– – Electronic Supplementary Material (ESI) for Chemical Communications.
This journal is © The Royal Society of Chemistry 2019

Supporting Information

# AB loop engineered ferritin nanocage for drug loading under benign experimental conditions

Wenming Wang,*[a] Lele Wang,[a] Guobang Li,[c] Guanghua Zhao,*[b] Xuan Zhao,[a] and Hongfei Wang*[a]

[a] Key Laboratory of Chemical Biology and Molecular Engineering of Education Ministry, Key Laboratory of Energy Conversion and Storage Materials of Shanxi Provence, Institute of Molecular Science, Shanxi University, Taiyuan 030006, China.

[b] College of Food Science & Nutritional Engineering, China Agricultural University, Beijing 100083, China

[c] Drug Discovery Center for Infectious Disease and State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin, People's Republic of China.

*Corresponding author.

E-mail address: wangwm@sxu.edu.cn

E-mail address: gzhao@cau.edu.cn

E-mail address: wanghf@sxu.edu.cn

## Experimental Section

### 1. Preparation of Ferritins

Protein sequences for each ferritin variant is provided above. rHuHF, rHuHFΔ2 and rHuHFΔ3 were cloned into the pGEX-6p-1 plasmid (GE Healthcare) using the restriction enzymes *Bam*HI and *Xho*I. Proteins were then expressed in *E.coli* BL21 (DE3) cells. The transformed cells were grown at 37 °C until $OD_{600}$ = 0.6, and protein expression was induced using 0.2 mM isopropyl-D-thiogalactopyranoside (IPTG). All of the proteins were expressed at 25 ºC for overnight. Harvested cells were resuspended in lysis buffer containing 50.0 mM Tris (pH 8.0) and 200.0 mM NaCl, followed by sonication at 4 °C and collection by ultracentrifugation for 30 min at 40,000 g at 4 °C to remove cell debris. The supernatant was then loaded twice onto a GST column (Sangon Biotech) pre-equilibrated with lysis buffer, and the GST-tag was removed by digestion with PreScission protease (GE Healthcare) overnight at 4 °C. The eluted rHuHF, rHuHFΔ2 and rHuHFΔ3 were further purified using HiTrap™ Q HP (GE Healthcare) and Superdex 200 10/300 GL columns (GE Healthcare). SDS-PAGE analysis revealed over 95 % purity of the final purified recombinant proteins. The purified proteins were then concentrated to 5.0 mg/mL in a buffer containing 50.0 mM Tris (pH 8.0).

### 2. Blue Native PAGE

Samples were run on 6 % native gels following the Blue Native PAGE manual[1]. rHuHF, rHuHFΔ2 and rHuHFΔ3 stored in different pH buffer solutions were mixed with 4× native PAGE sample buffer (200.0 mM Tris, pH 7.0, 50 % w/v glycerol, and 0.01 % w/v bromophenol blue G250). Gels were then stained with Coomassie Brilliant Blue R250. All of the native PAGE gels were repeated independently at least twice with each repeat giving similar results.

Different pH buffer solutions: 50.0 mM citric acid (pH 2.0); 50.0 mM citric acid (pH 3.0); 50.0 mM sodium acetate (pH 4.0); 50.0 mM sodium acetate (pH 5.0); 50.0 mM MES (pH 6.0); and 50.0 mM Hepes (pH 7.0).

**3. Dynamic Light Scattering (DLS)**

Dynamic light scattering (DLS) experiments were performed on a Malvern Zetasizer Nano ZS in disposable polystyrene micro cuvettes (VWR) using 1500 μL of freshly prepared sample solution. Generally, for different pH, proteins were diluted into buffers with different pH values and the final concentration of the protein was 0.25 mg/mL. These buffers are 50.0 mM citric acid (pH 2.0), 50.0 mM citric acid (pH 3.0), 50.0 mM sodium acetate (pH 4.0), 50.0 mM sodium acetate (pH 5.0), 50.0 mM MES (pH 6.0), 50.0 mM Hepes-NaOH (pH 7.0), 50.0 mM Tris-HCl (pH 8.0), 50.0 mM CAPSO (pH 9.0), 50.0 mM CAPS (pH 10.0), 50.0 mM CAPS (pH 11.0). After equilibration at 25 °C, three measurements were performed with the instrument optimizing the number of runs for each measurement. The refractive index (RI) of each dispersant (preset: water) was set to 1.330 and the viscosity to 0.8872 cP, respectively. The RI of each particle was set to 1.45. The absorption of the protein was set to 0.00, and both attenuator and measurement positions were controlled by the instrument and all measurements were performed at a scattering angle of 173°. All of the tests were repeated independently three times with each repeat giving similar results.

**4. Transmission Electron Microscopy (TEM)**

Six samples were prepared, with the final concentrations of 50.0 μg/mL. These were: (1) rHuHF in a buffer containing 50.0 mM citric acid (pH 3.0), (2) rHuHF in a buffer containing Hepes-NaOH (pH 7.0), (3) rHuHFΔ2 in a buffer containing 50.0 mM citric acid (pH 3.0), (4) rHuHFΔ2 in a buffer containing Hepes-NaOH (pH 7.0), (5) rHuHFΔ2 in a buffer containing 50.0 sodium acetate (pH 4.0), (6) rHuHFΔ2 in a buffer containing Hepes-NaOH (pH 7.0). Samples were placed on carbon-coated copper grids, after excess sample solution was removed with filter paper. These samples were stain using 2 % uranyl acetate for 40 s. TEM data were collected using a Field emission transmission electron microscope (Talos F200C) operating at 200 kV.

**5. Crystallization and X-ray data Collection**

rHuHFΔ2 and rHuHFΔ3 crystals were grown in buffer containing 50.0 mM Bicine (pH 9.0) and 2.0 M $MgCl_2$, using the sitting-drop vapor diffusion method in 24-well plates. Then, 1.0 μL drops of the protein samples (5.0 mg/mL) were mixed with an equal volume of reservoir solution and the mixtures were equilibrated against 200.0 μL reservoir solution at 18 °C, which was similar to rHuHF. The crystals were sent to SSRF (BL18U1) and diffraction data were collected to resolutions of 1.8 Å for rHuHFΔ2 and 2.3 Å for rHuHFΔ3, respectively. Crystals were harvested and cryo-protected in a well solution containing 20 % (v/v) glycerol, 50.0 mM Bicine (pH 9.0) and 2.0 M $MgCl_2$, and then flash-cooled in a dry nitrogen stream at 100 K for X-ray data collection.

**6. X-ray data Processing and Structure Determination**

X-ray data were processed, merged and scaled using the HKL-3000 (HKL Research).[2] The structures of rHuHFΔ2 and rHuHFΔ3 were determined by molecular replacement using coordinates of rHuHF (PDB code: 2FHA) as an initial model using Phaser program in the Phenix program package.[3] Structure refinement was conducted using the program Phenix.refine software.[4] The structures were rebuilt using COOT[5] which made the models manually adjusted. Model geometry was verified using the program MolProbity.[6] Most residues appeared in the favored region and no residues were disallowed in the Ramachandran plot. Structural figures were drawn using the program PyMOL.[7] Data collection and the final refinement statistics are summarized in Table S1.

**7. The Recovery Rate of rHuHFΔ3**

Recovery rate experiments were performed utilizing a Cary 50 UV-Vis spectrophotometer to record changes in absorbance at 280 nm. Undissociated rHuHFΔ3 (0.3 mg/mL) was firstly measured. Reassembled rHuHFΔ3 was then prepared by adjusting the pH value from 7.0 to 4.0 then back to 7.0 again, and then the absorption of reassembled rHuHFΔ3 was measured after centrifugation and filtration.

**8. Protein Thermal Shift**

The melting temperature of proteins was measured by thermal scanning coupled with fluorescence detection using a quantitative PCR instrument (CFX96, Bio-Rad). The fluorescence (excitation wavelength at 498 nm and emission wavelength at 610 nm) of the protein solutions (4.0 μM protein in 50.0 mM HEPES, pH 7.0, plus 5×SYPRO Orange dye (Sigma) in a final volume of 20.0 μL) was measured as a function of temperature at a climbing rate of 0.5 °C/min from 25 to 95 °C. The fluorescence was then recorded every 10 s. The data were analyzed using the Boltzmann equation in Origin version 8.0 (OriginLab Corp.).

**9. Encapsulation of Doxorubicin within rHuHFΔ3**

Doxorubicin encapsulation experiments were conducted using methods previously reported [8,9] with some modifications. Doxorubicin was added to rHuHFΔ3 (50.0 mM Tris, pH 8.0), and the final concentration of doxorubicin was 400.0 μM, which was ten times than rHuHFΔ3. After the pH was adjusted to 4.0 from 8.0 by 1.0 M HCl, the mixture was gently vortexed for 20 min at room temperature to ensure complete dissociation. Then, the mixture pH was increased to 7.0 using NaOH (1.0 M). Finally, in order to remove excess doxorubicin, the solution was dialyzed 5 times through a 100 kDa dialysis bag at 4 °C.

**10. Encapsulation of Curcumin within rHuHFΔ3**

Briefly, 200.0 μM curcumin (dissolved in DMSO) was slowly trickled (dropwise) into rHuHFΔ3 (5.0 mL, 2.0 μM) and the mixture was adjusted slowly to pH 4.0 with HCl (1.0 M). The solution was then stirred for 30 min to disassemble ferritin into subunits, and then the pH value was adjusted back to 7.0 using NaOH (1.0 M). Then, the resultant mixture was allowed to stand at 4 °C for 1 h. Finally, the resultant solution was dialyzed (100 kDa cutoff) against 20.0 mM Hepes (pH 7.0) three times at intervals of 6 h to remove free curcumin.

**Supplementary figures and table**

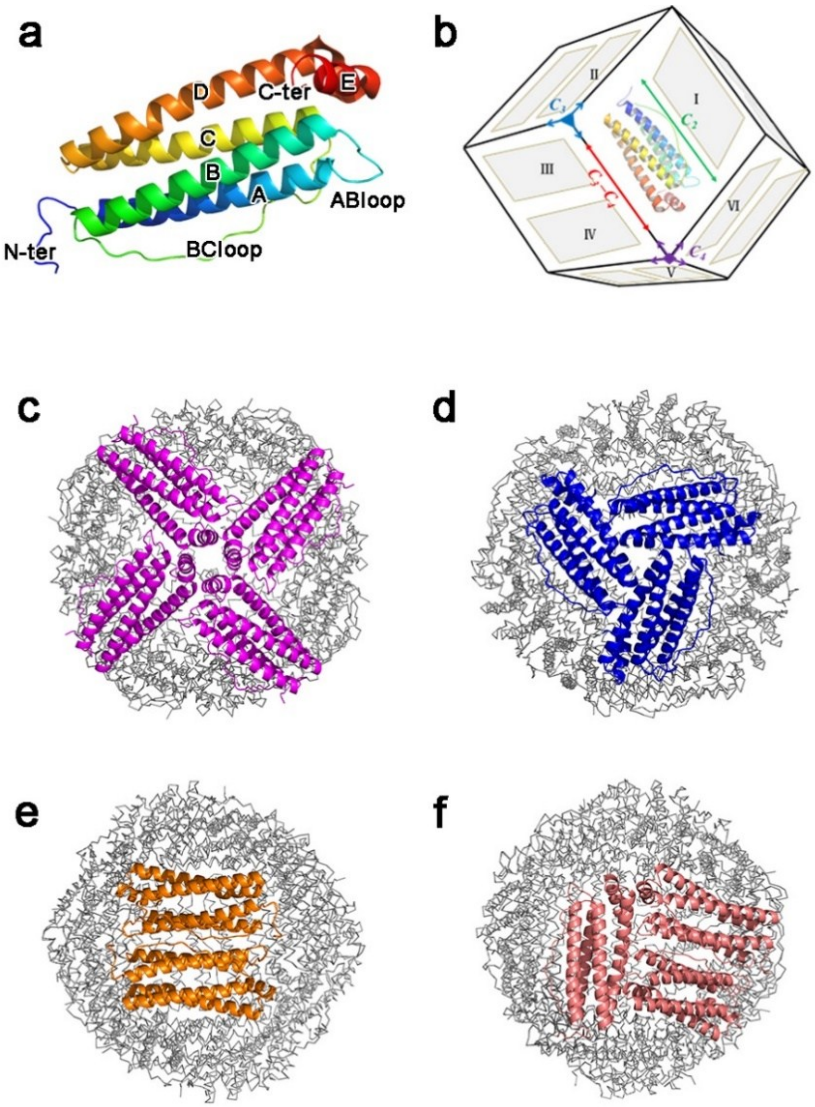


**Fig. S1** (a) Ribbon diagram of the subunit of human H chain ferritin (HuHF), showing the main secondary structure elements of ferritin subunits, each subunit is composed of a four-α-helix bundle containing two antiparallel helix pairs (A, B and C, D) and a fifth short helix (E helix). (b) Schematic representation of intersubunit interfaces of ferritin cage model with the approximate geometry of a rhombic dodecahedron symmetry, and each ferritin molecule contains four kinds of intersubunit interfaces responsible for its shell-like assembly, namely, (c) 6 of $C_4$ interfaces, (d) 8 of $C_3$ interfaces, (e) 12 of $C_2$ interfaces and (f) 24 of $C_3$-$C_4$ interfaces.

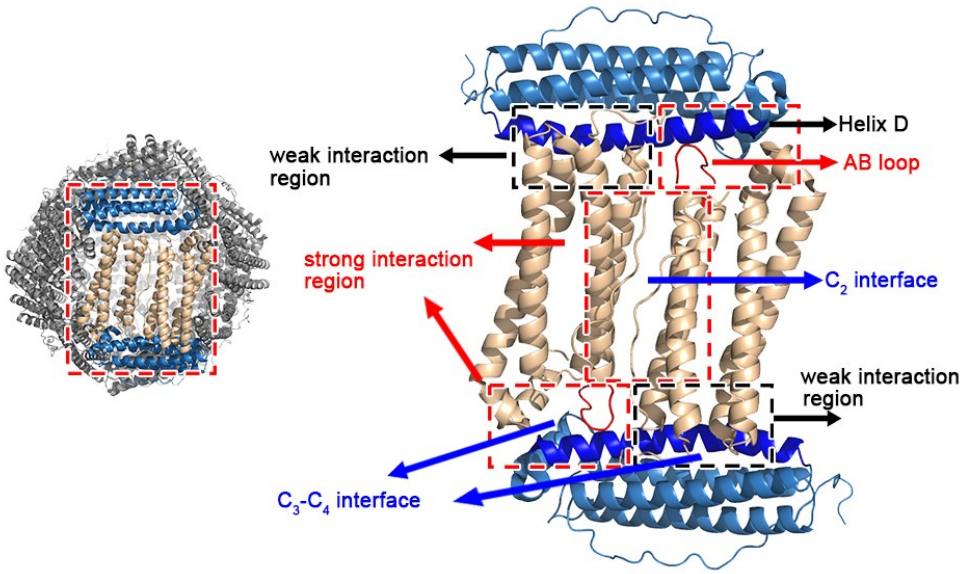


**Fig. S2** Interaction analysis diagram of $C_2$ and $C_3$-$C_4$ interface. Four ferritin subunits are in cartoon representation and colored in skyblue and wheat, $C_2$ and $C_3$-$C_4$ interface are also indicated with blue arrow. Three strong interaction regions were framed with red box while two weak interaction regions with black frames. Helix D and AB loop are colored with red and blue.

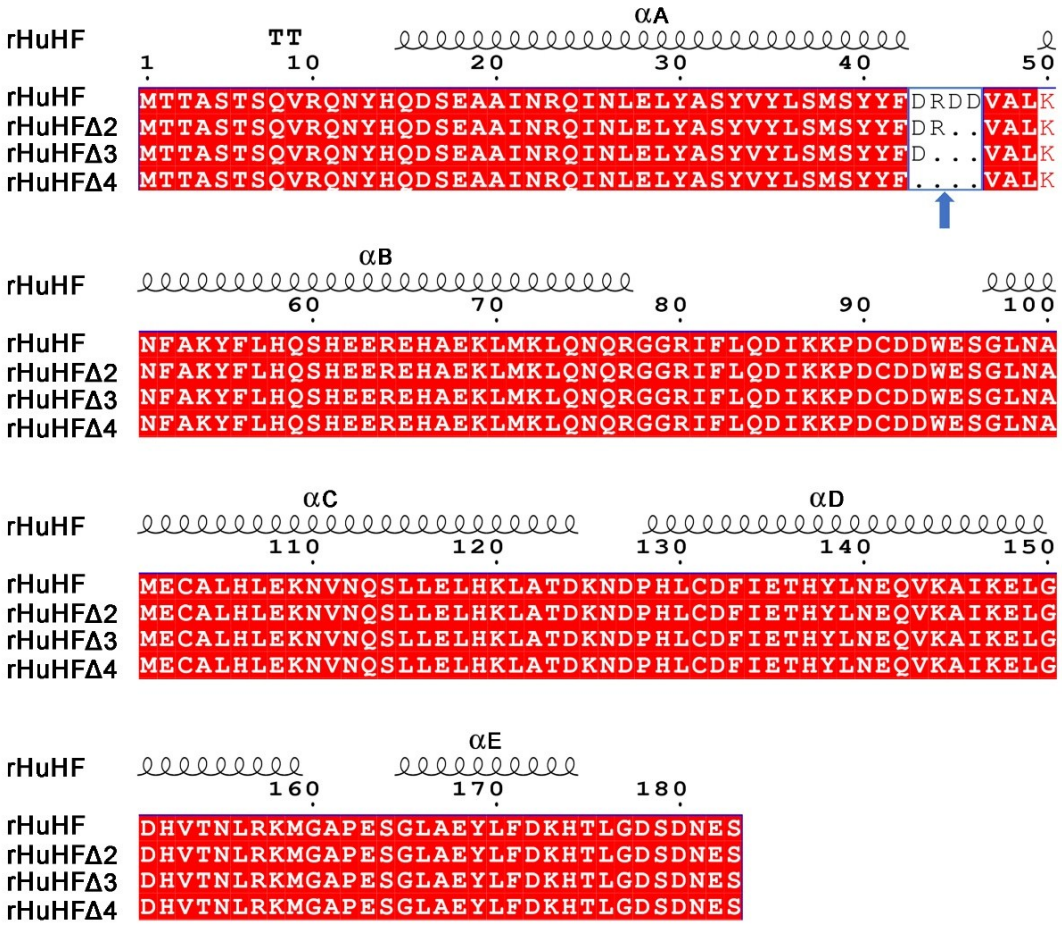


**Fig. S3** Sequence alignment of rHuHF, rHuHFΔ2 and rHuHFΔ3. Secondary structures of rHuHF are also shown. The residues reengineered in AB loops are indicated by blue arrows.

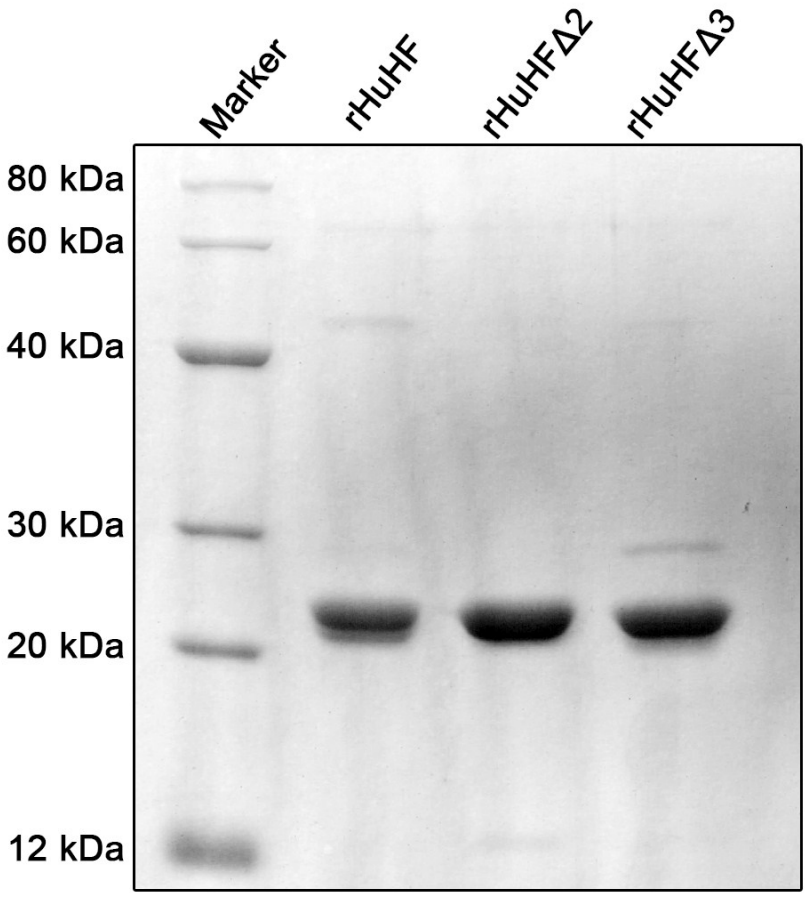


**Fig. S4** SDS polyacrylamide gel electrophoresis (SDS-PAGE) analyses of rHuHF, rHuHFΔ2 and rHuHFΔ3 after protein purification. Lane M, protein markers and their corresponding molecular masses.

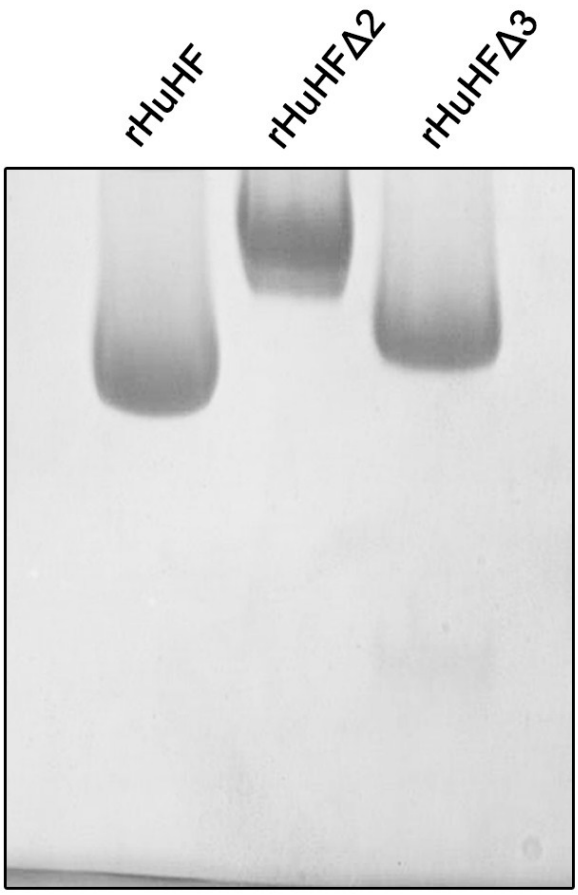


**Fig. S5** 6% Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) analyses of rHuHF, rHuHFΔ2 and rHuHFΔ3 after protein purification at pH 7.0. Assuming the charge of rHuHF subunit is 0, then the charge of rHuHFΔ2 subunit is +2 (45-DD-46 were deleted) and +48 for 24-mer nanocage, rHuHFΔ3 subunit is +1 (44-RDD-46 were deleted) and +24 for 24-mer nanocage. The difference of net charge makes rHuHF have the fastest electrophoresis rate, followed by rHuHFΔ3.

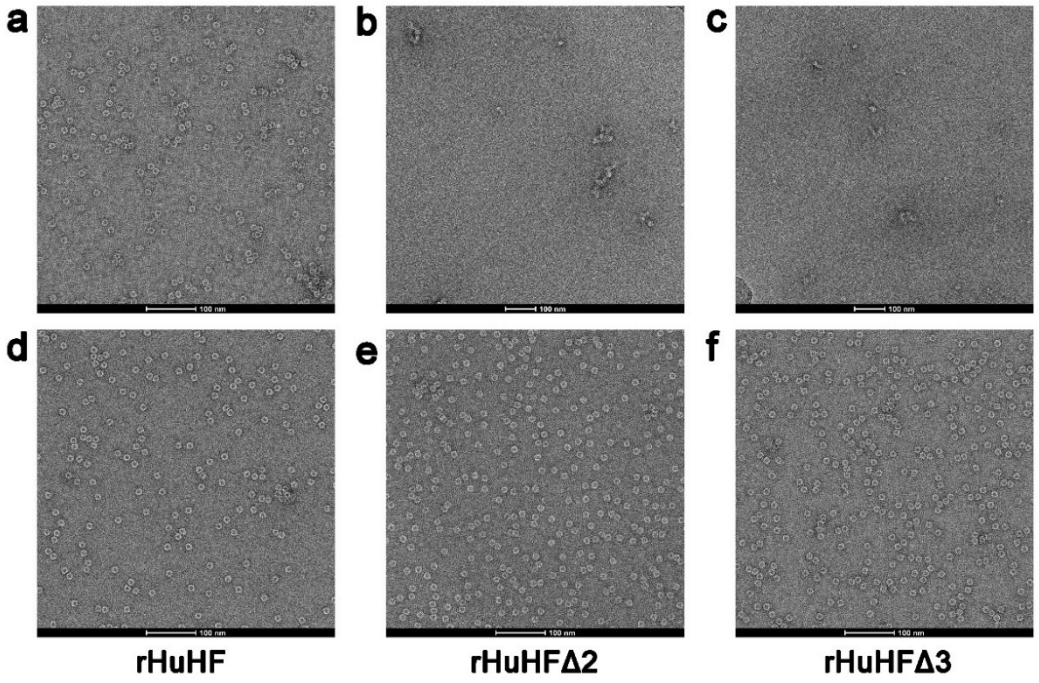


**Fig. S6** TEM images of (a) rHuHF at pH 3.0, (b) rHuHFΔ2 at pH 3.0, (c) rHuHFΔ3 at pH 4.0, (d) rHuHF at pH 7.0, (e) rHuHFΔ2 at pH 7.0, and (f) rHuHFΔ3 at pH 7.0.

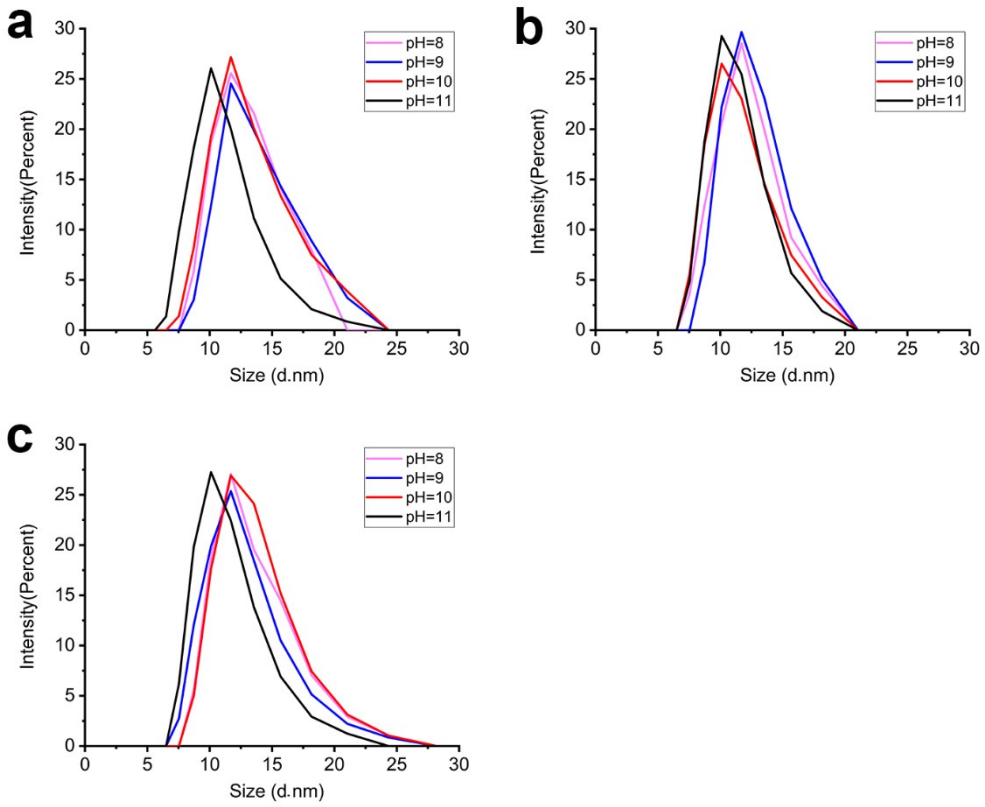


**Fig. S7** DLS spectra of (a) rHuHF, (b) rHuHFΔ2 and (c) rHuHFΔ3 under various alkaline pH conditions.

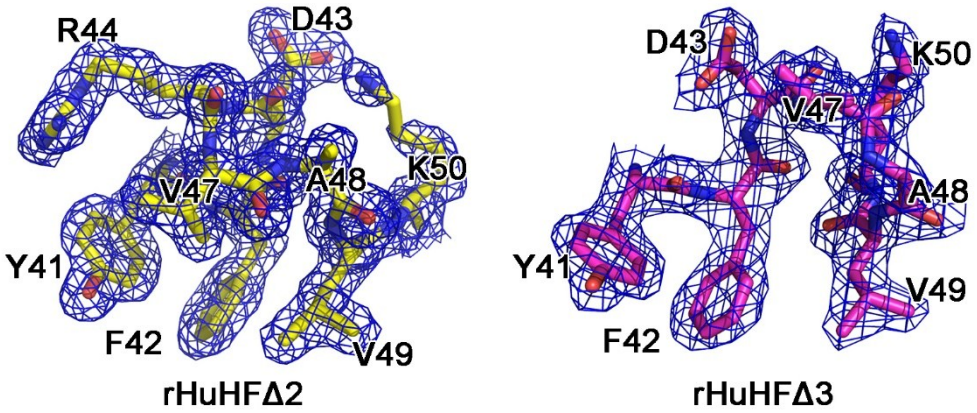


**Fig. S8** Electron density map of engineered AB loops of rHuHFΔ2 and rHuHFΔ3 is shown by 2Fo-Fc map (blue) contoured at 1 sigma (σ). Amino acid residues in AB loops are well built and labeled, the sequence number are consistent with rHuHF.

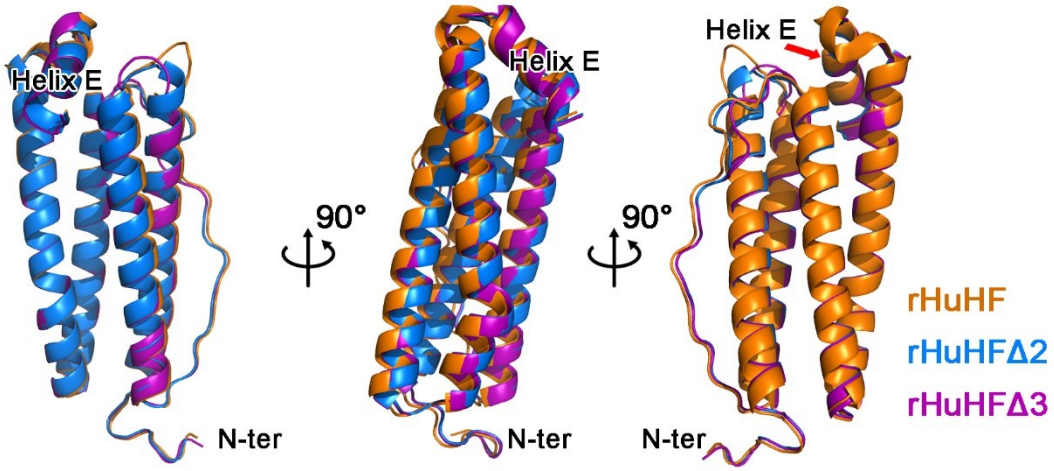


**Fig. S9** Structures alignment of rHuHF (PDB code: 2FHA), rHuHFΔ2 and rHuHFΔ3. Structures are in cartoon representation and colored in orange, marine and purple, respectively.

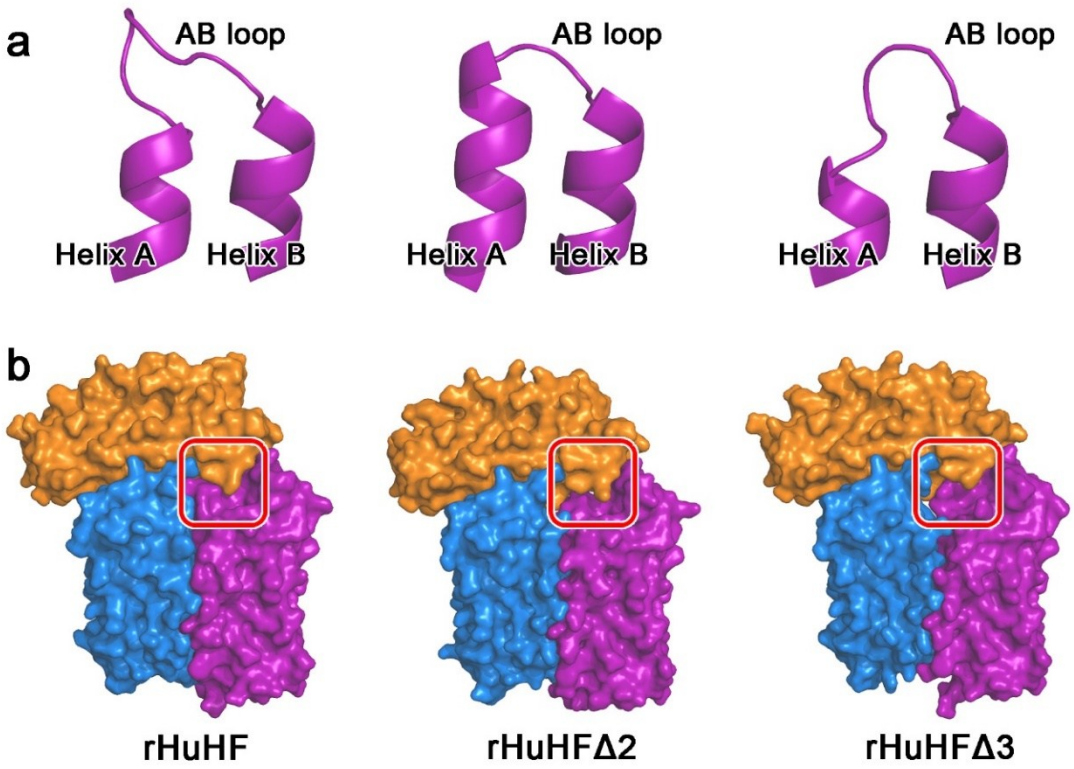


**Fig. S10** Structural comparison of reengineered AB loops from rHuHF, rHuHFΔ2 and rHuHFΔ3. (a) Comparison of the structure of AB loops. All structures are shown as cartoon diagrams. Helix A and Helix B are labeled. (b) Comparison of the $C_3$-$C_4$ interface of rHuHF, rHuHFΔ2 and rHuHFΔ3 after AB loops was reengineered. Three subunits colored orange, marine and purple, are shown in a surface diagram. Red boxes indicate the position of AB loops.

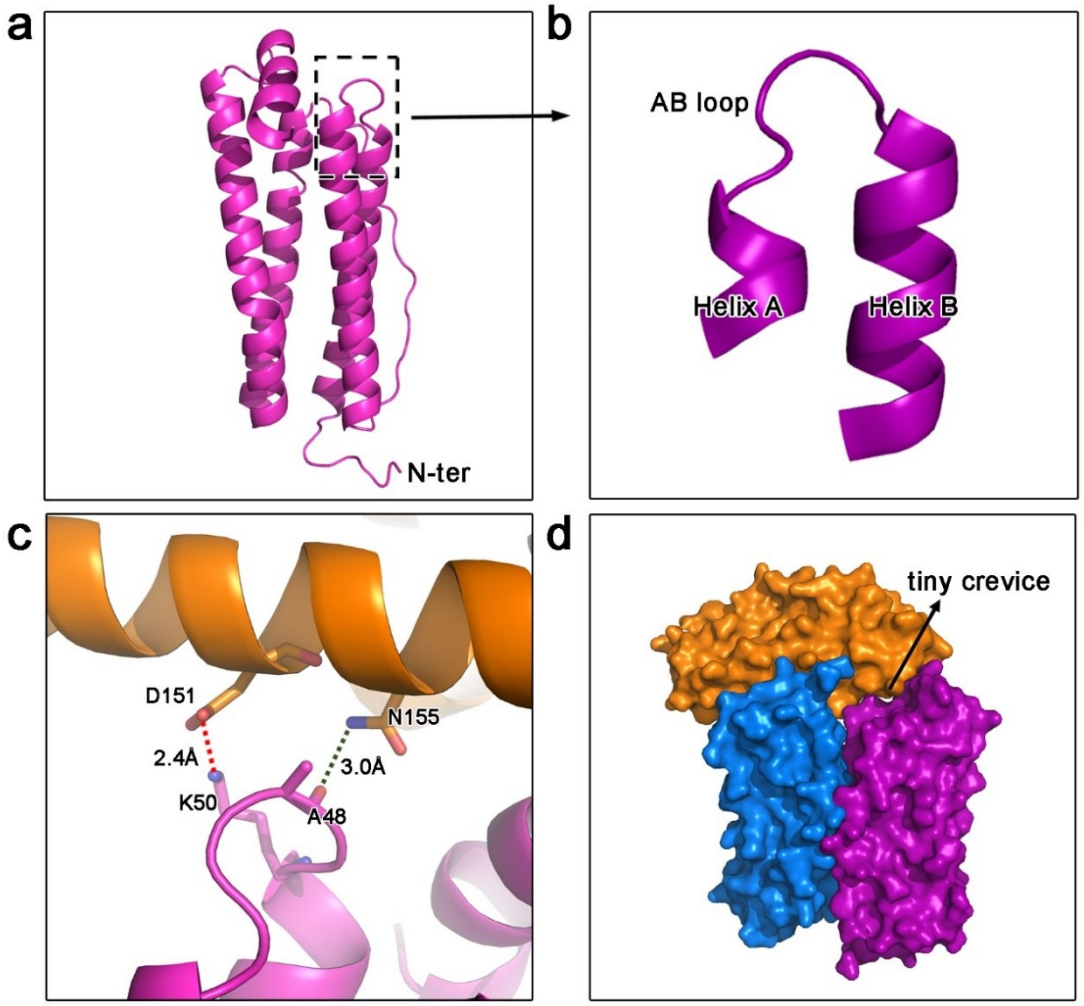


**Fig. S11** Crystal structure of rHuHFΔ4. (a) Overall structure of rHuHFΔ4 is shown as cartoon diagrams colored with purple with N-ter was labeled. (b) The close view of AB loop conformation from rHuHFΔ4. (c) The interaction between AB loop and Helix D. Amino acid residues involved are labeled and shown in sticks. Dotted red lines and green lines represent salt bridges and hydrophobic contacts with distances labeled. (d) A tiny crevice is appeared on the surface of rHuHFΔ4. Three protein subunits are shown in surface.

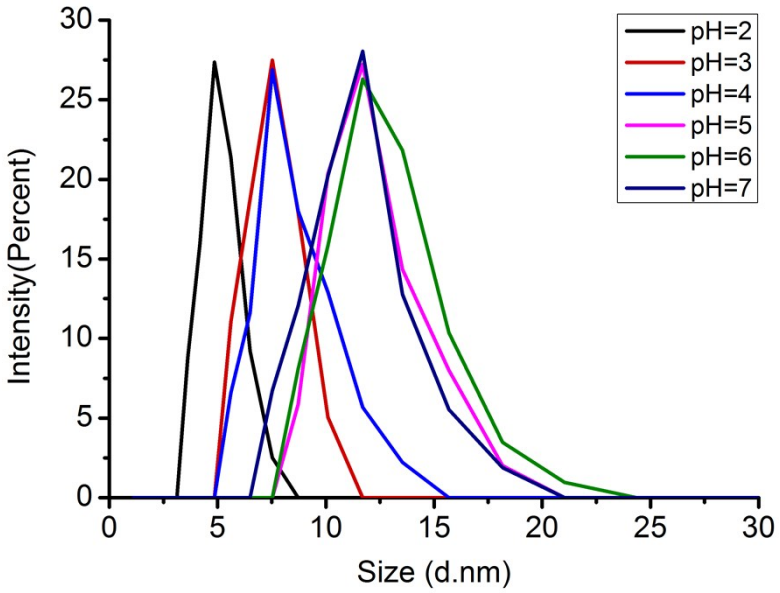


**Fig. S12** DLS spectra of rHuHFΔ4 under various pH conditions.

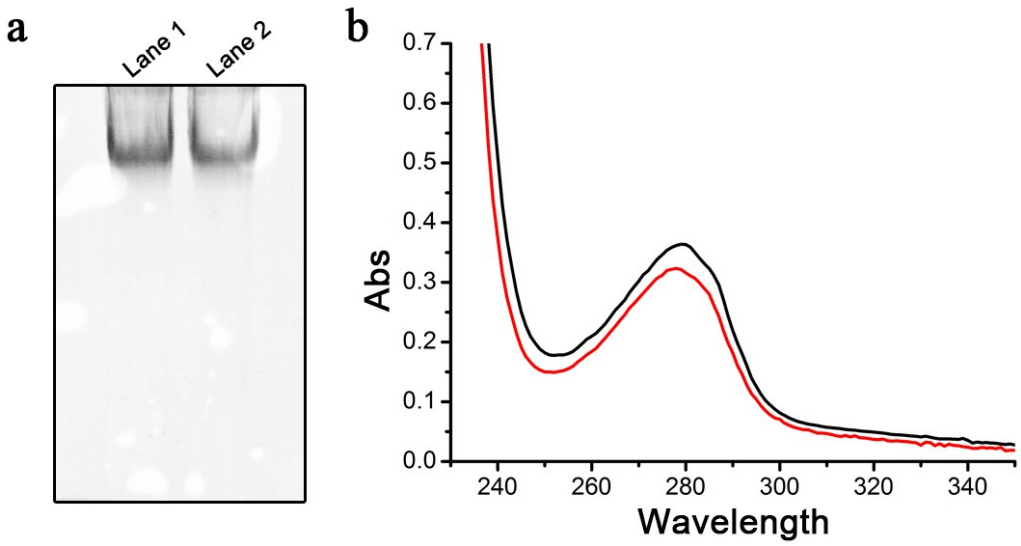


**Fig. S13** Analyses of the reassembly properties of reengineered rHuHFΔ3. (a) BN-PAGE analyses of the reassembly properties of rHuHFΔ3 at pH 4.0. Lane 1, rHuHFΔ3 at pH 7.0. Lane 2, reassembled rHuHFΔ3 with adjusting pH 4.0 back to 7.0. (b) The recovery rate of rHuHFΔ3 is calculated by the UV (280 nm) absorption. Red line: reassembled rHuHFΔ3. Black line: undissociated rHuHFΔ3.

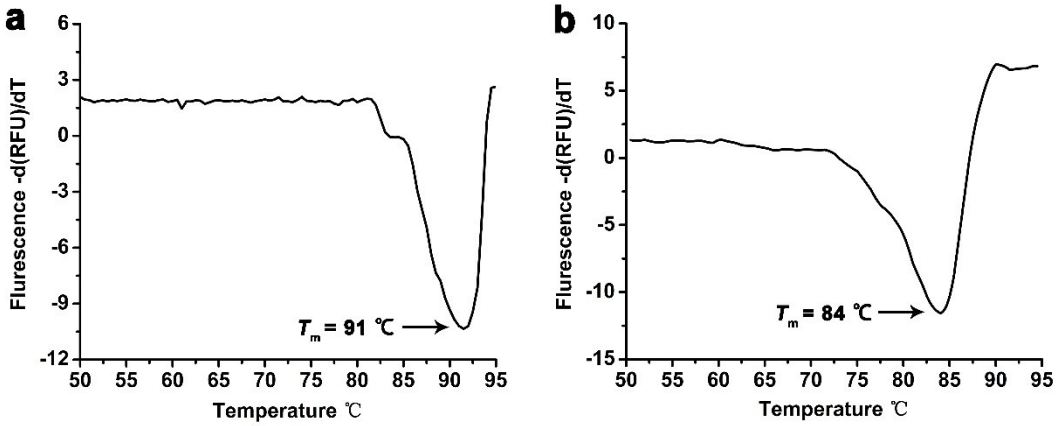


**Fig. S14** Thermal denaturation curves of (a) rHuHF and (b) rHuHFΔ3 in the temperature range from 25 ˚C to 95 ˚C measured by protein Thermal Shift.

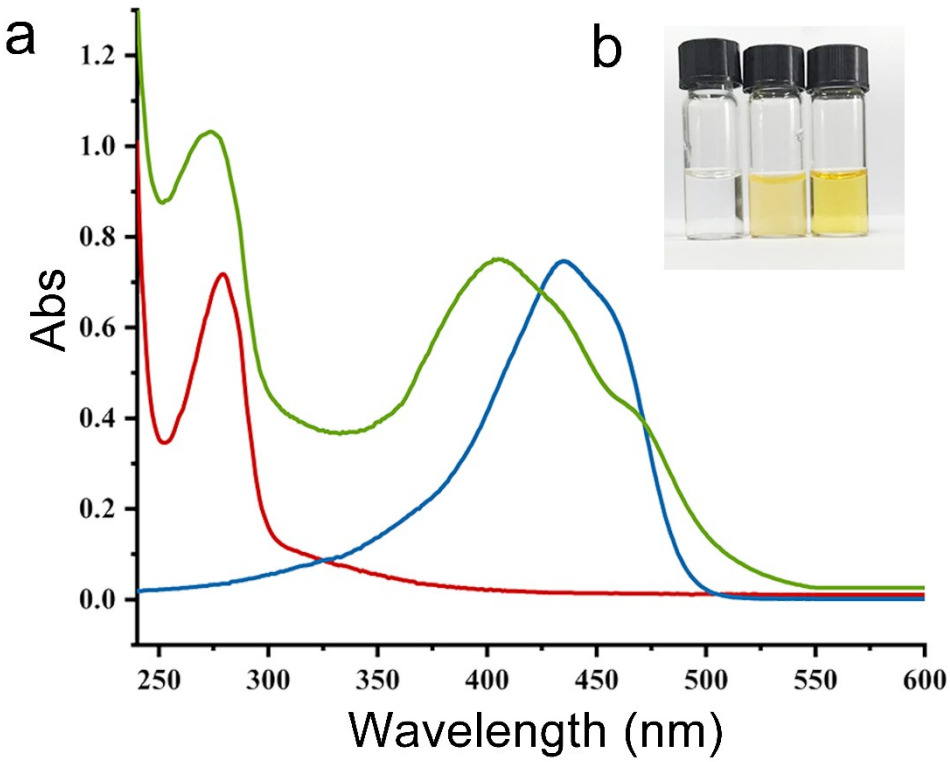


**Fig. S15** (a) UV/Vis spectra of rHuHFΔ3 (red line), curcumin (blue line), curcumin-containing rHuHFΔ3 (green line). (b) Photographs of rHuHFΔ3 (left), curcumin (middle) and curcumin-containing rHuHFΔ3 (right).

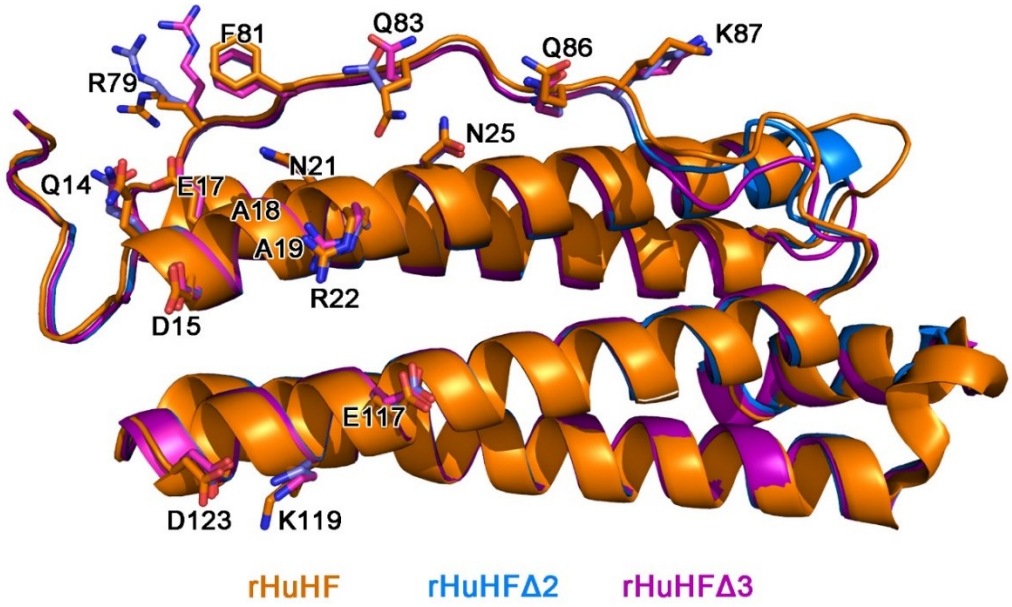


**Fig. S16** Alignment of amino acid residues from rHuHF participated in the interaction with human ferritin–transferrin receptor 1 (CD71 or TfR1). rHuHF, rHuHFΔ2 and rHuHFΔ3 are colored in orange, marine and purple. Amino acid residues are shown as sticks diagrams, respectively.

Table S1. Crystallographic properties and data collection and model refinement statistics for rHuHFΔ2, rHuHFΔ3, rHuHFΔ4.

| **Parameters** | **rHuHFΔ2** | **rHuHFΔ3** | **rHuHFΔ4** |
|---|---|---|---|
| **PDB code** | **6KE2** | **6KE4** | |
| a (Å) | 182.448 | 182.448 | 182.991 |
| b (Å) | 182.448 | 182.448 | 182.991 |
| c (Å) | 182.448 | 182.448 | 182.991 |
| α, β, γ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Space group | *F432* | *F432* | *F432* |
| Wavelength used(Å) | 0.9779 | 0.9779 | 0.9779 |
| Resolution(Å) | 50.00-1.80 (1.83-1.80) | 50.00-2.30 (2.38-2.30) | 50.00-2.10 (2.14-2.10) |
| No. of all reflections | 1126571 | 313697 | 1643163 |
| No. of unique reflections | 24788 | 12159 | 15873 |
| Completeness (%) | 99.90(100) | 99.90(100) | 99.84(100) |
| Average I/σ (I) | 53.8(13.8) | 13.6 (2.3) | 47.5(11.5) |
| $R_{merge}$[a] (%) | 7.3(20.0) | 8.0(42.4) | 7.7(29.1) |
| No. of reflections used (σ(F) > 0) | 24762(2440) | 11756(1063) | 15873(1533) |
| $R_{work}$[b] (%) | 15.44(15.68) | 19.50(22.02) | 17.59 (17.94) |
| $R_{free}$[b] (%) | 17.95(19.81) | 23.64(29.37) | 20.75 (23.61) |
| r.m.s.d. bond distance(Å) | 0.006 | 0.008 | 0.002 |
| r.m.s.d. bond angle(º) | 0.776 | 0.700 | 0.470 |
| Average B-factor(Å$^2$) | 12.18 | 20.46 | 17.79 |
| No. of protein atoms | 1434 | 1398 | 1365 |
| No. of solvent atoms | 260 | 169 | 146 |
| Rotamer outliers (%) | 0 | 0 | 0 |
| Ramachandran plot | | | |
| Res. in favored regions (%) | 98.20 | 98.80 | 98.77 |
| Res. in generously allowed region (%) | 1.80 | 1.20 | 0.62 |
| Res. in disallowed region (%) | 0 | 0 | 0 |

[a] $R_{merge} = \Sigma_h\Sigma_l \mid I_{ih} - \langle I_h \rangle \mid / \Sigma_h\Sigma_l \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of the observations $I_{ih}$ of reflection h.

[b] $R_{work} = \Sigma( \mid\mid F_p(obs)\mid - \mid F_p(calc)\mid\mid) / \Sigma\mid F_p(obs)\mid$; $R_{free}$ is an R factor for a pre-selected subset (5%) of reflections that was not included in refinement.

[c] Numbers in parentheses are corresponding values for the highest resolution shell.

## References


1 H. Schagger, *Methods in cell biology,* 2001, **65,** 231-244.

2 W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, *Acta Crystallogr. Sect. D: Biol. Crystallogr,* 2010, **62,** 859-866.

3 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M.D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr,* 2007, **40,** 658-674.

4 P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallogr., Sect. D: Biol. Crystallogr,* 2002, **58,** 1948-1954.

5 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr,* 2004, **60,** 2126-2132.

6 V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr., Sect. D: Biol. Crystallogr,* 2010, **66,** 12-21.

7 W. L. DeLano, *DeLano Scientific; San Carlos, CA, USA,* 2002.

8 L. Chen, G. Bai, R. Yang, J. Zang, T. Zhou, and G. Zhao, Food Chem, 2014, **149,** 307-312.

9 L. Chen, G. Bai, S. Yang, R. Yang, G Zhao, C. Xu, and W Leung, Food Research International, 2014, **62,** 1147-1153.