

## Supporting Information

# 2D Square Arrays of Protein Nanocages through Channel-Directed Electrostatic Interaction with Poly( $\alpha$ , L-lysine)

Rui Yang, Lingli Chen, Senpei Yang, Chenyan Lv, Xiaojing Leng, and Guanghua Zhao \*

*CAU & ACC Joint-Laboratory of Space Food, College of Food Science & Nutritional Engineering, China Agricultural University, Beijing Key Laboratory of Functional Food from Plant Resources, Beijing, 100083, China*

*E-mail: [gzhao@cau.edu.cn](mailto:gzhao@cau.edu.cn)*

## Experimental Section

### 1. Protein Synthesis.

Recombinant soybean seed H-1 ferritin (rH-1) and H-2 ferritin (rH-2) were cloned, expressed, and purified as previously described.<sup>1,2</sup> The cDNA sequences of the H-1 $\Delta$ E subunit was subjected to PCR with primers, N-term (5'-GGAATTCCATATGGCCTCAACGGTGCCTCTC-3') and C-term (5'-CGCGGATCCTCACCTTCTCAACTGAGCC-3'). Mutagenesis of the H-2 cDNA was performed with the description of the Fast mutagenesis system kit (TransGen Biotech, Beijing, China). The specific site-directed mutation was designed for substitution of the H-2 His193 and His197 located at the E-helix of rH-2 by Glu or Ala (H-2\_H193E/H197E or H-2\_H193A/H197A) or substitution of the H-2 E165, E167, and E171 by Ile or Ala (H-2\_E165I/E167A/E171A). Recombinant proteins H-1 $\Delta$ E (rH-1 $\Delta$ E), ASF (rASF), H-2\_H193E/H197E (rH-2\_H193E/H197E), H-2\_H193A/H197A (rH-2\_H193A/H197A), and H-2\_E165I/E167A/E171A (rH-2\_E165I/E167A/E171A) were prepared according to our reported method.<sup>3</sup> The resulting inclusion bodies of the expressed rH-1 $\Delta$ E and rASF were denatured, refolded, and purified according to the reported method.<sup>4</sup> Reconstructed heteropolymeric proteins were conducted by mixing dissociated rH-2 (or rASF, or rH-2\_H193E/H197E, or rH-2\_H193A/H197A, or rH-2\_E165I/E167A/E171A) and rH-1 $\Delta$ E subunits in a 1:1 ratio at pH 11.4, followed by dialyzing against 50 mM 3-(N-morpholino) propanesulfonic acid (Mops) at pH 7.0, 4 °C,

in a dialysis bag (12-14 kDa MWCO). The concentrations of all types of ferritin were determined according to the Lowry method with BSA as a standard sample.

## 2. Ferritin Lattices Formation.

For protein self-assembly experiments, different volumes (0.0–3.0  $\mu\text{L}$ ) of PLL<sub>15</sub> (500  $\mu\text{M}$ ) were mixed with 500  $\mu\text{L}$  of 0.5  $\mu\text{M}$  apo rmSSF solution (50 mM Mops, 10.0 mM urea pre-cultivated with protein for 6 h, pH 7.0) to make a ratio of 0.0–6.0 PLL/ferritin in an eppendorf tubule. The reaction mixture was stirred for 10 min at room temperature  $\sim 25$  °C. Then, a cultivation time of 0, 30, 60, 90, 120, and 150 min at 25 °C was, respectively, conducted to obtain PLL<sub>15</sub>-induced ferritin assemblies. PLL<sub>5</sub>, PLL<sub>7</sub>, PLL<sub>10</sub> and PLL<sub>50</sub> were used as control samples to analyse the effect of PLL length on rmSSF self assembly following the same procedures as that of PLL<sub>15</sub>. PLR<sub>15</sub>, PLLG<sub>15</sub>, and PLG<sub>15</sub> were used as control samples to determine the effect of charge density on rmSSF self-assembly following the same procedures as that of PLL<sub>15</sub>. For protein self-assembly of rASF-H-1 $\Delta$ E (rmSSF\_H193E/H197E, rmSSF\_H193A/H197A, or rmSSF\_E165I/E167A/E171A) induced by PLL<sub>15</sub>, the reaction followed the same procedures for rmSSF self-assembly.

## 3. Preparation of Holo ferritin and Iron Release.

Holo rmSSF\_H193E/H197E sample (1.0  $\mu\text{L}$ ) was produced by adding approximately 600 iron atoms per ferritin in 6 increments at intervals of 0.5 h in 50 mM Mops, 50 mM NaCl, at pH 7.0. Iron release from holo ferritin was investigated as follows. Briefly, the assay system (1 ml of total volume) contained 0.5  $\mu\text{M}$  holo ferritin, 500  $\mu\text{M}$  ferrozine and 50 mM NaCl in 50 mM Mops buffer (pH 7.0). Reactions were carried out at 25 °C and were initiated by the addition of ascorbic acid (1 mM). The development of  $[\text{Fe}(\text{ferrozine})_3]^{2+}$  was measured by recording the increase in absorbance at 562 nm, and the iron release was measured using  $\epsilon_{562} = 27.9 \text{ mM}^{-1} \text{ cm}^{-1}$ .<sup>5</sup> The initial rate of iron release ( $v_o$ ) was obtained as previously described.<sup>5</sup>

## 4. TEM Analyses.

Liquid samples were diluted with 50 mM Mops buffer (pH 7.0) prior to placing on carbon-coated copper grids and excess solution removed with filter paper. Then, proteins were stained using 2% uranyl acetate for 5 min, with Fe (III) cores observation as an exception. TEM micrographs were imaged at 80 kV through a Hitachi H-7650 scanning electron microscope.

## 5. Fluorescence Titration.

Fluorescence titration experiments were performed using the Cary Eclipse spectrophotometer (Varian, USA). Measurements were carried out by an adding of 0.0–3.0  $\mu\text{L}$  of PLL (500  $\mu\text{M}$ ) to 500  $\mu\text{L}$  of the protein (0.5  $\mu\text{M}$ ) in 50 mM Mops, 10.0 mM urea, pH 7.0, followed by a reaction time of 10 min at 25 °C. Path lengths for

excitation and emission were 1.0 and 0.2 cm at wavelengths of 290 and 330 nm, respectively. A slit width of 5 nm and 10 nm was set for excitation and emission, respectively. To obtain the stoichiometry and binding constant, the data were fitted to Equation (1)<sup>6</sup> for the binding of PLL<sub>15</sub> to n independent binding sites on the protein.

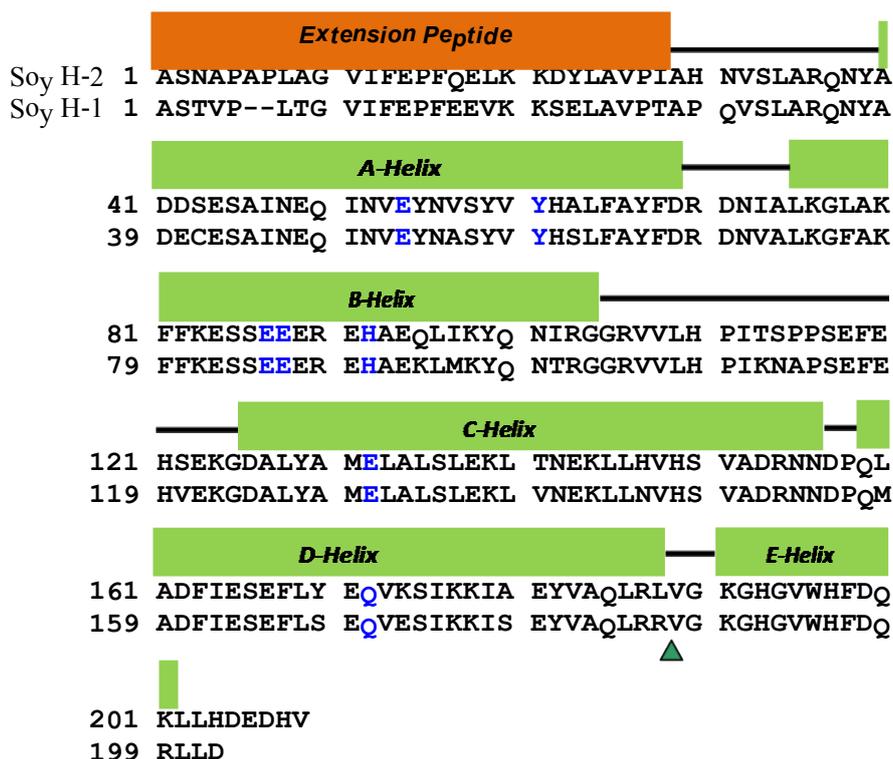
$$I = I_0 - \left( \frac{I_0 - I_\infty}{2n[P]_0} \right) \times [1/K + [PLL_{15}]_0 + n[P]_0 - \sqrt{(1/K + [PLL_{15}]_0 + n[P]_0)^2 - 4n[P]_0[PLL_{15}]_0}] \quad (1)$$

Here n is the binding site number and K is the apparent binding constant, [P]<sub>0</sub> and [PLL<sub>15</sub>]<sub>0</sub> are the 24-mer protein (rmSSF) and PLL<sub>15</sub> concentrations, and I<sub>0</sub> and I<sub>∞</sub> are the relative fluorescence intensities in the absence and presence of PLL<sub>15</sub> when the binding sites are fully saturated, respectively.

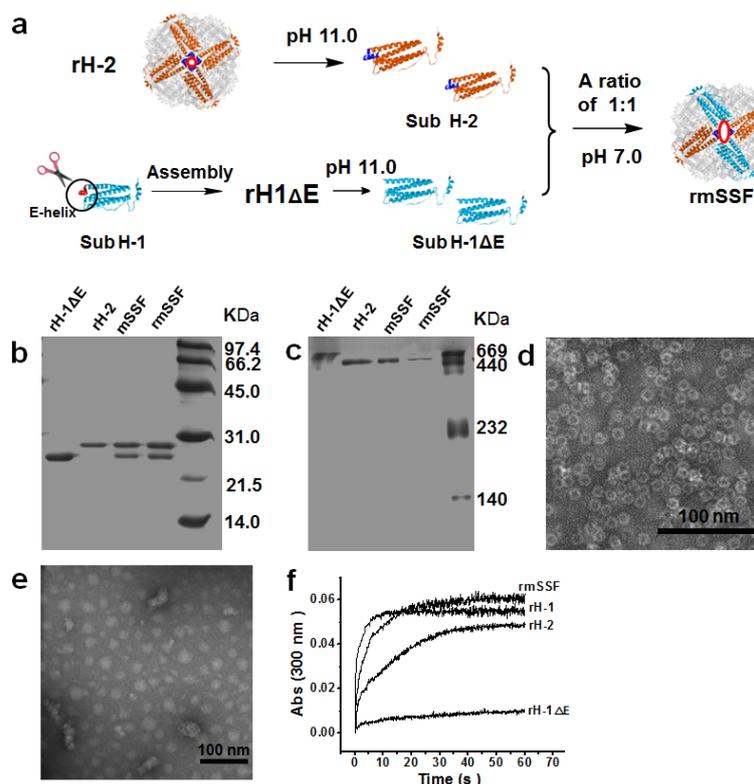
## 6. Dynamic Light Scattering (DLS) Analyses.

DLS experiments were performed at 25 °C using a Viscotek model 802 dynamic light scattering instrument (Viscotek, Europe) as described previously.<sup>7</sup> The OmniSIZE 2.0 software was used to calculate the size distribution of samples. All samples, unless stated otherwise, the final concentration at 0.5 μM, were allowed to stand for 2 h prior to DLS measurement to ensure that the reactions were complete.

Supplementary figures (Fig. S1-11)

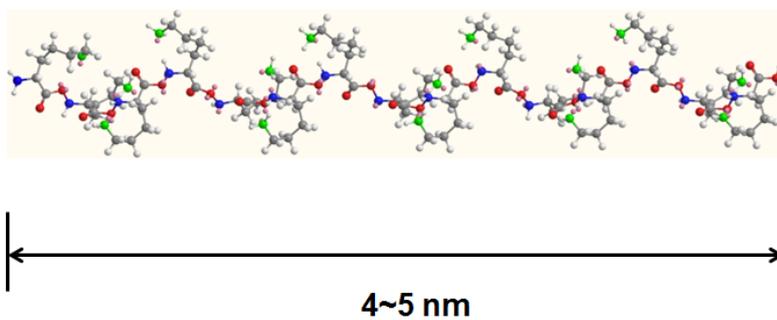


**Fig. S1** Amino acid sequences of mature soybean seed ferritin subunits. Top, amino acid sequence of the H-2 subunit. Bottom, amino acid sequence of the H-1 subunit. Conserved residues between the two subunits are shown in black. Residues in blue indicate the deduced ferroxidase center. The N-terminal EP residues of the H-2 subunit are shown boxed in orange. The mature regions of both subunits are downstream from here. The cleavage site for conversion of the H-1 to H-1ΔE subunit is indicated by a green arrowhead.

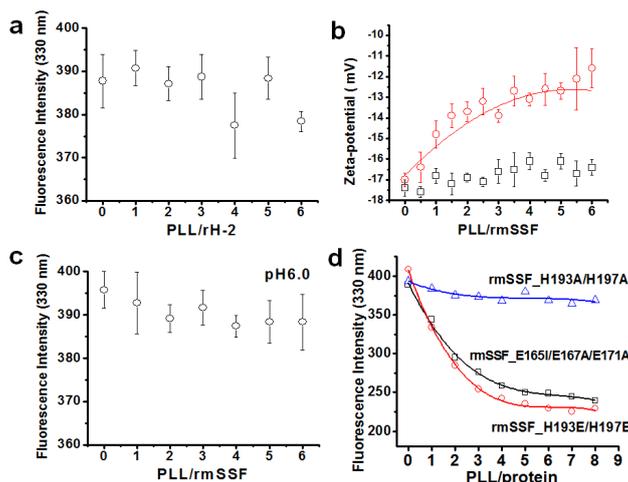


**Fig. S2** Preparation and characterization of rmSSF cage. (a) A scheme for rmSSF formation. (b) SDS-PAGE and (c) Native-PAGE analyses of rH-1ΔE, rH-2, mSSF, and rmSSF. (d) TEM images of rH-2. (e) TEM images of rH-1ΔE. (f) Kinetic curves of  $\text{Fe}^{2+}$  oxidation by  $\text{O}_2$  catalyzed by different types of ferritins. Iron initial rates were  $0.35 \pm 0.02 \mu\text{M Fe subunit}^{-1} \text{s}^{-1}$  for rmSSF, a value midway between values of rH-1 ( $0.74 \pm 0.02 \mu\text{M Fe subunit}^{-1} \text{s}^{-1}$ ) and rH-2 ( $0.17 \pm 0.01 \mu\text{M Fe subunit}^{-1} \text{s}^{-1}$ ), while rH-1ΔE only has marginal ferroxidase activity due to lack of ferritin-like structure.

Through genetic engineering and chemical approaches, we have synthesized two different subunits of normal H-2 (uncleaved 28-kDa subunit form) and H-1ΔE where the C-terminal 16 amino acid residues corresponding to the E helix are removed (Fig. S1, ESI†). Heteropolymeric rmSSF with expanded 4-fold channels was prepared by taking advantages of the reversible dissociation and reassembly characteristic of ferritin in different pH environments as shown in Fig. S2a, ESI†. We dissociated both recombinant H-1ΔE (rH-1ΔE) and recombinant H-2 (rH-2) at pH 11.0, then mixed these two dissociated subunits in a 1:1 ratio, and finally slowly decreased the pH value of the mixture to 7.0 to induce rmSSF reassembly. SDS and native PAGE analyses showed that rmSSF has an apparent MW nearly the same as rH-2 and mature soybean seed ferritin (mSSF), but is composed of two subunits of 28.0 and 26.5 kDa with a 1:1 ratio (Fig. S2b and c, ESI†), indicative of a successful preparation. TEM analysis showed that both rH-2 (Fig. S2d, ESI†) and rmSSF (Figure 2b) have identical shell-like structure with exterior diameter of approximately 12 nm, whereas rH-1ΔE alone is not able to form such structure (Fig. S2e, ESI†). Further characterization of rmSSF is carried out by the fact that rmSSF exhibits a greater catalysing activity of iron oxidation than its analogue, rH-2 shown in Fig. S2f, ESI†.



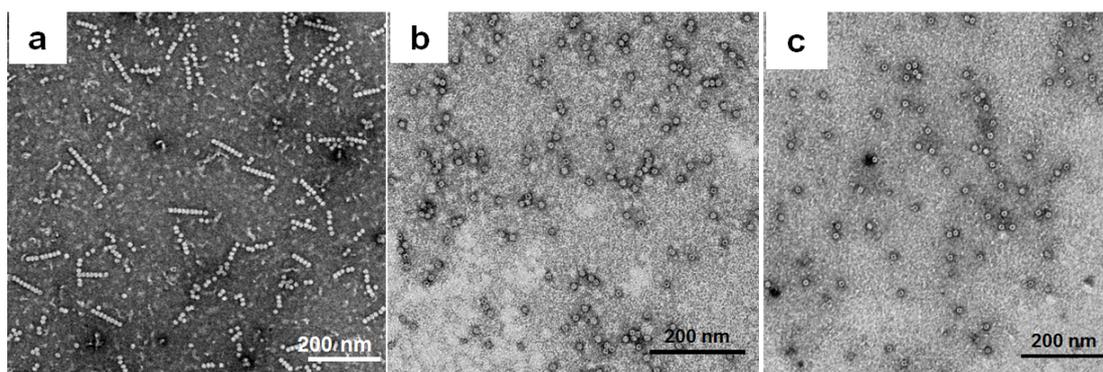
**Fig. S3** A schematic diagram of PLL<sub>15</sub> represented by software *Chem3D Ultra 12* in a minimum energy state.



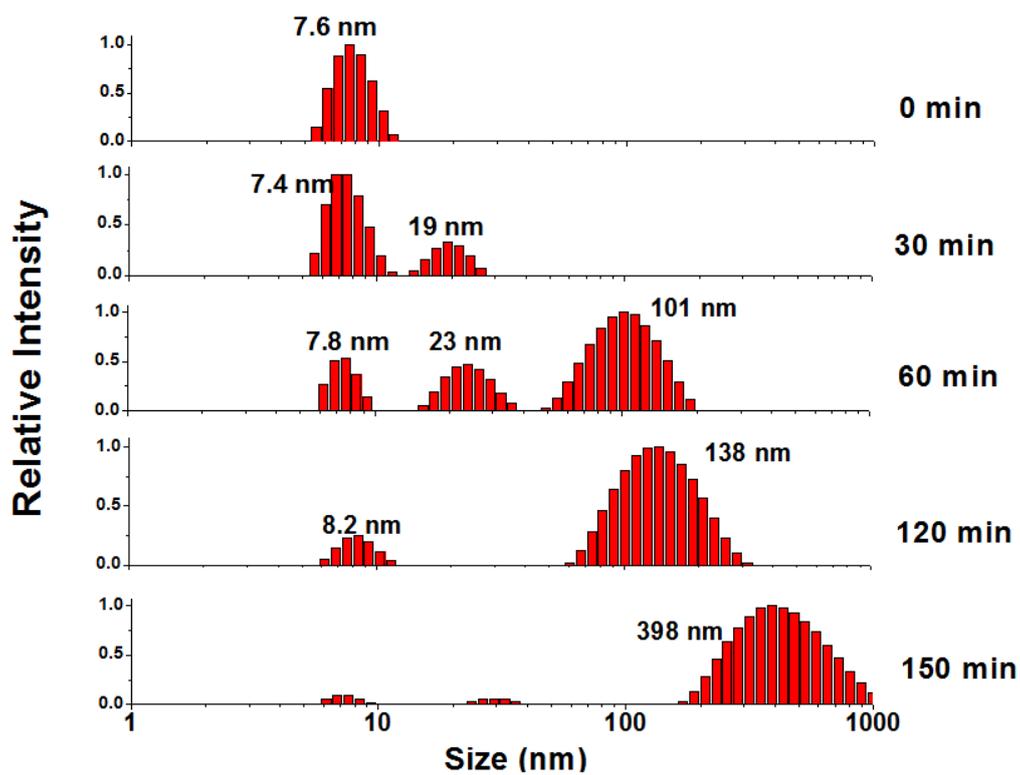
**Fig. S4** (a) Fluorescence titration and (b) Zeta-potential analysis of rmSSF or rH-2 upon treatment with different amounts of PLL<sub>15</sub> in the presence of urea (10.0 mM). (c) Fluorescence titration of rmSSF upon treatment with different amounts of PLL<sub>15</sub> in the presence of urea (10.0 mM) at pH 6.0. (d) Fluorescence spectrum of three rmSSF mutants upon treating with different amounts of PLL<sub>15</sub> in the presence of 10.0 mM of urea, respectively. Conditions: [apo rmSSF, rmSSF\_H193E/H197E, rH-2\_H193A/H197A, or rmSSF\_E165I/E167A/E171A] = 0.5 μM, [PLL<sub>15</sub>] = 0-3.0 μM, in 50 mM Mops (pH 6.0 or 7.0), 25 °C.

As shown in Figure 3a, two His residues (His193 and His197) from H-2 subunit of rmSSF are located at the 4-fold channels close to Trp196. If PLL<sub>15</sub> molecules bind to rmSSF through the 4-fold channels, we would expect no such fluorescence quenching at pH 6.0 due to positively electrostatic exclusion between PLL<sub>15</sub> and His residues ( $pK_a \sim 6.4$ ). Indeed, fluorescence quenching was slightly observed with rmSSF upon the addition of PLL<sub>15</sub> to the protein at pH 6.0 (Fig. S4c, ESI†), again demonstrating that the binding of PLL<sub>15</sub> to rmSSF is carried out through the 4-fold channels. To obtain more evidence, rmSSF\_H193A/H197A, an rmSSF mutant where His197 and His193 of H-2 subunit were both mutated into Ala residue, was prepared to “block” the channel by changing its hydrophilic environment to hydrophobic one. The same fluorescence titration was carried out with this mutant, and results showed that no fluorescence quenching was observed (Fig. S4d, ESI†), indicative of the importance of the charged residues around the channel for such binding. On the other hand, rmSSF\_H193E/H197E, another rmSSF mutant where His197 and His193 of H-2 subunit were both mutated into Glu residue ( $pK_a \sim 4.0$ ), was prepared, followed by fluorescence titration with PLL<sub>15</sub> under the same conditions. Results showed that PLL<sub>15</sub> can quench the fluorescence of this mutant at pH 6.0 with a PLL<sub>15</sub>/protein stoichiometry of 3:1 (Fig. S4d, ESI†), exhibiting a completely different PLL<sub>15</sub>-binding behavior from its analogue, rmSSF. Such distinction can be well explained by different charges on their 4-fold channels; namely, the 4-fold channels of rmSSF take positive charge at pH 6.0, whereas rmSSF\_H193E/H197E has negative charges on its 4-fold channels due to a much lower  $pK_a$  value ( $\sim 4.0$ ) of Glu residues. It appears that the electrostatic attraction between PLL<sub>15</sub> molecules and residues located the 4-fold channels further enhances the interaction of PLL<sub>15</sub> and rmSSF. These results confirm the conclusion that PLL<sub>15</sub> molecules interact with rmSSF through the 4-fold channels. Indeed, the fluorescence of a 3-fold channel mutant of rmSSF,

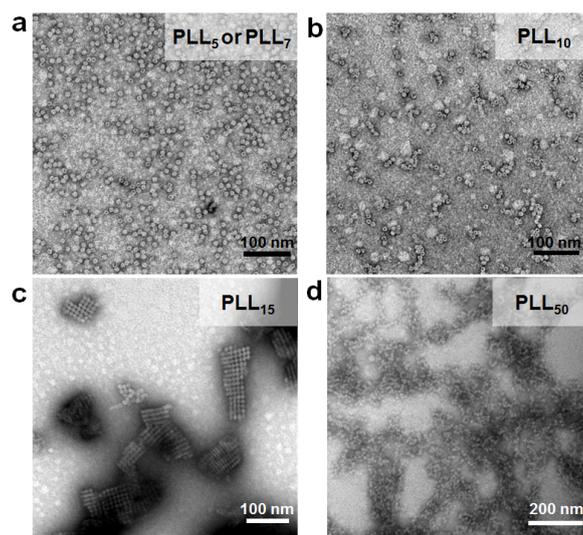
rmSSF\_E165I/E167A/E171A (E165, E167, and E171 locating on the 3-fold channel of H-2 subunit were mutated into Ile or Ala residues) whose 4-fold channels are not altered, can be likewise quenched by PLL<sub>15</sub> in the presence of urea (10.0 mM) (Fig. S4d, ESI<sup>†</sup>), indicating that this 3-fold channel mutant did not influence the specific binding between protein and PLL, and the 3-fold channels are not the binding site of PLL. All these results proved that it is the 4-fold channels rather than the 3-fold channels are the binding site of PLL<sub>15</sub>.



**Fig. S5** (a) TEM image of rmSSF upon treatment with PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 for 2 h in the absence of urea. (b) TEM image of rH-2 upon treatment with PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 for 2 h in the absence of urea. (c) TEM image of rH-2 upon treatment with PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 for 2 h in the presence of urea (10.0 mM). Conditions: [apo rmSSF or rH-2] = 0.5  $\mu$ M, [PLL<sub>15</sub>] = 1.5  $\mu$ M, in 50 mM Mops (pH 7.0), 25 °C.

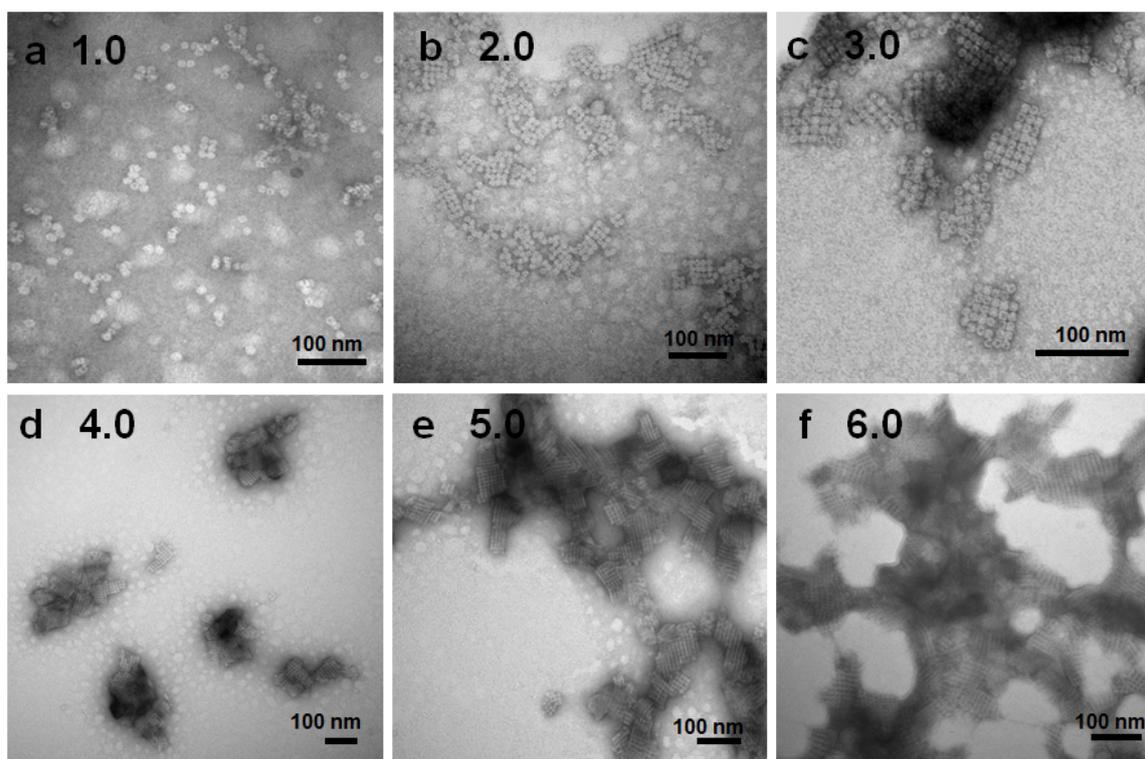


**Fig. S6** Time courses of size distribution of rmSSF assemblies upon treatment with PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 in the presence of urea (10.0 mM). Conditions: [apo rmSSF] = 0.5  $\mu$ M, [PLL<sub>15</sub>] = 1.5  $\mu$ M, in 50 mM Mops (pH 7.0), 25 °C.



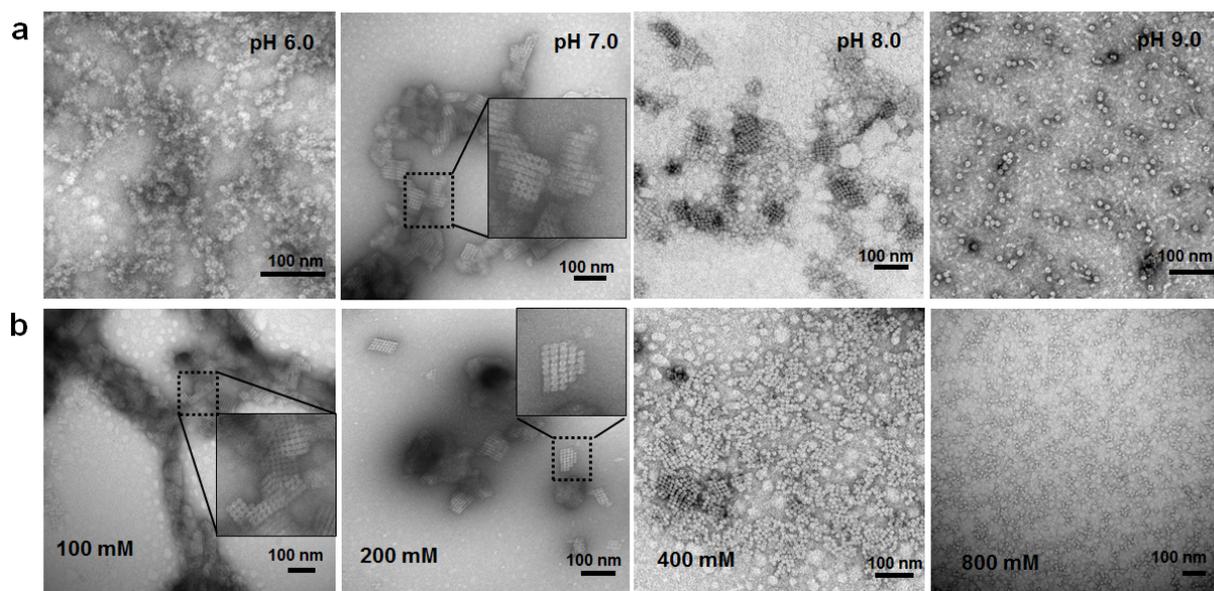
**Fig. S7** TEM images of rmSSF assemblies induced by PLL molecules (PLL<sub>5</sub>, PLL<sub>7</sub>, PLL<sub>10</sub>, PLL<sub>15</sub>, and PLL<sub>50</sub>) (PLL/protein = 3:1) in the presence of urea (10.0 mM).

PLL molecules with less polymerization degrees such as 5, 7, and 10 (PLL<sub>5</sub>, PLL<sub>7</sub>, and PLL<sub>10</sub>), medium polymerization degree of 15 (PLL<sub>15</sub>), and a larger polymerization degree (PLL<sub>50</sub>) were incubated with apo rmSSF in the presence of urea, respectively, at the ratio of 3:1 under the same experimental condition. TEM analyses showed that no such protein lattice was formed but occupying by monoferritins with PLL<sub>5</sub> or PLL<sub>7</sub> (Fig. S7a, ESI†), or some small irregular polymers with PLL<sub>10</sub> (Fig. S7b, ESI†). Subsequently, only PLL<sub>15</sub> molecules with medium size can trigger such lattices formation (Fig. S7c, ESI†). As for PLL<sub>50</sub> treatment with rmSSF in the presence of urea (10.0 mM), results showed that non-specifically protein polymerization appeared, and instead various irregular protein assemblies were observed as shown in Fig. S7d, ESI†. Thus the length of the polypeptide about 4-5 nm is very strict for the ferritin regular array. The reason may lie in the structures of both ferritin and cationic polypeptide. The crystal structure shows the thickness of the ferritin shell is around 2 nm. The length of PLL<sub>15</sub> molecule is approximate 4-5 nm (Fig. S3, ESI†), which is about 2 times longer than the protein shell thickness. Such length is long enough to bridge two protein cages together by inserting into their corresponding two 4-fold channels through electrostatic attraction. In contrast, the length of the PLL with the polymerization degree  $\leq 10$  is less than 4 nm, which is too short to bridge the two protein cage together. Based on these results, we believe that an appropriate length of PLL is prerequisite for the fabrication of the protein lattices from the standpoint of PLL structure.



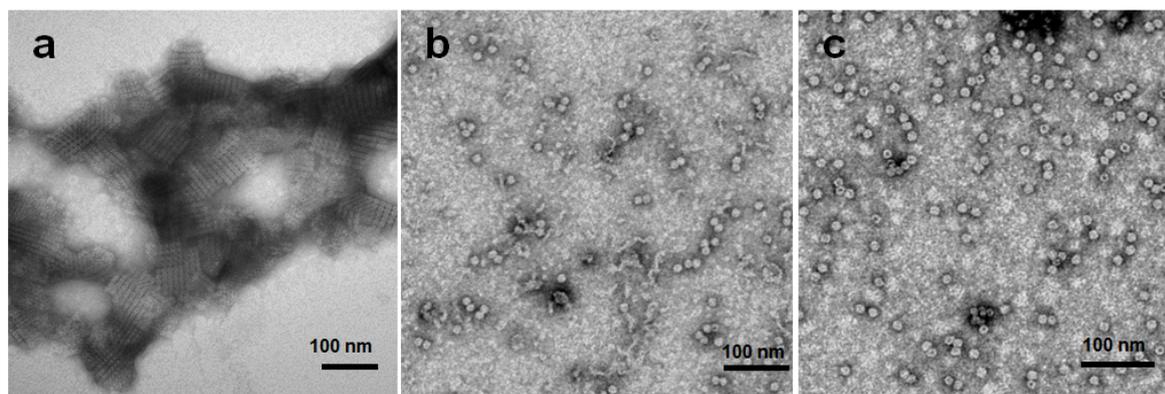
**Fig. S8** TEM images of rmSSF assemblies upon treatment with different amounts of PLL<sub>15</sub> for 2 h in the presence of urea (10.0 mM), and the ratio of PLL<sub>15</sub>/protein varies from 1/1 to 6/1. Conditions: [apo rmSSF] = 0.5  $\mu$ M, [PLL<sub>15</sub>] = 0.5-3.0  $\mu$ M, in 50 mM Mops (pH 7.0), 25 °C.

TEM imaging reveals that the PLL<sub>15</sub>/protein ratio plays an important role in the size and aggregation extent of ferritin lattices in Fig. S8, ESI†. As the PLL<sub>15</sub>/apo rmSSF ratio increases from 1.0 to 6.0 at fixed reaction time as 2 h, the morphology of ferritins changes from mainly monodispersed ferritins, to small planar lattices usually with unbonded monoferritins, and to larger ferritin aggregate species (Fig. S8a–c, ESI†). Beyond a PLL<sub>15</sub>/protein ratio of 3:1, ferritin lattices often cross-linked with each other, forming lattices stack clusters (Fig. S8d–f, ESI†), accompanying with an obvious extension of the regular lattice size. Thus, the increase in the PLL/protein ratio (> 2) not only influences the size of the lattices but also promotes the protein aggregation.



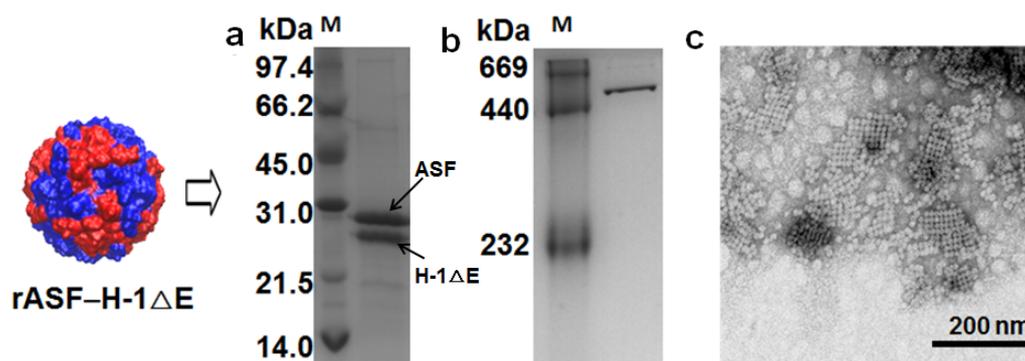
**Fig. S9** (a) TEM images of rmSSF assemblies induced by PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 in the presence of urea (10.0 mM) at different pH values. (b) TEM images of rmSSF assemblies induced by PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 in the presence of urea (10.0 mM) at different ionic strengths (presented by different concentrations of NaCl). Conditions: [apo rmSSF] = 0.5  $\mu$ M, [PLL<sub>15</sub>] = 1.5  $\mu$ M, in 50 mM Mops (pH 7.0), 25 °C.

We found that only at neutral pH (7.0 or 8.0) (Fig. S9a, ESI<sup>†</sup>) and low or medium ionic strength (0–400 mM) (Fig. S9b, ESI<sup>†</sup>), regular lattices of rmSSF can be observed. By contrast, both low pH (6.0) or high pH (9.0) and high ionic strength (800 mM NaCl) inhibited the generation of such two dimensional protein cage array, approving the above conclusion that the electrostatic interaction is responsible for the binding of PLL<sub>15</sub> with rmSSF.



**Fig. S10** TEM images of rmSSF assemblies induced by (a) PLR<sub>15</sub>, (b) PLLG<sub>15</sub>, and (c) PLG<sub>15</sub> in a 3:1 ratio (polypeptide/protein) in the presence of urea (10.0 mM). Conditions: [apo rmSSF] = 0.5  $\mu$ M, [PLL<sub>15</sub>, PLR<sub>15</sub>, or PLLG<sub>15</sub>] = 1.5  $\mu$ M, in 50 mM Mops (pH 7.0), 25 °C.

To better understand the assembling mechanism of protein lattices, effect of the charge density of polypeptide on such assembly was also investigated. As expected, we found that PLL<sub>15</sub> analogue, poly( $\alpha$ , L-arginine) with polymerization degree of 15 (PLR<sub>15</sub>), could also trigger nearly the same protein lattices as PLL<sub>15</sub> does (Fig. S10a, ESI<sup>†</sup>). Other analogues with lower charge density, however, poly( $\alpha$ , L-lysine-glycine) (PLLG<sub>15</sub> where two amino acid lysine and glycine are arranged alternatively) and poly( $\alpha$ , L-glycine) (PLG<sub>15</sub>), lose such ability (Fig. S10b and c, ESI<sup>†</sup>), demonstrating that the charge density of the polypeptide greatly influences the observed lattices formation.



**Fig. S11** Applicability of the channel-directed electrostatic interactions in rASF-H-1ΔE. (a–c) show (a) SDS-PAGE, (b) native-PAGE analyses of rASF-H-1ΔE, and (c) TEM image of rASF-H-1ΔE protein lattices induced by PLL<sub>15</sub> in a PLL<sub>15</sub>/protein ratio of 3:1 in the presence of urea (10.0 mM). Conditions: [apo rASF-H-1ΔE] = 0.5 μM, [PLL<sub>15</sub>] = 1.5 μM, in 50 mM Mops (pH 7.0), 25 °C.

Another ferritin which consists of 12 of rH-1ΔE subunits and 12 subunits of recently reported *adzuki* bean seeds ferritin (ASF, a homopolymeric plant ferritin),<sup>8</sup> which is named as rASF-H-1ΔE (Fig. S11a and b, ESI†), was prepared in a similar way. We found that PLL<sub>15</sub> also could bridge this ferritin to generate protein lattices in the presence of urea (Fig. S11c, ESI†). Therefore, this strategy with rationally engineered principal can be applied to other hybrid protein building block material.

## References

- 1 C. Li, X. Fu, X. Qi, X. Hu, N. D. Chasteen, G. Zhao, *J. Biol. Chem.* 2009, **284**, 16743–16751.
- 2 T. Masuda, F. Goto, T. Yoshihara, T. Ezure, T. Suzuki, S. Kobayashi, M. Shikata, S. Utsumi, *Protein Expr. Purif.* 2007, **56**, 237–246.
- 3 R. Yang, L. Chen, T. Zhang, S. Yang, X. Leng, G. Zhao, *Chem. Commun.* 2014, **50**, 481–483.
- 4 O. W. Wuytswinkel, G. Savino, J. F. Briat, *Biochem. J.* 1995, **305**, 253–261.
- 5 M. J. Hynes, M. O. Coinceanainn, *J. Inorg. Biochem.* 2002, **90**, 18–21.
- 6 F. Bou-Abdallah, G. Zhao, G. Biasiotto, M. Poli, P. Arosio, N. D. Chasteen, *J. Am. Chem. Soc.* 2008, **130**, 17801–17811.
- 7 H. Yang, X. Fu, M. Li, X. Leng, B. Chen, G. Zhao, *Plant Physiol.* 2010, **154**, 1481–1491.
- 8 M. Li, S. Yun, X. Yang, G. Zhao, *Biochim. Biophys. Acta* 2013, **1830**, 2946–2953.