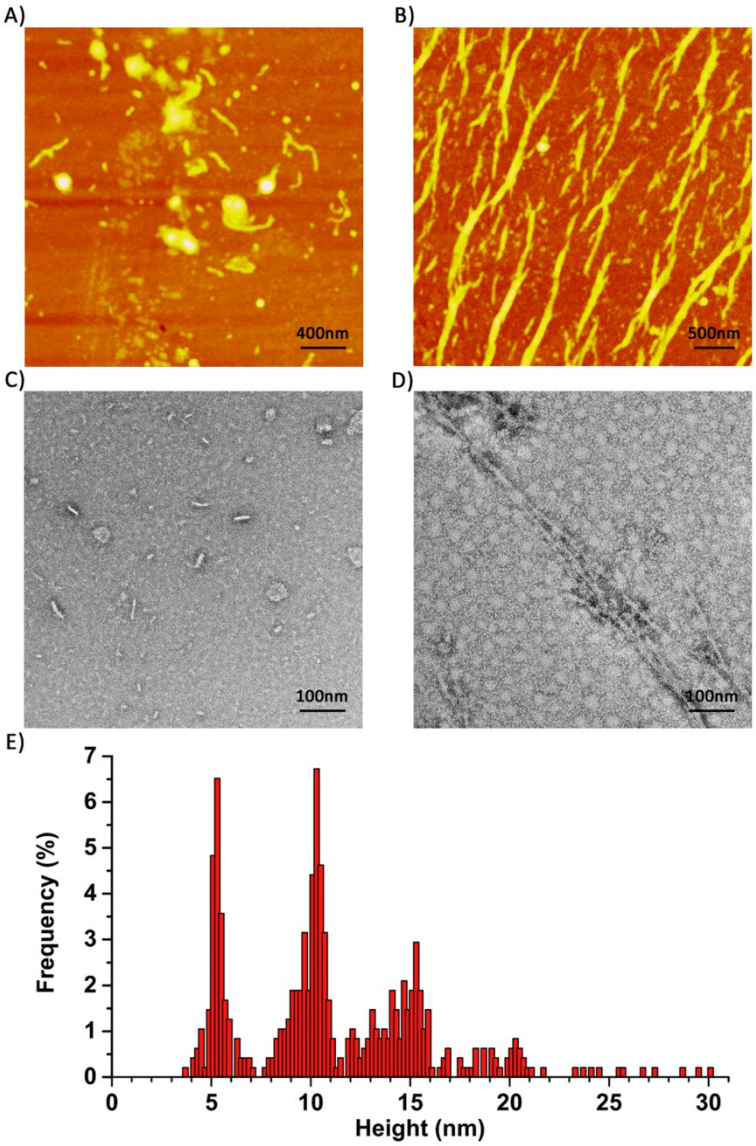

Fig. S11 AFM and TEM images of protein nanowires in oxidized conditions. A) and C) dispersed single nanowires at a low protein concentration. B) and D) bundles formed by further assembled nanowires. E) Height distribution histogram of linear architectures.

11. AFM and TEM characterization of nanorings based on redox-sensitive conversion of FGIG self-assembly

The morphology conversion was realized though altering oxidative and reductive conditions, and after adding TCEP to oxidized assembling system nanorings were observed (Fig. S12).

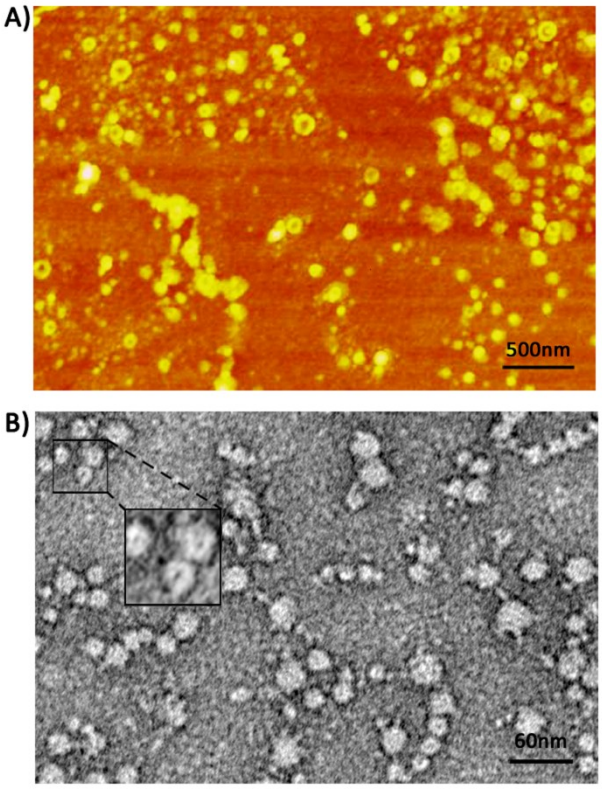

Fig. S12 A) AFM and B) TEM images of protein nanorings in reduced conditions.

12. TEM characterization of oxidative assembly morphology with different protein concentrations.

In order to investigate the morphology transformation of FGIG self-assemblies was not influenced by the protein concentrations but induced through conformation change of building blocks, we tested six kinds of samples from a high concentration to a low concentration *via* TEM. As Fig. S13A and B showed, assemblies in high concentrations were bundle-like structures. When decreasing protein concentrations (Fig. S13C-F), bunched nanowires got dispersed, but no ring-like architectures appeared.

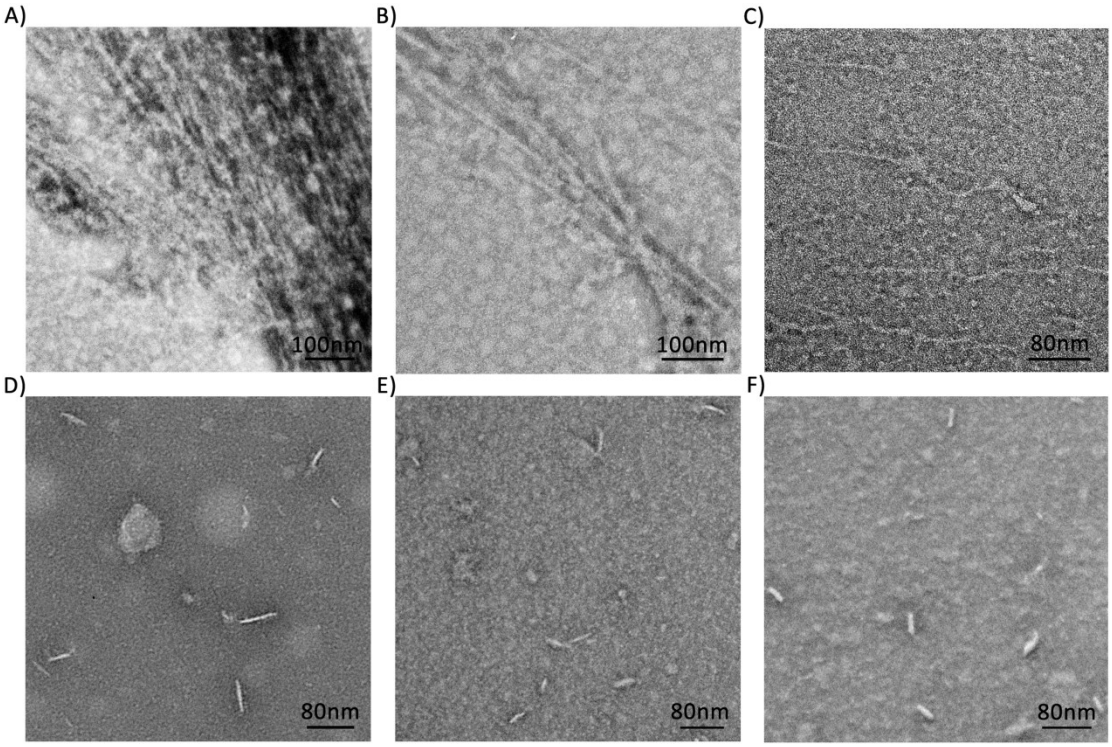

Fig. S13 TEM characterization of FGIG(O) assemblies with different protein concentrations. From (A) to (F), the protein concentrations were respectively 10 μM, 5 μM, 1 μM, 0.5 μM, 0.1 μM and 0.05 μM.

13. The reversibility characterization of FGIG self-assembly morphology transformation *via* AFM

The conversion of protein self-assemblies between oxidation state and reduction state was reversible (Fig. S14).

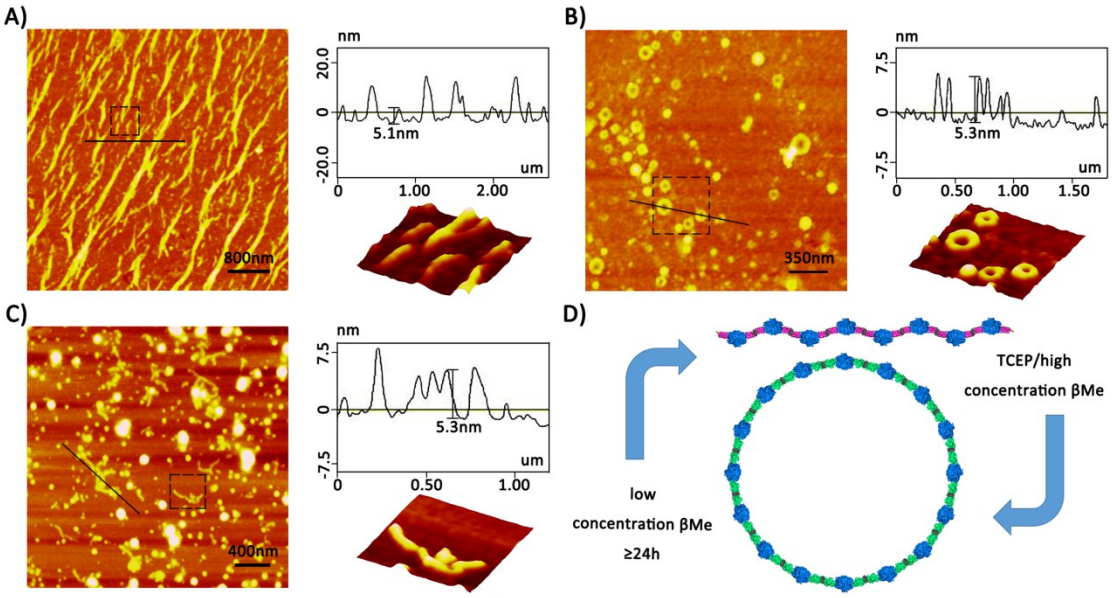

Fig. S14 AFM characterization of alterable FGIG assemblies between nanowires and nanorings. Morphology transformed from A) nanowires to B) nanorings after adding TCEP. C) When system turned to oxidative environment again assemblies recover linear structure. D) Assembly model of interconversion between two-morphology states.

References for supporting information

(1) M. E. Bush, N. D. Bouley, A. R. Urbach, *J. Am. Chem. Soc.* 2005, **127**, 14511-14517.
(2) Z. Yassin, M. J. Clemente-Jiménez, R. Téllez-Sanz, L. García-Fuentes, *Int. J. Biol. Macromol.* **2003**, *31*, 155-162.
(3) K. T. Nam, B. R. Peelle, S. W. Lee, A. M. Belcher, *Nano Lett.* **2004**, *4*, 23-27.