Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

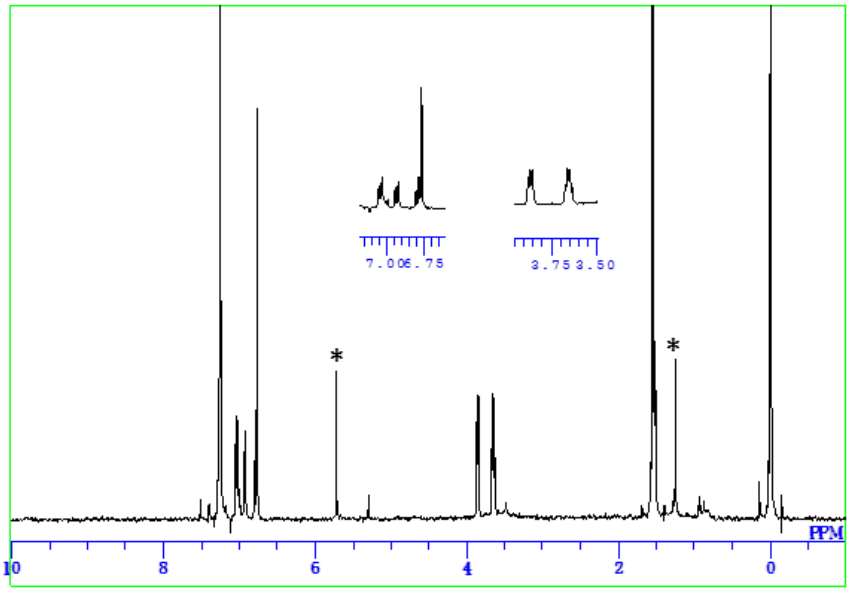
Construction of an enterobactin analogue with symmetrically arranged monomer subunits of ferritin

Hiroshi Nakajima,[a] Mio Kondo,[b] Taiki Nakane,[c] Satoshi Abe,[c] Takahiro Nakao,[b] Yoshihito Watanabe[d] and Takafumi Ueno[c]

[a]*Division of Molecular Materials Science, Graduate School of Science, Osaka City University, 3-3-138 Sugimoto, Osaka, 558-8585, Japan. E-mail: nakajima@sci.osaka-cu.ac.jp*
[b]*Department of Chemistry, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8602, Japan.*
[c]*Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259-B55, Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan. E-mail: tueno@bio.titech.ac.jp*
[d]*Research Centre of Materials Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan. E-mail: yoshi@nucc.cc.nagoya-u.ac.jp*

All chemicals were purchased from Wako Pure Chemicals, Nacalai tesque, and Sigma Aldrich and were used without further purification. The reproducibility of all data shown in this manuscript was confirmed by repeating independent experiments at least four times.

Preparation of ((N-(2-(2,5-dioxo-1H-pyrrol-1-yl)ethyl)-2,3-dihydroxybenzamide (Cat). 2,3-dihydroxybenzoicacid (34.5 mg, 0.224 mmol) was dissolved in dry THF (5 mL) under Ar atmosphere. To this solution, 1-Hydroxy-benzotriazole[1] (68.7 mg, 0.508 mmol), N-ethyl-N-isopropylpropan-2-amine (0.76 g/L, 0.085 ml, 0.500 mmol), and *N*-(2-aminoethyl)maleimide hydrochloride (FW: 176.60, 50.4 mg, 0.285 mmol) were added and stirred for more than 6hr at 25℃. The solvent was then evaporated, and the residue was dissolved in 1M HCl (5 mL). The solution was extracted with 5 mL of ethylacetate for three times. The ethylacetate solution was washed with 5 mL of saturated NH_4Cl solution for three times and dried with Na_2SO_4. After removal of ethylacetate, the residual solid was applied to a silica gel column and eluted with a mixed solution of CH_2Cl_2:MeOH = 20:1) to afford a pale yellow solid (Yield 22.6mg, 37%). ^{1}H-NMR ($CDCl_3$): δ 6.77 (s, 2H, maleimide), 6.79-7.03 (m, 3H, ArH), 3,86 (t, 2H, CH_2-N(maleimide)), 3.65 (t, 2H, CH_2NH). MS FAB: m/z ($M-H^+$): calc. 275, obsd. 275.



^{1}H-NMR spectrum of Cat. Signals marked with an asterisk show some impurity.

Modification of the A119C/C126A variant with Cat.

The A119C/C126A variant was dissolved in 20 mM MES-NaOH buffer (pH 7.0) containing 10%(v/v) DMSO, 2 mM Tris(2-carboxyethyl)phosphine (TCEP), and 150 mM NaCl. Concentration of the protein variant was adjusted to 10 μM by monitoring the absorbance at 280 nm (ε_{280} = 4.6×10^5 M^{-1}cm^{-1}). To 4 mL of the variant solution, 400 μL of 10 mM DMSO solution of Cat was added. The reaction mixture was stirred for 30 min at 25°C, followed by removal of the unreacted Cat and exchange of the buffer with 20 mM MES-NaOH buffer (pH 7.0) through a desalting column, PD-10 (GE Healthcare). The elution containing the modified protein variant was concentrated by ultrafiltration if needed. Characterization of the product was done by MALDI-TOF Mass spectrometry, CD spectroscopy, and native PAGE (Fig. S1 and S2).

Determination of a stability constant for the trisCat ligands with Fe^{3+}.

From the result of EDTA titration to Fe-Cat-Fn (Fig. S3), an equilibrium constant (K) was determined to be 4.0×10^{-2}, where

$$K = \frac{[trisCat][Fe - EDTA^-]}{[Fe\text{ - }trisCat^{3-}][EDTA]} \quad \textbf{(1)}$$

$$[EDTA] = [EDTA^{4-}] + [EDTAH^{3-}] + [EDTAH_2{}^{2-}] + [EDTAH_3{}^-] + [EDTAH_4]$$

$$[trisCat] = [trisCat^{6-}] + [trisCatH^{5-}] + [trisCatH_2{}^{4-}] + [trisCatH_3{}^{3-}] + [trisCatH$$
$$[trisCatH_5{}^-] + [trisCatH_6]$$

We define four equilibrium constants as below,

$$K_{E_1} = \frac{[EDTAH_3{}^-][H^+]}{[EDTAH_4]} \qquad K_{E_2} = \frac{[EDTAH_2{}^{2-}][H^+]}{[EDTAH_3{}^-]} \qquad K_{E_3} = \frac{[EDTAH^{3-}][H^+]}{[EDTAH_2{}^{2-}]} \qquad K_{E$$

Then,

$$[EDTA] = \alpha_4[EDTA^{4-}], \quad \alpha_4 = \left(1 + \frac{[H^+]}{K_{E_4}} + \frac{[H^+]^2}{K_{E_4}K_{E_3}} + \frac{[H^+]^3}{K_{E_4}K_{E_3}K_{E_2}} + \frac{[H^+]^4}{K_{E_4}K_{E_3}K_{E_2}K_{E_1}}\right) \quad (2)$$

K_{E1} to K_{E4} are known to be 1.0×10^{-2}, 2.1×10^{-3}, 6.9×10^{-7}, 5.5×10^{-11} /M, respectively.

Likewise,

$$K_{C_1} = \frac{[trisCatH_5{}^-][H^+]}{[trisCatH_6]} \qquad K_{C_2} = \frac{[trisCatH_4{}^{2-}][H^+]}{[trisCatH_5{}^-]} \qquad K_{C_3} = \frac{[trisCatH_3{}^{3-}][H^+]}{[trisCatH_4{}^{2-}]}$$

$$K_{C_4} = \frac{[trisCatH_2{}^{4-}][H^+]}{[trisCatH_3{}^{3-}]} \qquad K_{C_5} = \frac{[trisCatH^{5-}][H^+]}{[trisCatH_2{}^{4-}]} \qquad K_{C_6} = \frac{[trisCat^{6-}][H^+]}{[trisCatH^{5-}]}$$

Then,

$$[trisCat] = [trisCat^{6-}](1 + \frac{[H^+]}{K_{C_6}} + \frac{[H^+]^2}{K_{C_6}K_{C_5}} + \frac{[H^+]^3}{K_{C_6}K_{C_5}K_{C_4}} + \frac{[H^+]^4}{K_{C_6}K_{C_5}K_{C_4}K_{C_3}} + \frac{[H^+]^5}{K_{C_6}K_{C_5}K_{C_4}K_{C_3}K_{C_2}} + \frac{[H^+]^6}{K_{C_6}K_{C_5}K_{C_4}K_{C_3}K_{C_2}K_{C_1}})$$

The average proton dissociation constant for the *o*-hydroxyl group of catecholamide moiety in enterobactin (K_{Cav1}) has been determined to be 4.0×10^{-8}M.[2] We apply this value to K_{C1} to K_{C3}. Since that of *m*-hydroxyl group (K_{Cav2}) has not been reported, we adopt a proton dissociation constant of the *m*-hydroxyl group in dihydroxy bezamide (7.9×10^{-13}M)[3] for K_{C4} to K_{C6}. Then, the equation above is simplified as

$$[trisCat] = \alpha_6[trisCatH_3{}^{6-}]$$

$$\alpha_6 = (1 + \frac{[H^+]}{(K_{Cav_2})} + \frac{[H^+]^2}{(K_{Cav_2})^2} + \frac{[H^+]^3}{(K_{Cav_2})^3} + \frac{[H^+]^4}{(K_{Cav_2})^3(K_{Cav_1})} + \frac{[H^+]^5}{(K_{Cav_2})^3(K_{Cav_1})^2} + \frac{[H^+]^6}{(K_{Cav_2})^3(K_C$$

) **(3)**

$K_{Cav1} = 4.0\times10^{-8}$M, $K_{Cav2} = 7.9\times10^{-13}$M

With eqs. 2 and 3, eq. 1 is transposed to $$K = \frac{[trisCat^{6-}][EDTA\text{-}Fe^-]\alpha_6}{[Fe\text{-}trisCat^{3-}][EDTA^{4-}]\alpha_4} \quad \textbf{(4)}$$

The absolute complex stability constant of EDTA with Fe^{3+} (K_{SE}) is

$$K_{SE} = \frac{[EDTA\text{-}Fe^-]}{[EDTA^{4-}][Fe^{3+}]} \quad \textbf{(5)}$$

Likewise, the absolute complex stability constant of trisCat with Fe^{3+} (K_{SC}) is

$$K_{SC} = \frac{[Fe - trisCat^{3-}]}{[trisCat^{6-}][Fe^{3+}]} \quad \textbf{(6)}$$

From eqs. 4,5, and 6, $$K_{SC} = \frac{\alpha_6 K_{SE}}{\alpha_4 K} \quad \textbf{(7)}$$

Appling the known K_{SE} ($1.3\times10^{25}M^{-1}$) [4] and obtained K (4.0×10^{-2}) at pH 7 (i.e. $[H]^+ = 10^{-7}M$), K_{SC} is calculated to be $1.2\times10^{40}M^{-1}$ by eq. 7.

Crystallization of FeCat-Fn and CdCat-Fn.

Crystallization of FeCat-Fn and CdCat-Fn were performed with a hanging drop vapor diffusion method as described in previous reports.[5] A solution of FeCat-Fn (or CdCat-Fn) was concentrated to approximately 19 mg mL^{-1} and the drops were prepared by mixing an equal volume (3 μL) of a protein solution (20 mM Tris/HCl pH8.0, 0.15 M NaCl aq.) and the precipitant solution (1 M $(NH_4)_2SO_4$, 20 mM $CdSO_4$), and equilibrated against the precipitant solution (1 mL) at 20 °C. The crystals were obtained within a day.

X-ray Crystal Analysis.

Before the data collection of FeCat-Fn, single crystal was immersed in a precipitant solution containing 25% (w/w) glycerol and subsequently frozen in liquid nitrogen. X-ray diffraction data of FeCat-Fn was collected at 100K at beamline BL26B1 at SPring-8 using X-ray wavelength of 1.0 Å. In order to distinguish Fe atoms from Cd atoms, which are essential for crystallization, X-ray diffraction data of Fe-Cat-Fn was also collected using 1.73830 and 1.79022 Å which represent the peak wavelength and remote wavelength of Fe X-ray absorption, respectively. X-ray diffraction data of CdCat-Fn was collected at 100K at beamline BL38B1 at SPring-8 using X-ray wavelength of 1.0 Å. The data were processed with HKL2000 programs in the cubic F432 space group. The crystal parameters and data collection statistic are summarized in Table S1.

Refinement.

The structures were determined by molecular replacement with MOLREP using an apo-Fr structure (PDB ID: 1DAT) as the initial model. Refinement of the protein structure was performed using REFMAC5[6] in the CCP4 suite. Rebuilding was performed using COOT[7] based on sigma weighted (2Fo-Fc) and (Fo-Fc) electron density maps. About 5% of the observed data were excluded from the refinements and used to calculate the free R (Rfree) as a monitor of model bias. Water molecules were positioned to fit residual (Fo-Fc) density peaks with a lower cut-off of 3σ. The models were subjected to quality analysis during the

various refinement stages with omit maps and RAMPAGE.[8] The refinement statistics are summarized in Table S2. Atomic coordinates are deposited in the Protein Data Bank under accession numbers 5CZU, and 5AXS for Fe-Cat-Fn and Cd-Cat-Fn, respectively. Lys172 of Cat-Fn was replaced to Ala because corresponding electron density was missing

References

1 L. A. Carpino, *J. Am. Chem. Soc.*, 1993, **115**, 4397.
2 L. D. Loomis and K. N. Raymond, *Inorg. Chem.*, 1991, **30**, 906.
3 W. R. Harris, C. J. Carrano and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 2213.
4 R. M. Smith, A. E. Martell and R. J. Motekaitis, *Inorg. Chim. Acta*, 1985, **99**, 207.
5 T. Ueno, M. Abe, K. Hirata, S. Abe, M. Suzuki, N. Shimizu, M. Yamamoto, M. Takata and Y. Watanabe, *J. Am. Chem. Soc.*, 2009, **131**, 5094.
6 G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr., Sect. D* **1997**, 53, 240-255.
7 P. Emsley, K. Cowtan, *Acta Crystallogr., Sect. D* **2004**, 60, 2126-2132.
8 S. C. Lovell, I. W. Davis, W. B. Adrendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, *Proteins-Structure Function and Genetics* **2003**, *50*, 437-450.

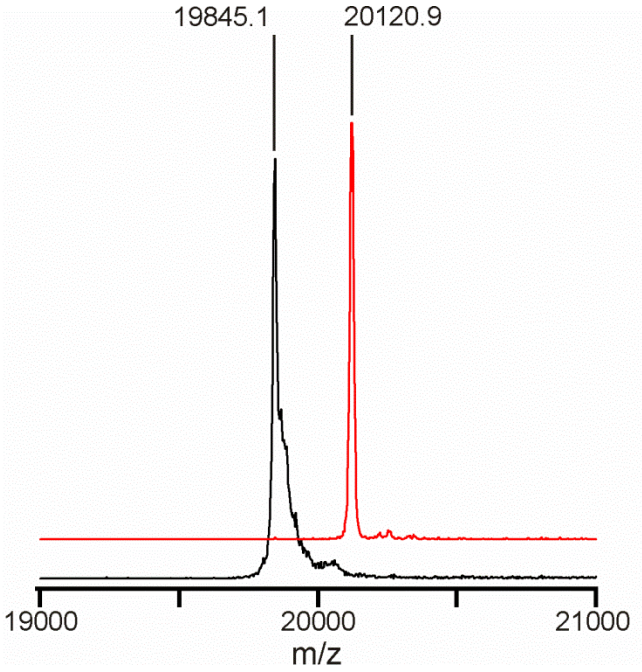


Fig. S1 MALDI-TOF mass spectra of the

A119C/C126A variant before (black line) and after (red line) the modification with Cat. The red line is offset for reader's convenience.

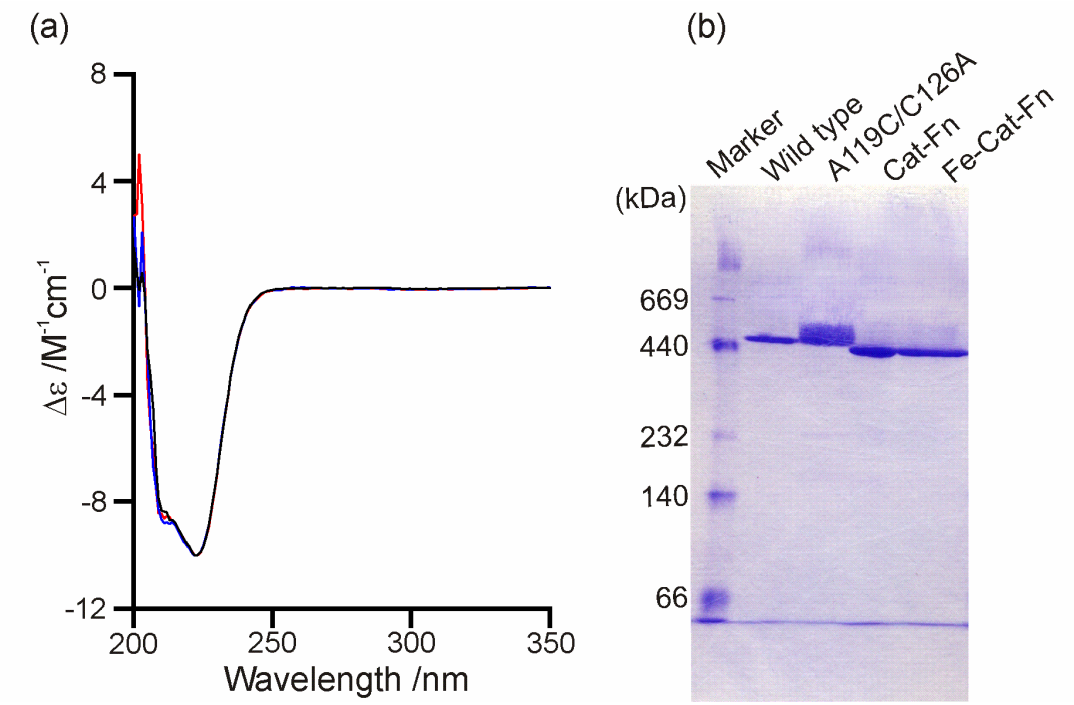


Fig. S2 (a) CD spectra of intact ferritin (black line), the A119C/C126A variant (blue line), and the variant modified with Cat (Cat-Fn, red line). (b) Results of native PAGE. All proteins show virtually same migration distance. Fe-Cat-Fn represents Cat-Fn in the ferric iron bound form. Native PAGE markers: 669kDa, Thyroglobulin; 440, Ferritin; 232, Catalase; 140, Lactate dehydrogenase; 66, Albumin

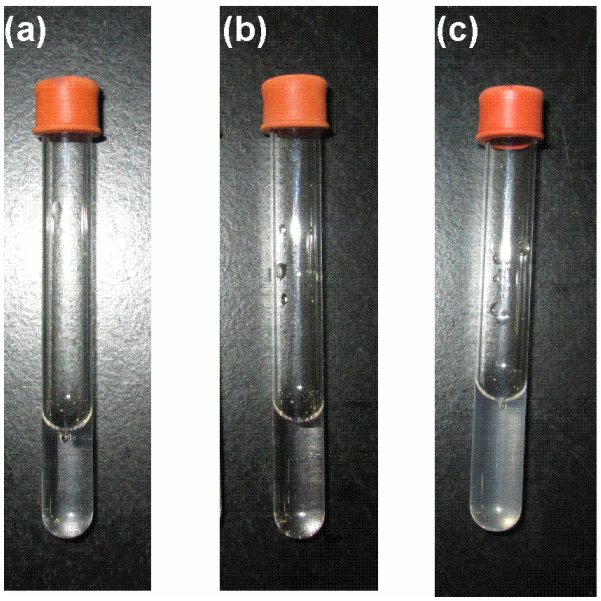


Fig. S3 The pictures of Cat-Fn solution taken during the $FeCl_3$ titration. (a) + 0eq. $FeCl_3$ (b) + 1eq. $FeCl_3$ (c) + 2eq. $FeCl_3$. The concentration of Cat-Fn was adjusted to 4.2×10^{-6} M.

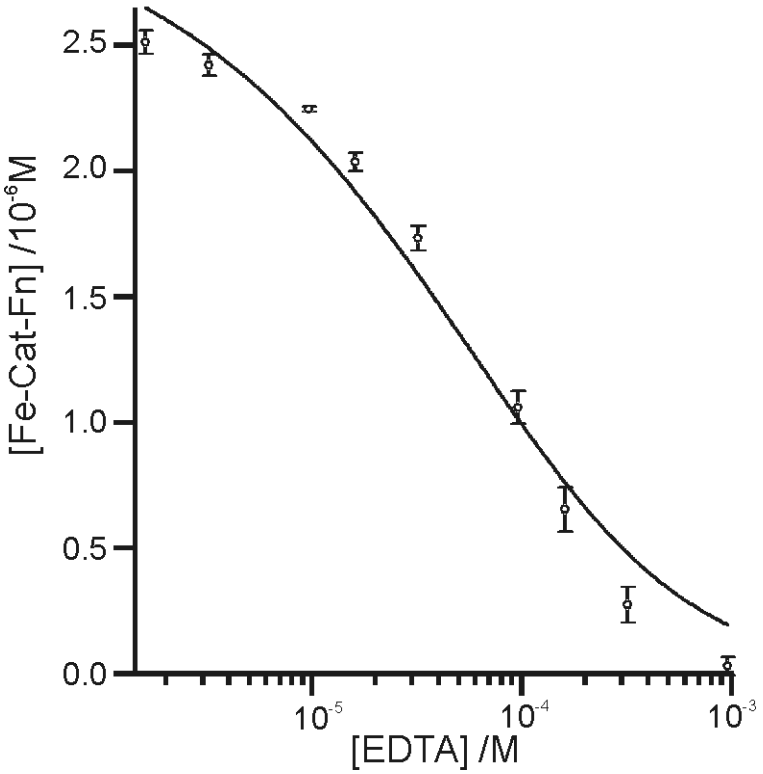


Fig. S4 Titration of EDTA to Cat-Fn in the Fe^{3+} bound form (Fe-Cat-Fn). The initial concentration of Fe-Cat-Fn was adjusted to 2.5×10^{-6} M. As there are eight C_3 axis channels in ferritin, actual concentration of the Fe-trisCat complex should be 20×10^{-6} M. Each plot is a mean value calculated from four sets of independent experiments. The plots were fitted with a regression curve obtained from the balanced equation between [Fe-Cat-Fn] and [EDTA].

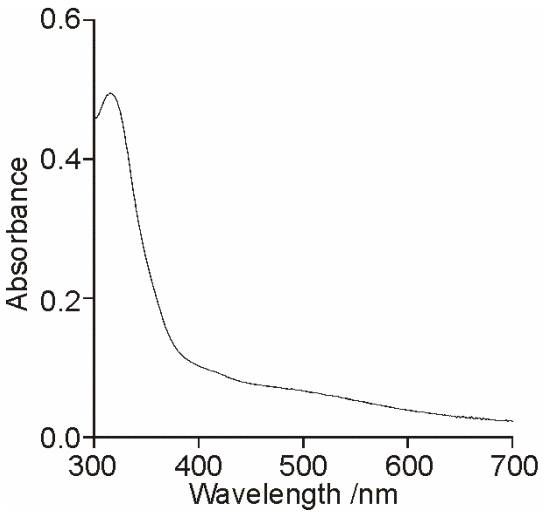


Fig. S5 UV-Vis spectrum of Fe-Cat-Fn after dialysis against Fe-free 20mM MES-NaOH buffer (pH 7.0) for 3 days at 25°C.

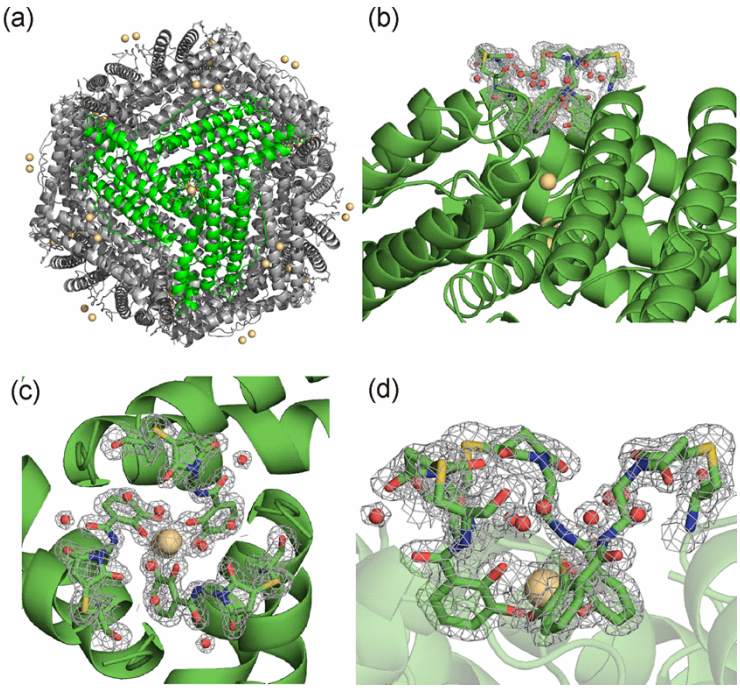


Fig. S6 X-ray Crystal structures of Cat-Fn in the Cd^{2+} ion bound form. The Cd ions are illustrated with ivory-colored spheres. (a) Overall structure. (b) and (C) Side and top views of the Cat modified C_3 axis channel. (d) Close view of Cd-trisCat.

Table S1. Crystallographic data

Data collection	Fe-Cat-Fn			Cd-Cat-Fn
	1 Å	**Fe Peak**	**Fe Remote**	**1 Å**
X-ray wavelength, Å	1	1.73830	1.79022	1
Space group	*F432*	*F432*	*F432*	*F432*
Crystal cell (Å)				
$a = b = c$	181.070	181.171	181.176	180.878
Resolution range (Å)	60.0-1.60 (1.63-1.60)	60.0-1.97 (2.00-1.97)	60.0-2.01 (2.04-2.01)	40.0-1.67 (1.70-1.67)
Observations	666,123	171,719	158,631	616,177
Unique reflections	63,276	34,046	32.025	55,550
Completeness (%)	99.9 (100)	100 (99.6)	99.9 (98.4)	100 (100)
I/σ	58.0 (6.1)	29.9 (2.6)	28.5 (2.4)	53.4 (11.3)
R_{merge}	4.8 (29.7)	5.6 (30.5)	5.9 (28.5)	7.1 (29.1)
Redundancy	10.5 (5.4)	5.1 (2.1)	5.0 (1.9)	11.1 (11.2)

Values in parentheses are for the highest-resolution shell.

Table S2. Refinement statistics

	Fe-Cat-Fn	**Cd-Cat-Fn**
Resolution range (Å)	36.96 – 1.60	34.81 – 1.67
Reflection used	32282	28477
R-factor (%)	0.173	0.162
Free *R*-factor (%)	0.208	0.188
R.m.s. deviations from		
Bond length (Å)	0.0268	0.0274
Angle (°)	2.9215	2.6223
Ramachandran plot		
most favored	97.6	98.8
allowed	1.8	1.2
outlier	0.6	0.0