Supplementary Information

Structural basis for heme transfer reaction in heme uptake machinery from Corynebacteria

Norifumi Muraki,[a,b,c] Chihiro Kitatsuji,[a] Yasunori Okamoto,[a,d] Takeshi Uchida,[e,f] Koichiro Ishimori [e,f] and Shigetoshi Aono*[a,b,c]

[a]*Department of Creative Research, Exploratory Research Center on Life and Living Systems (ExCELLS), National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji-cho, Okazaki 444-8787, Japan. E-mail: aono@ims.ac.jp*
[b]*Institute for Molecular Science, National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji-cho, Okazaki 444-8787, Japan*
[c]*Department of Structural Molecular Science, The Graduate University for Advanced Studies, 38 Nishogo-naka, Myodaiji-cho, Okazaki 444-8585, Japan*
[d]*Present address: Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3 Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan*
[e]*Graduate School of Chemical Sciences and Engineering, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan*
[f]*Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan*

Materials and methods

The *htaA* and *htaB* genes from *Corynebacterium glutamicum*, whose codons were optimized for their expression in *E. coli*, were synthesized by FASMAC. These genes were designed based on their amino acid sequences (Fig. S1) reported in NCBI Reference Sequences (NCgl0377 and NCgl0381 for HtaA and HtaB, respectively). HtaA-N, HtaA-C, and HtaB prepared in this study are composed of Gln36~Thr219, Ser364~Gly529, and Lys35~Val228, respectively. Each of these genes that was obtained by PCR was inserted at NdeI-XhoI site in pET-22b(+) to prepare the expression vectors for these proteins with a C-terminal His_6 tag. The expression vectors for the variant proteins (H111A-HtaA-N, H434A-HtaA-C, and H113A-HtaB) were prepared by QuickChange Site-directed Mutagenesis Kit (Stratagene). The introduction of the cognate mutation was confirmed by DNA sequencing.

E. coli BL21(DE3) cells transformed by the expression vector were grown in 5 ml of Luria-Bertani broth (LB) containing 50 μg/mL ampicillin and 1% (w/v) glucose for 4 h at 37 ºC. The 5 mL cultures were inoculated into 300 mL of Terrific Broth (TB) containing 50 μg/mL ampicillin, which were grown at 37 ºC with 140 rpm shaking by a rotary shaker. After 4 h of cultivation, isopropyl β-D-thiogalactopyranoside was added to a final concentration of 0.1 mM to induce protein expression, and then the cultivation was continued for overnight at 20 ºC with 80 rpm shaking. The grown cells were harvested by centrifugation and stored at -80 ºC until use.

The recombinant proteins were purified with metal affinity, anion exchange, and gel filtration column chromatography. Briefly, *E. coli* cells were resuspended in 50 mM Tris-HCl (pH 8.5) and lysed by sonication. The supernatant obtained by centrifugation was applied to a HisTrap HP column (GE Healthcare), which was pre-equilibrated in 50 mM Tris-HCl buffer (pH 8.5) containing 500 mM NaCl. The fractions containing the cognate sample were combined and applied to a HiTrap-Q column (GE Healthcare) and finally purified by a Superdex200 HiLoad 16/600 column (GE Healthcare), which was pre-equilibrated in 50 mM Tris-HCl buffer (pH 8.5) containing 200 mM NaCl.

HtaA-N, HtaA-C, and HtaB were concentrated in 50 mM Tris-HCl (pH 8.5) containing 200 mM NaCl by a centrifugal filter unit (Amicon Ultra, Merck) to 20, 47, and 20 mg/ml, respectively, for the crystallization. The crystals were obtained in 1.4 M tri-sodium citrate, 0.1 M HEPES-NaOH, 3% 2-methyl-2,4-pentanediol for HtaA-N, in 3.0 M sodium acetate for HtaA-C, and in 38% 2-ethoxyethanol, 0.1 M Tris (pH 7.5), 0.05 M calcium acetate for HtaB as precipitants. The holo and apo forms of H434A-HtaA-C were concentrated in 50 mM Tris-HCl (pH 8.0) containing 200 mM NaCl to 52 and 28 mg/ml, and were

crystallized in 4.5 M sodium formate, 0.1 M HEPES (pH 7.0), 7.0% sucrose and in 17% PEG3350, 0.1 M Tris (pH 8.0), 13 % 2-propanol, respectively.

The diffraction data were obtained at 100 K on beamline BL44XU at Spring-8. The data analyses were carried out as reported previously for the determination of crystal structures.[1,2] The statistics of data collection and refinement are summarized in Table S1.

The electronic absorption and resonance Raman spectra were measured as reported previously.[3,4]

References

1. N. Muraki and S. Aono, *Chem. Lett.*, 2016, **45**, 24.
2. N. Muraki, C. Kitatsuji, M. Ogura, T. Uchida, K. Ishimori and S. Aono, *Int. J. Mol. Sci.*, 2016, **17**, 829.
3. C. Kitatsuji, M. Ogura, T. Uchida, K. Ishimori and S. Aono, *Bull. Chem. Soc. Jpn.*, 2014, **87**, 997.
4. Y. Okamoto, H. Sawai, M. Ogura, T. Uchida, K. Ishimori, T. Hayashi and S. Aono, *Bull. Chem. Soc. Jpn.*, 2014, **87**, 1140.

Table S1. Statistics of X-ray data collection and structure refinement of HtaA-N, HtaA-C, and HtaB.

	HtaB	HtaA-N	HtaA-N FeSAD	HtaA-C	HtaA-C FeSAD	HtaA-C H434A Heme-bound	HtaA-C H434A Domain-swapped
Data collection							
Wavelength (Å)	0.90000	0.90000	1.70000	0.90000	1.70000	0.90000	0.90000
Resolution range (Å)	38.09 – 1.70 (1.76 – 1.70)	32.05 – 2.00 (2.07 – 2.00)	56.41 – 2.8 (2.9 – 2.8)	31.03 – 1.30 (1.35 – 1.30)	43.61 – 1.96 (2.03 – 1.96)	33.66 – 1.4 (1.45 – 1.4)	31.81 – 2.04 (2.12 – 2.04)
Space group	*C*2	$P3_112$	$P3_112$	$C222_1$	$C222_1$	$P2_12_12_1$	*C*2
Unit cell							
a, b, c (Å)	235.2 33.7 126.6	65.1 65.1 182.0	65.1 65.1 182.0	60.3 63.3 88.2	60.0 63.1 87.2	39.0 43.2 199.4	105.9 40.2 87.0
α, β, γ (°)	90 121.1 90	90 90 120	90 90 120	90 90 90	90 90 90	90 90 90	90 114.0 90
Total reflections	346427 (33750)	215950 (21340)	241864 (23765)	306261 (29990)	84603 (8055)	491325 (47629)	79006 (7286)
Unique reflections	94191 (9232)	30290 (3011)	11225 (1103)	41800 (4125)	12170 (1191)	66450 (6437)	21370 (2022)
Multiplicity	3.7 (3.7)	7.1 (7.1)	21.5 (21.5)	7.3 (7.3)	7.0 (6.8)	7.4 (7.4)	3.7 (3.6)
Completeness (%)	98.88 (98.10)	99.87 (100.00)	99.95 (100.00)	99.89 (99.64)	99.75 (98.84)	98.19 (96.87)	98.71 (93.86)
I/sigma(I)	11.73 (2.40)	11.01 (3.03)	22.96 (6.06)	12.14 (4.28)	17.20 (8.21)	17.73 (3.46)	15.06 (1.71)
Wilson B-factor	19.35	31.65	46.66	9.64	20.14	13.39	41.32
R_{meas}	0.086 (0.660)	0.104 (0.667)	0.114 (0.597)	0.106 (0.570)	0.077 (0.199)	0.072 (0.670)	0.062 (0.959)
R_{pim}	0.045 (0.341)	0.038 (0.248)	0.024 (0.128)	0.039 (0.209)	0.029 (0.076)	0.026 (0.243)	0.032 (0.498)
CC1/2	0.997 (0.810)	0.997 (0.792)	0.999 (0.955)	0.997 (0.908)	0.998 (0.99)	0.999 (0.884)	0.999 (0.602)
Refinement							
R_{work}	0.177 (0.229)	0.195 (0.269)		0.160 (0.218)		0.181 (0.230)	0.204 (0.332)
R_{free}	0.198 (0.245)	0.221 (0.314)		0.180 (0.249)		0.202 (0.247)	0.246 (0.370)
Number of atoms							
Proteins	3157	2664		1235		2480	2354
Ligands	174	86		43		90	4
Waters	300	53		117		171	11
r.m.s. deviations							
Bond lengths (Å)	0.017	0.006		0.005		0.005	0.007
Bond angles (°)	1.49	0.99		1.02		1.17	0.75
Ramachandran plot							
Favored (%)	97.46	95.77		98.16		96.07	97.82
Allowed (%)	2.54	3.66		1.23		3.32	2.18
Outliers (%)	0	0.56		0.61		0.6	0
Average B-factor							
Proteins	24.09	42		11.91		17.14	53.5
Ligands	21.11	28.59		11.75		19.03	55.95
Waters	30.52	34.29		21.48		24.17	41.49
PDB ID	6jsb	6js9		6jsa		6jsc	6jsd

(a)

MKLAPRMRMRSPKTFAALASLALVIGLGQVPIAQAQTEYRTASDGSLNWGFRQSFRNYI
QTGVAKGSITLGDGASDNGGNFAFTPRTNGTTVTSDSQGTVEFNGSVHFLGHQAEDKWI
LDTTMSDIKMVFNGSSAQLVVDLVAREFKGTTYDDIGEYIISDDIVLADVSLNSAADFS
QDSIDLSGTTDLTAAGAQAFGGFYETGEALDPTGGSLTISSTTTAPSTSTTSTSASTSG
GTADCSSGALGVVTTGTNDGMLGTIQEVNNTFAIWNNLIVNTERMFCNIDTLKARFDTD
DSSDSATSATSGTTASTGTTAATTAGTTGTTGTASTASGTSGTSGTSGTAATVAGTTPT
DNGVCTASGSLGVTQASAQWGVKASFQNYIRGSIANGSWTLNGVGFDNQQFQFSGNSGA
VDAENKTGSINFPGSIHFTGHGGILDMQIANIEISFNGNSGELIADVVSSDMDGNSTNY
GRTVVGTLNFSALNVSATEASGSASVSLSQSGSQAFADFYTPGTQLDPISFSATLGGDA
SCATGSTSTTGAAATANTDNTEGVAGEESTTPANQNSQFQIRQAAADSTGLDTTTTMLL
ILAAFVVAGGSMTRFTVGNPTGK

(b)

MNKLATRALVALTGSAIAMTGLTVVSANAAEKTGKCRVVTTTGTADWSVRESFNNYLEG
PIANGAAYKYHGGIEVRDGVETTGTKSAREFTWPVLGSEEGAVKLGGGVHWTGHNHYSG
DDESQAPDNFILDLDFSNPTVKFDGNEGTLLVDFKSREFVDTKTVADFLTGTQAELATI
TFDEPIDLTQENVTVTGQTKLTATGVDVMGTFYPEGEALAPITLNLTNEVVCDEPETPV
EPEVPVEPETPVDPETSVDPETPVDPETSVDPEKPGDDNKDDGSNSSSNGDILGILGIL
AALGGVGALVYNFLVASGFLAAFK

Fig. S1. Amino acid sequences of (a) HtaA and (b) HtaB from *Corynebacterium glutamicum*. The putative signal peptide and the membrane spanning regions are shown in red and blue, respectively. HtaA-N (residues 36-219), HtaA-C (residues 364-529), and HtaB (residues 35-228) used in this study consist of the amino acids shown with the solid, the dotted, and the wave underlines, respectively.

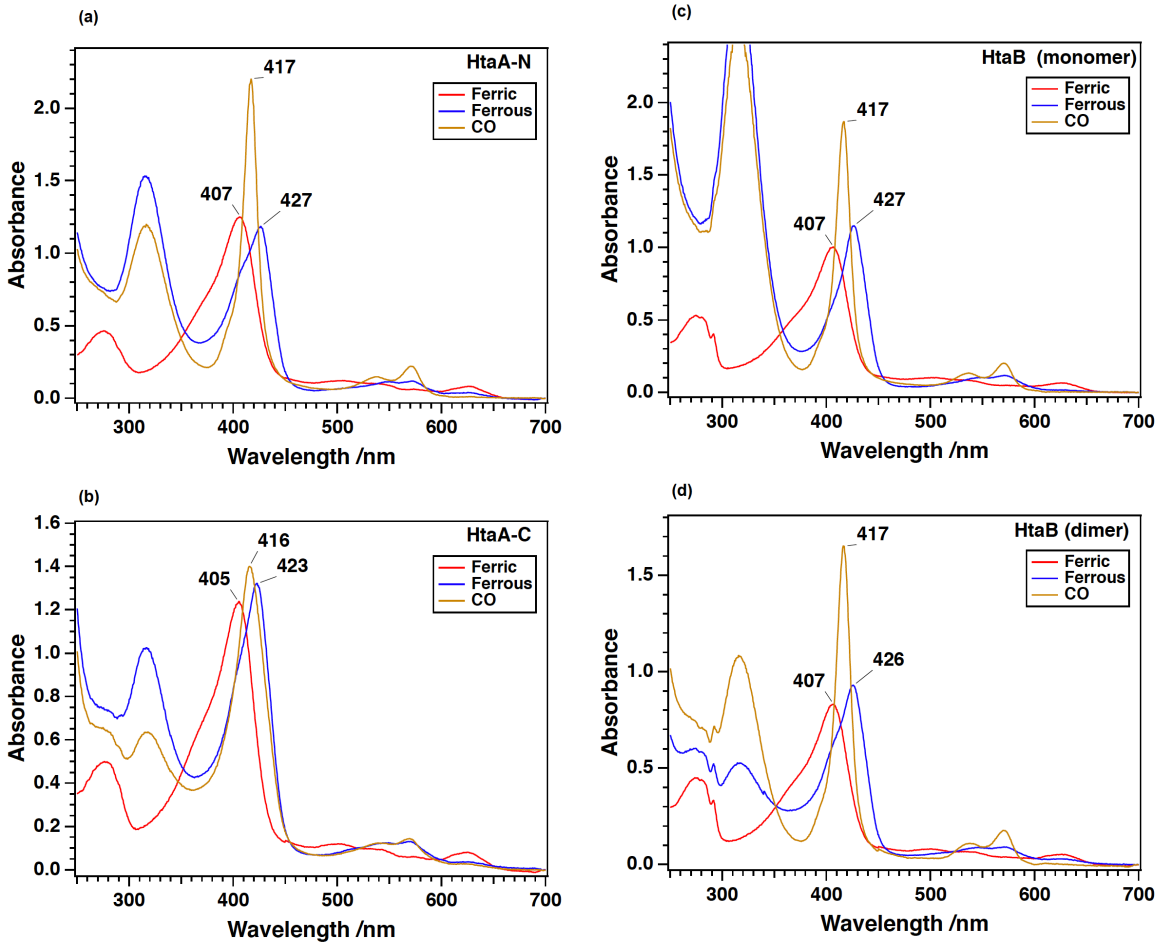


Fig. S2. Electronic absorption spectra for the holo form of (A) HtaA-N, (B) HtaA-C, (C) HtaB in monomer, and (D) HtaB in dimer. The ferric, ferrous, and ferrous CO-bound forms are shown in red, blue, and brown, respectively.

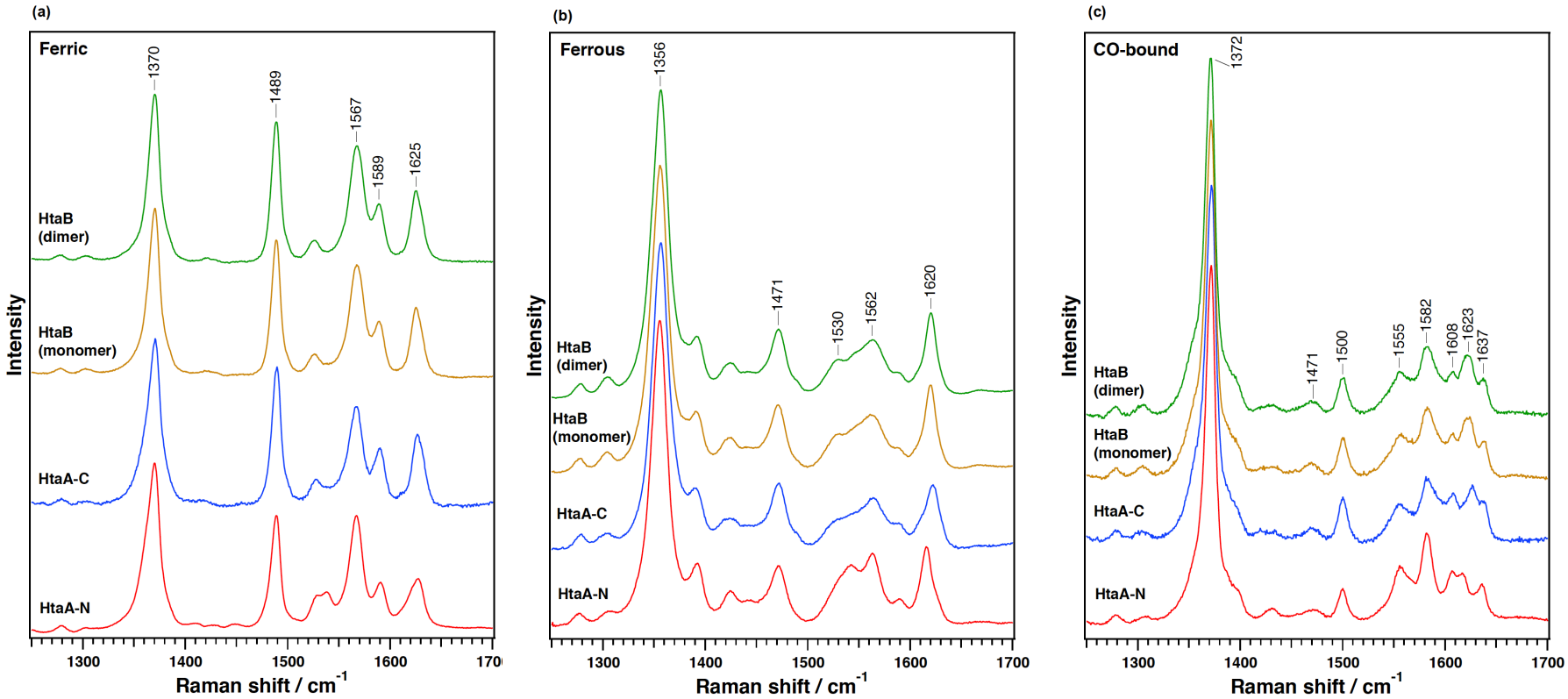


Fig. S3. Resonance Raman spectra of HtaA-N (red), HtaA-C (blue), HtaB in monomer (brown), and HtaB in dimer (green) in the ferric (a), ferrous (b), and ferrous CO-bound (c) forms. These spectra were obtained by excitation with 413.1 nm light from a krypton ion laser .

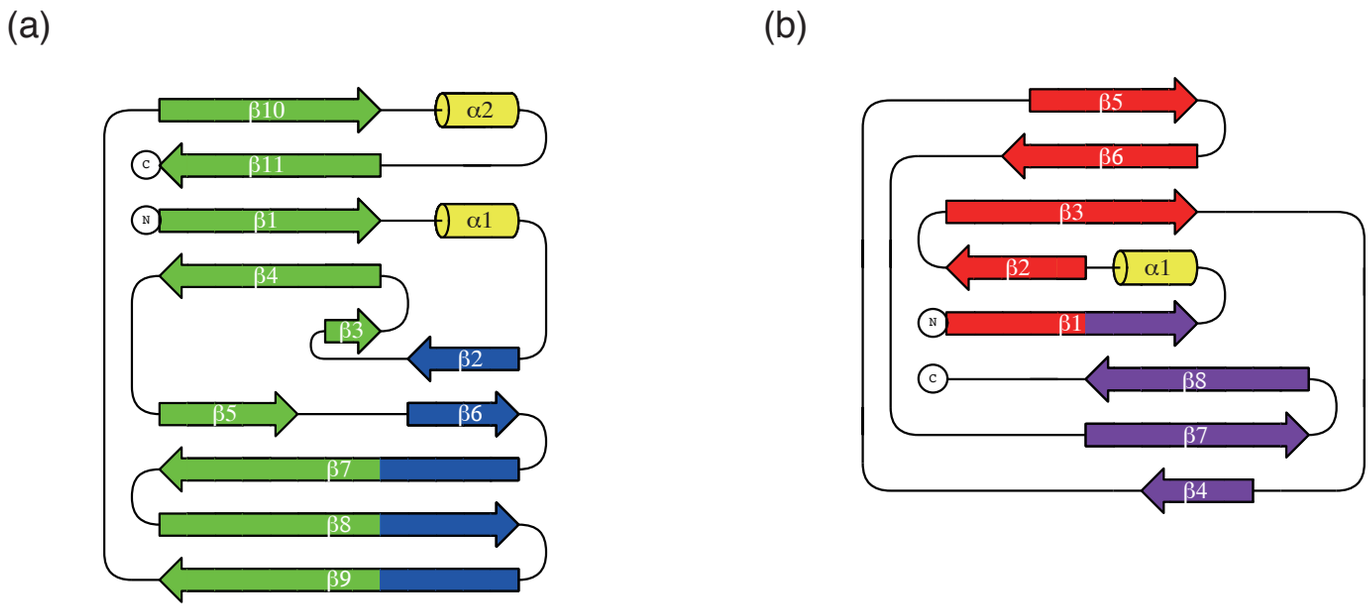


Fig. S4. Topology diagram of the protein fold of (a) the CR domains (HtaA-N, HtaA-C, and HtaB) and (b) the NEAT domain (IsdC, PDB code: 2O6P). “N” and “C” show the N- and C-terminus, respectively. The α helices and the β-strands are represented as cylinders and arrows, respectively. The regions composing the β-barrel in the CR domains are shown in green. The β strands shown in blue, red, and purple form individual β-sheets.

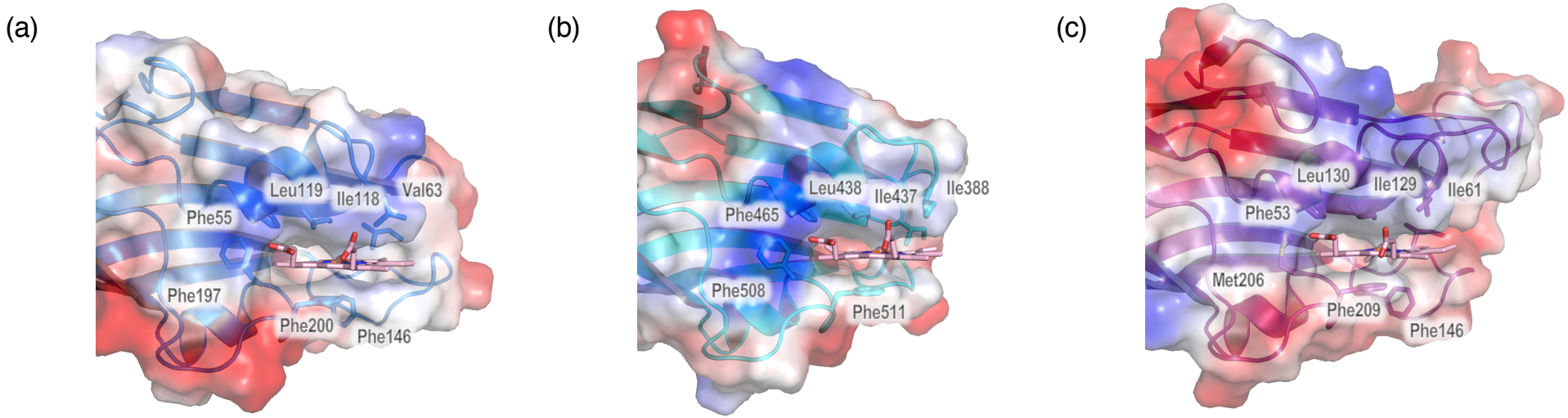


Fig. S5. Hydrophobic residues in heme binding pocket of (a) HtaA-N, (b) HtaA-C, and (c) HtaB. Only one orientation of the heme is shown in (c) for clarity.

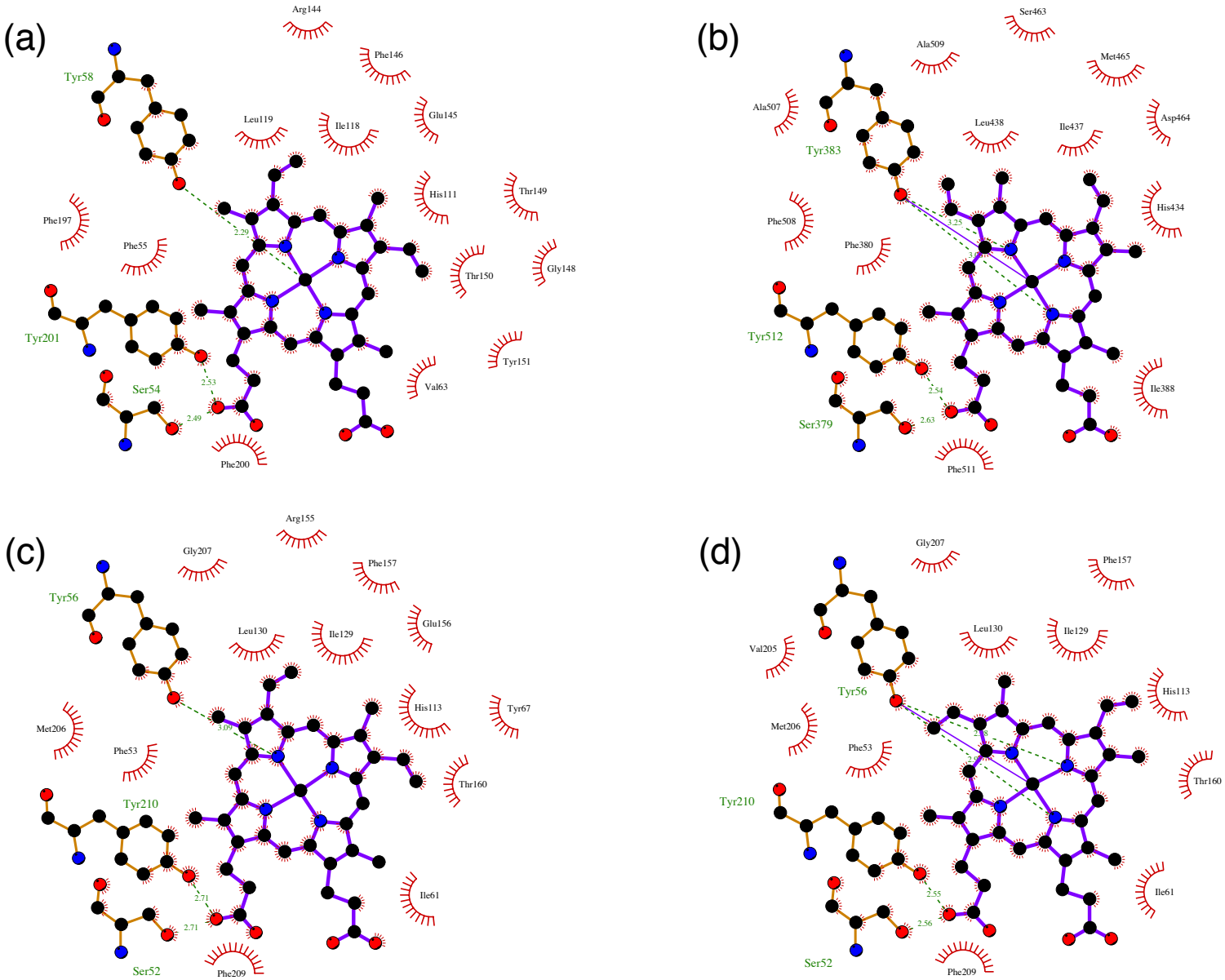


Fig. S6. Lig-plot indicating the interactions between heme and surrounding residues for (a) HtaA-N, (b) HtaA-C, and (c) HtaB with the orientation I of heme, and (d) HtaB with the orientation II of heme.

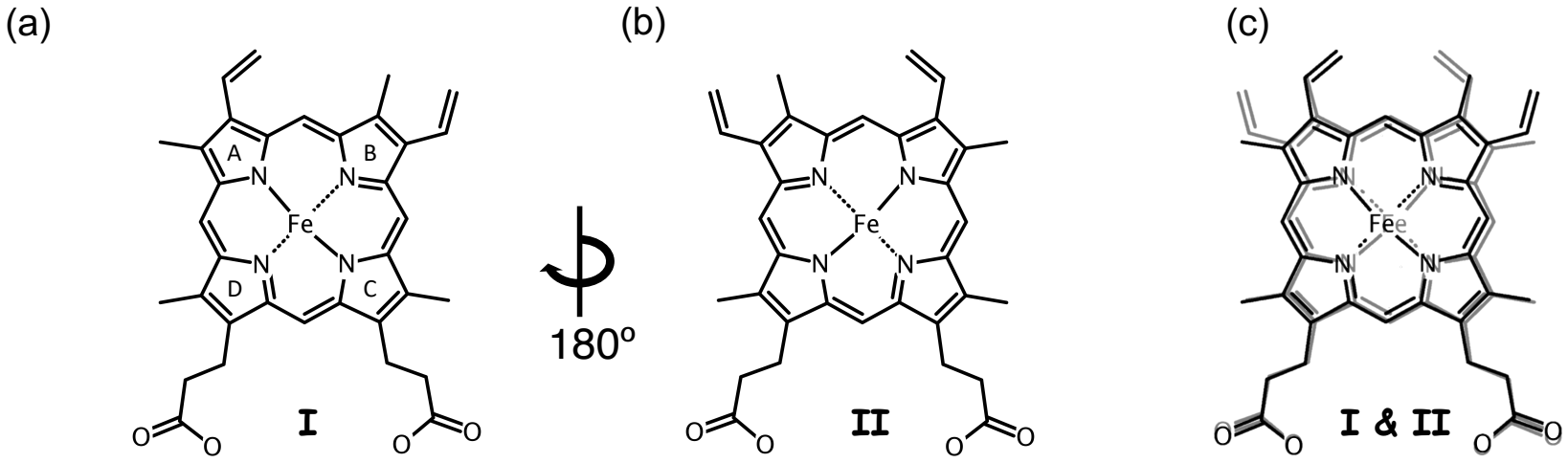


Fig. S7. The observed orientation of heme in (a) HtaA-N, (b) HtaA-C, and (c) HtaB, which are viewed from the vacant axial position. The orientations of heme bound to HtaB is modeled with a mixture of the orientations of I and II.

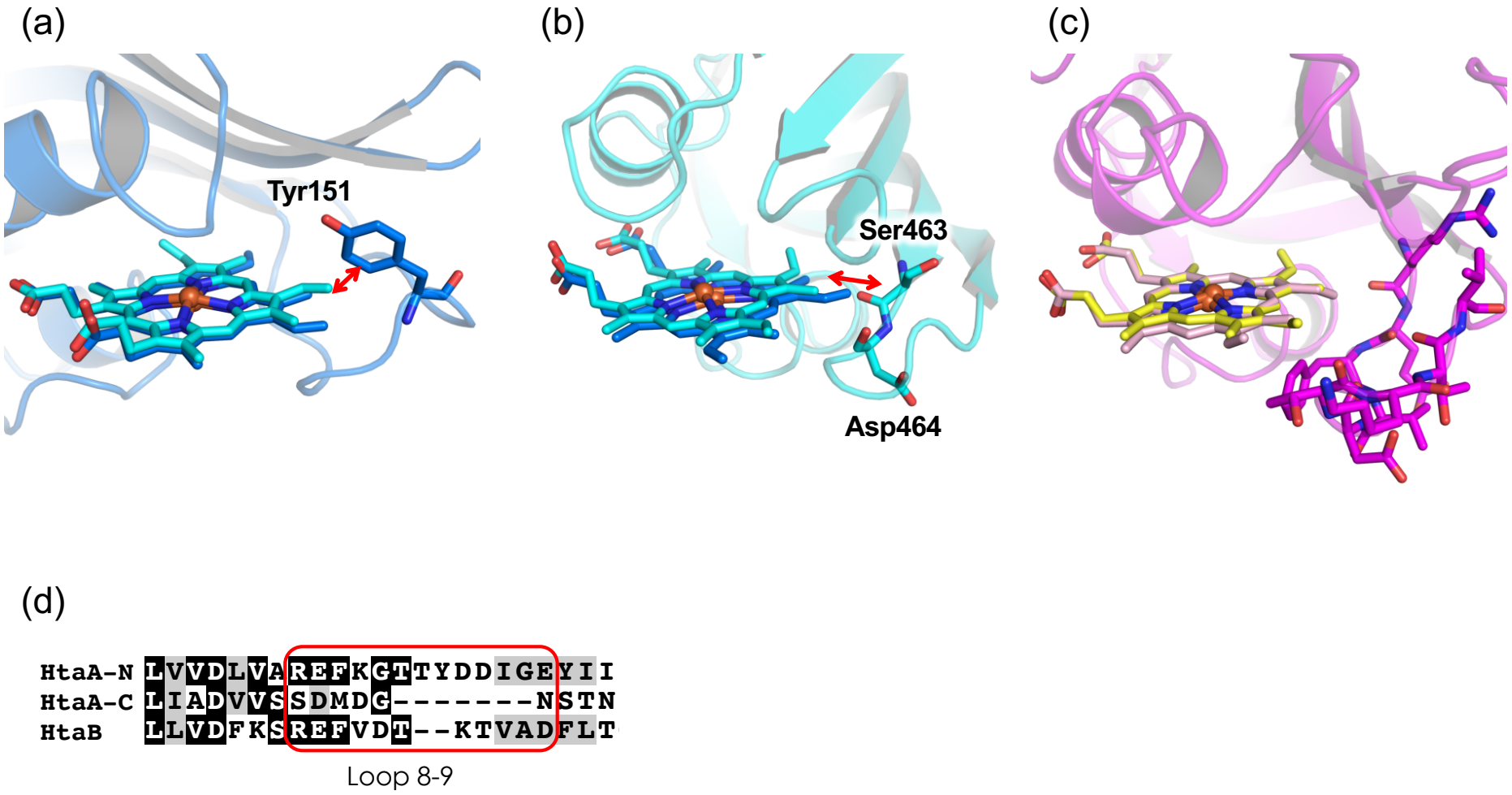


Fig. S8. The orientation of heme in (a) HtaA-N, (b) HtaA-C, and (c) HtaB, in which the orientation of heme observed in X-ray crystallographic analyses is shown in blue and cyan for HtaA-N and HtaA-C, respectively. The opposite orientation of heme is superimposed in cyan and blue for HtaA-N and HtaA-C, respectively. In the case of HtaB, both orientations of heme shown in (c) were observed in its crystal structure. (d) Amino acid sequences alignment around the loop 8-9 among HtaA-N, HtaA-C, and HtaB. The loop 8-9 is shown in a red box.

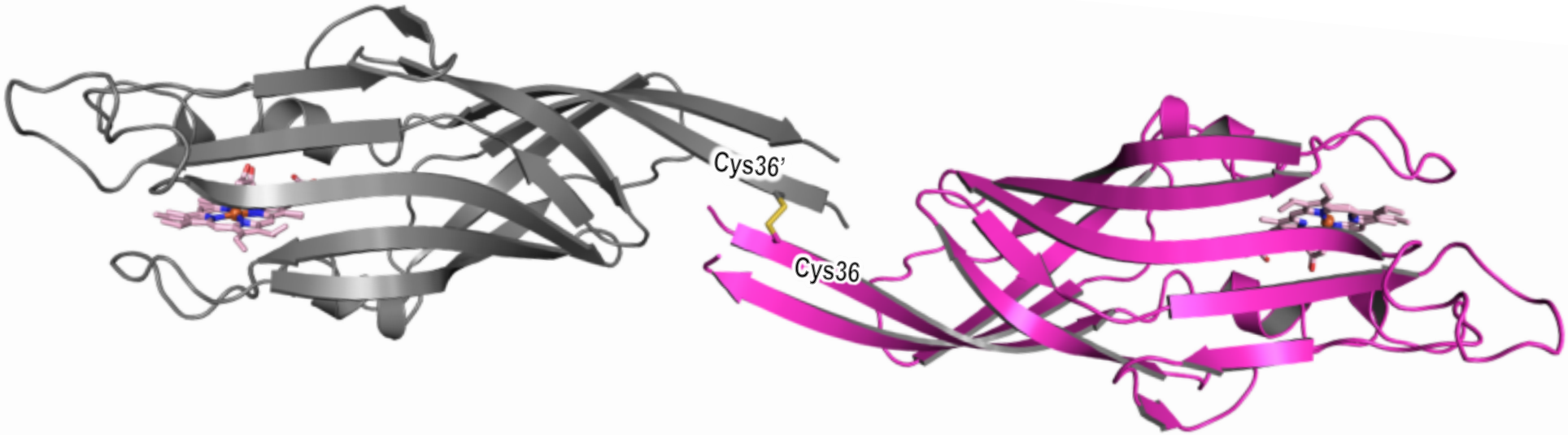


Fig. S9. The dimeric structure of HtaB. Each protomers are shown in different colors.

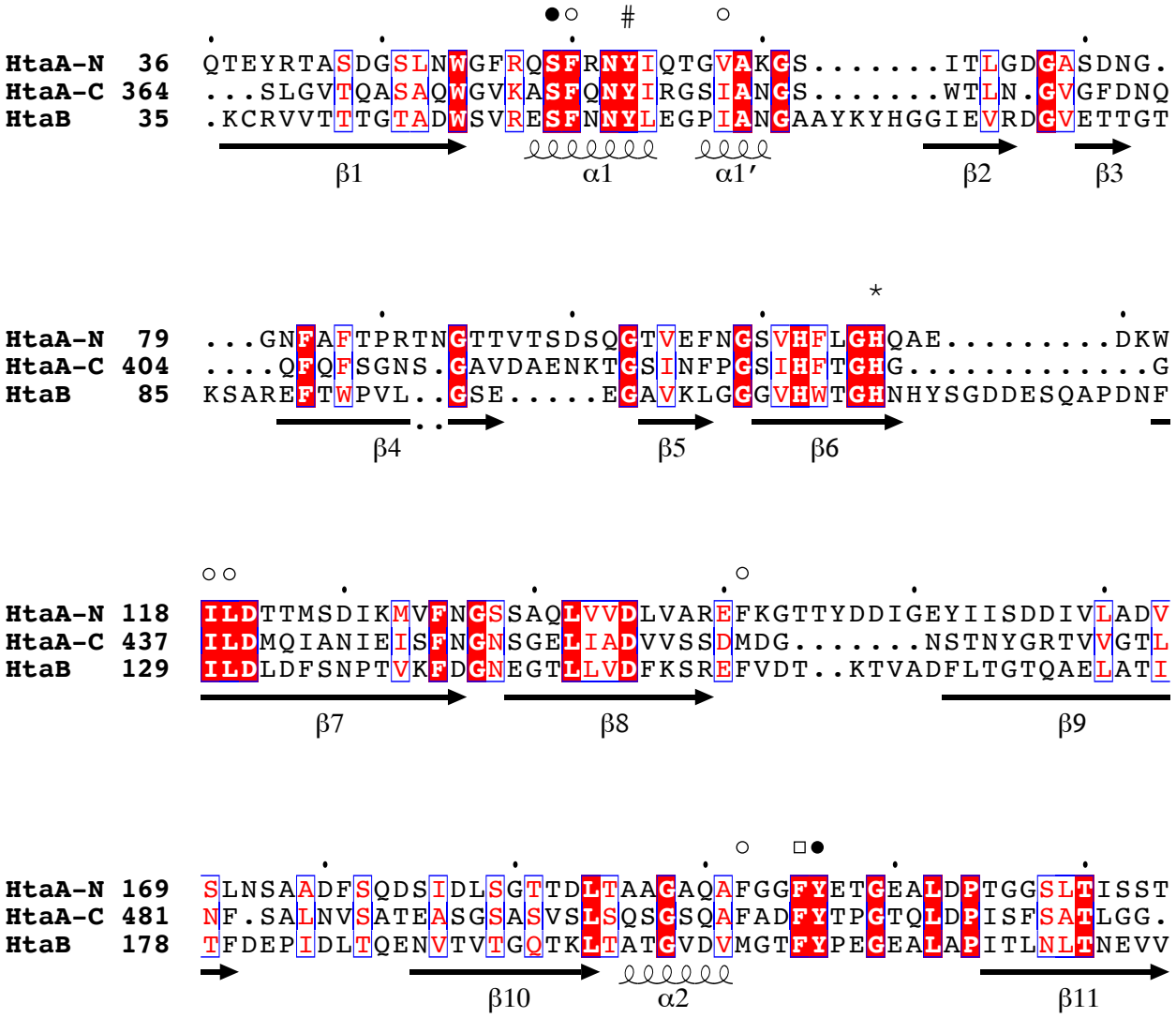
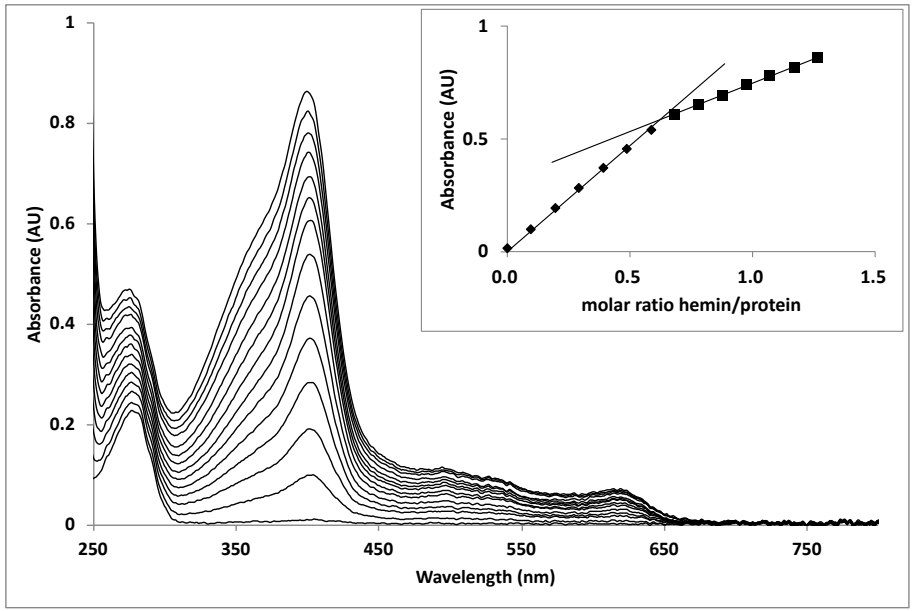
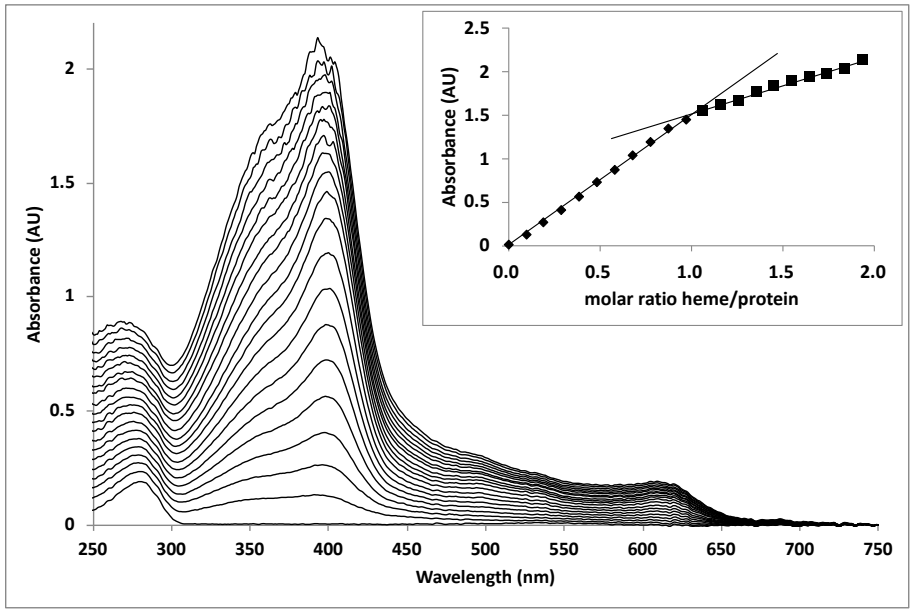


Fig. S10. Amino acid sequences alignment among HtaA-N, HtaA-C, and HtaB based on their crystal structures. The secondary structures are shown below the sequence, in which α helix and β strand are shown in an arrow and a coil, respectively. The identical and similar residues are shown in white and red, respectively. The axial ligand of heme, His interacting with the axial ligand, Tyr forming a π–π stacking with heme, the residues hydrogen bonding to heme propionate, and the hydrophobic residues interacting with heme in the heme pocket are marked with a hash tag, an astarisk, an open square, a closed circle, and an open circle, respectively.

(a) H111A-HtaA-N



(b) H434A-HtaA-C



(c) H113A-HtaB

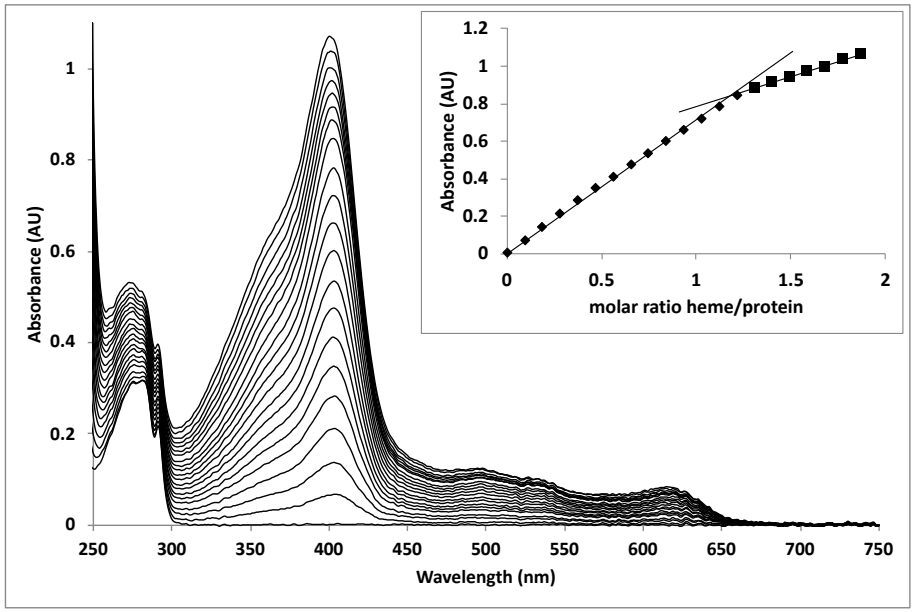


Fig. S11. Spectral change upon titration of hemin for (a) H111A-HtaA-N, (b) H434A-HtaA-C, and (c) H113A-HtaB. Titration curves are shown in the inset.

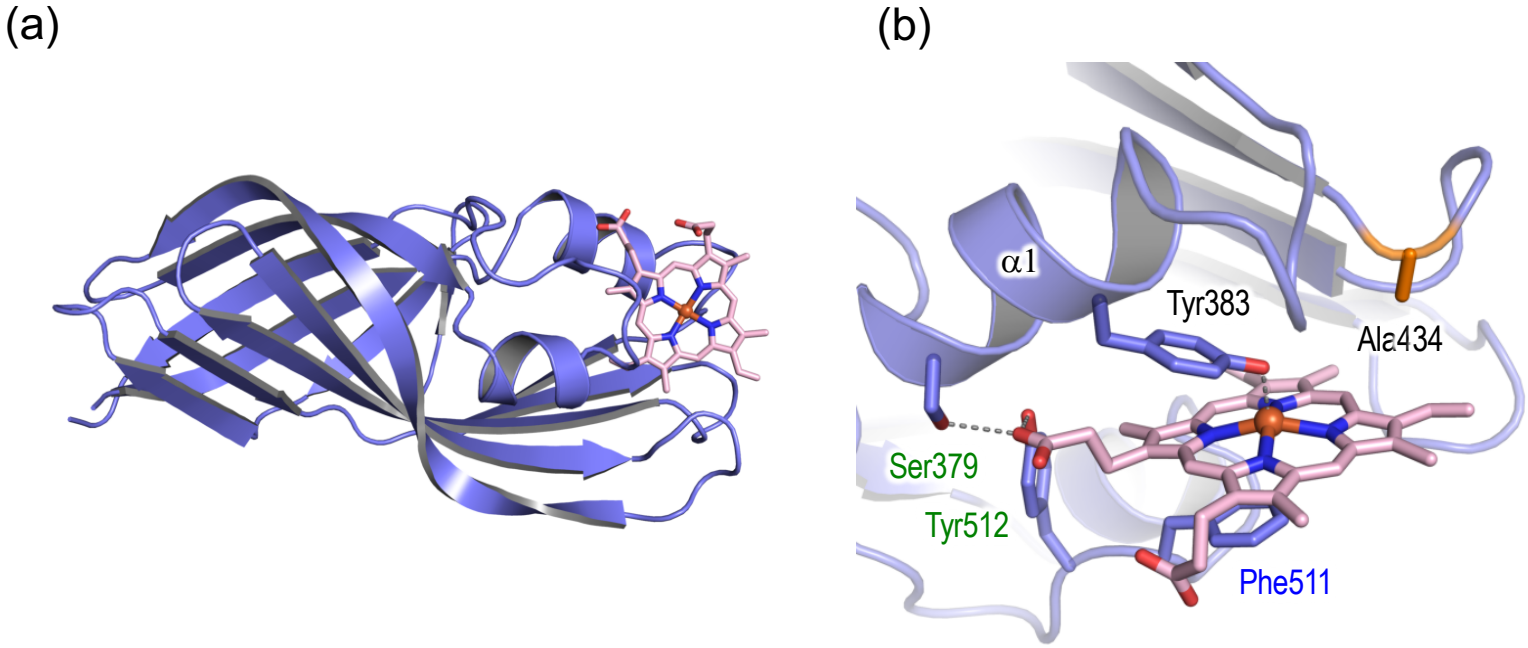


Fig. S12. (a) Overall structure of H434A-HtaA-C in the holo form. (b) Close-up view of heme binding region of this variant.

(a)

(b)

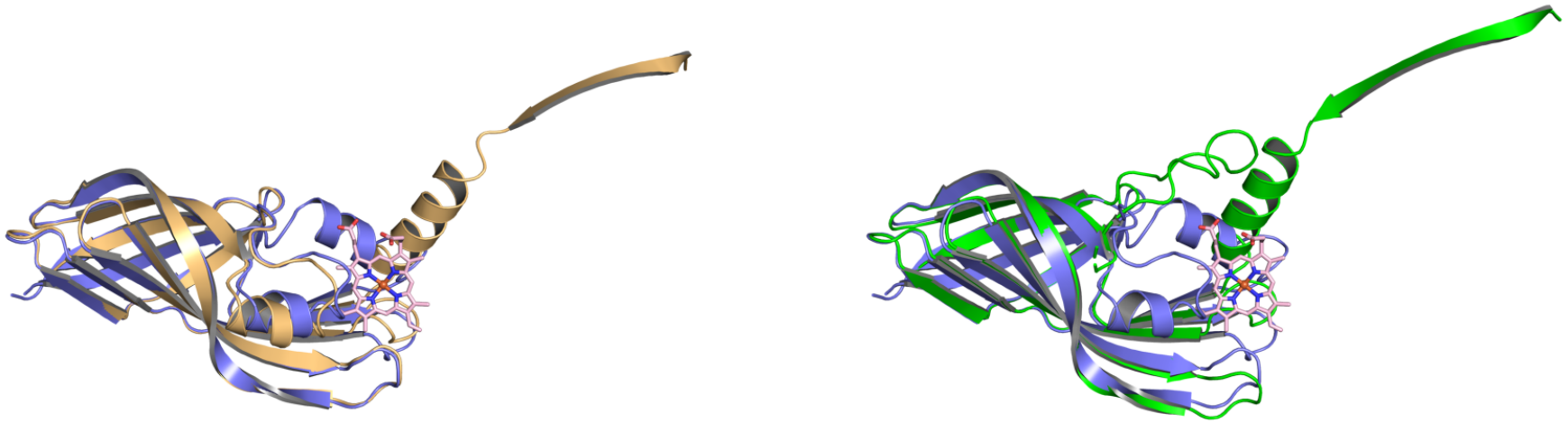


Fig. S13. Superposition of the structures (a) between the chainAof apo-H434A-HtaA-C (ocher) and holo-H434A-HtaA-C (blue) and (b) between the chain B of apo-H434A-HtaA-C (light green) and holo-H434A-HtaA-C (blue).

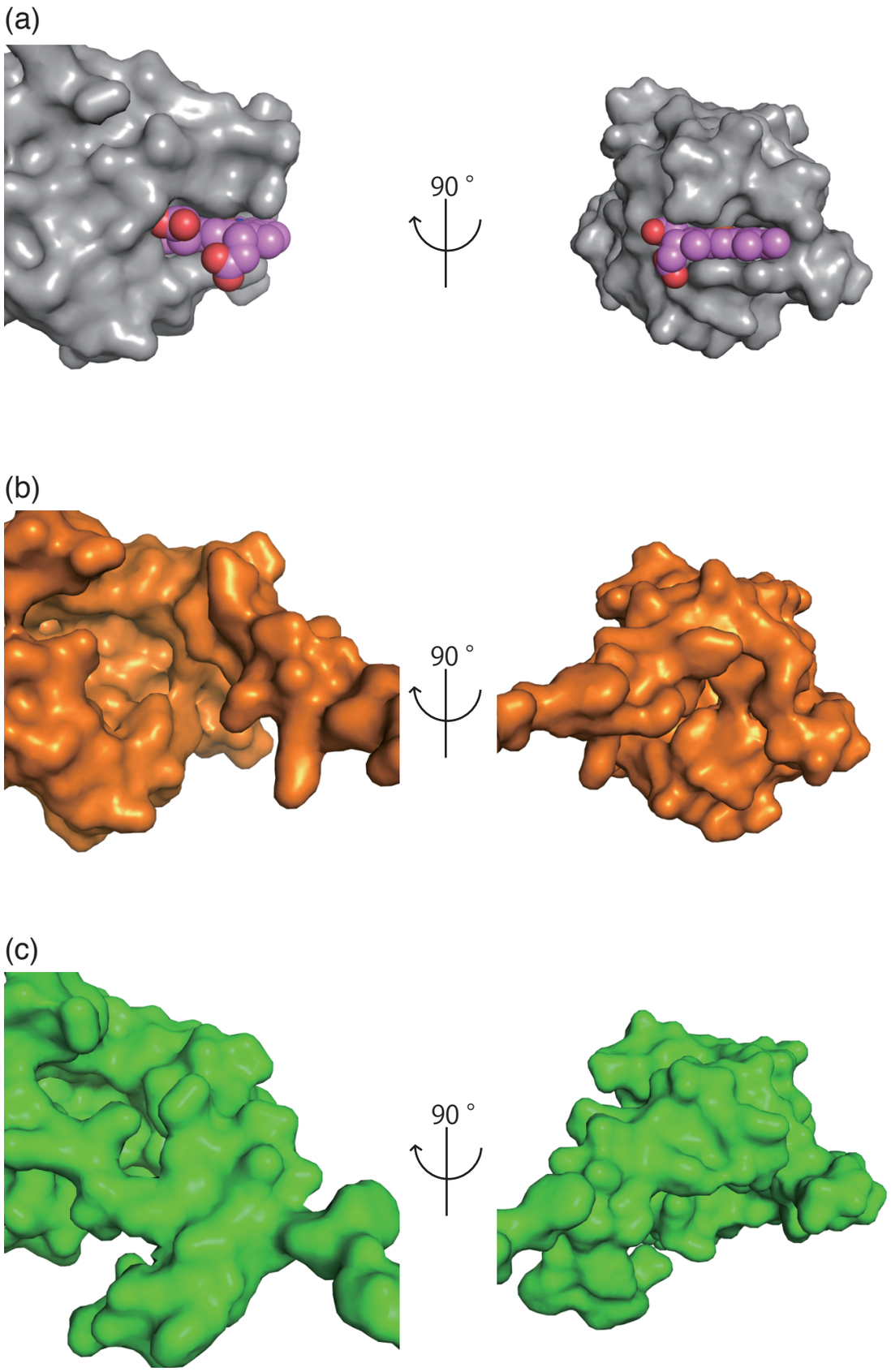


Fig. S14. Structural differences around the heme pocket in (a) the holo-form of HtaA-N, (b) the apo-form of H434A-HtaA-C (chain A), and (c) the apo-form of H434A-HtaA-C (chain B). The protein part and heme are shown in the surface representation and space-filling models, respectively.

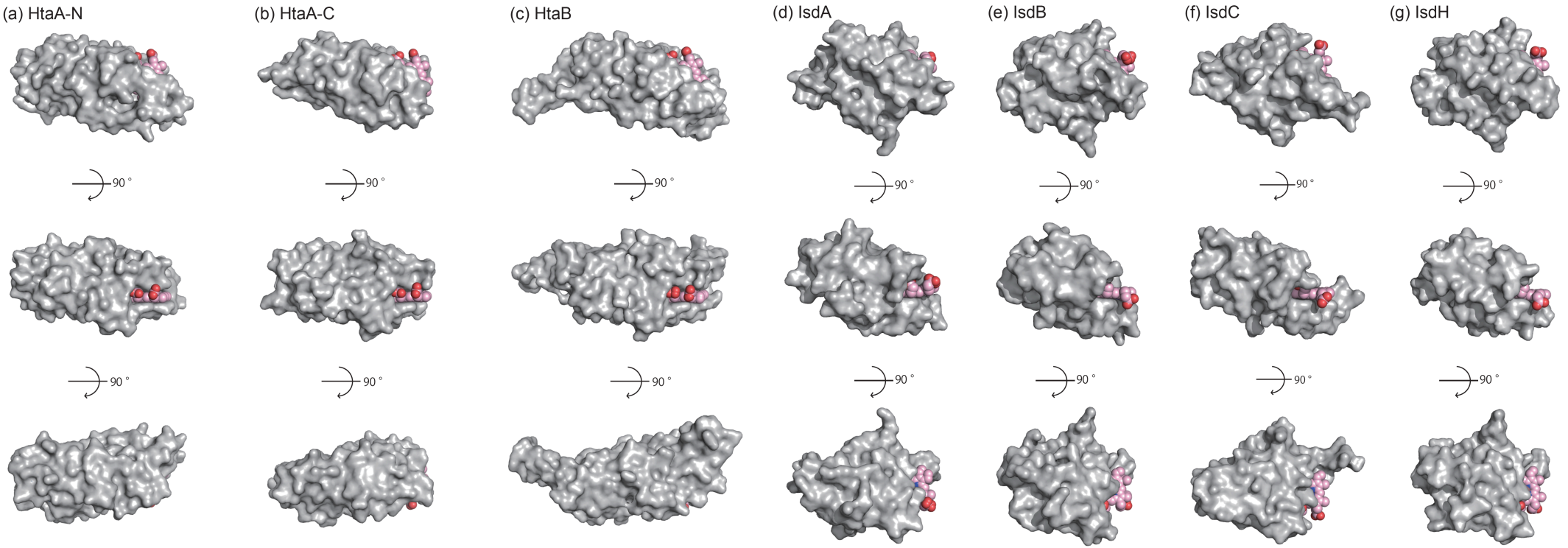


Fig. S15. Structural differences for the heme exposure to solvent in (a) HtaA-N, (b) HtaA-C, and (c) HtaB, (d) IsdA (PDB code: 2ITF), (e) IsdB (PDB code: 3RTL), (f) IsdC (PDB code: 2O6P), and (g) IsdH (PDB code: 3VTM). The protein part and heme are shown in the surface representation model (gray) and the space-filling model (pink and red), respectively.

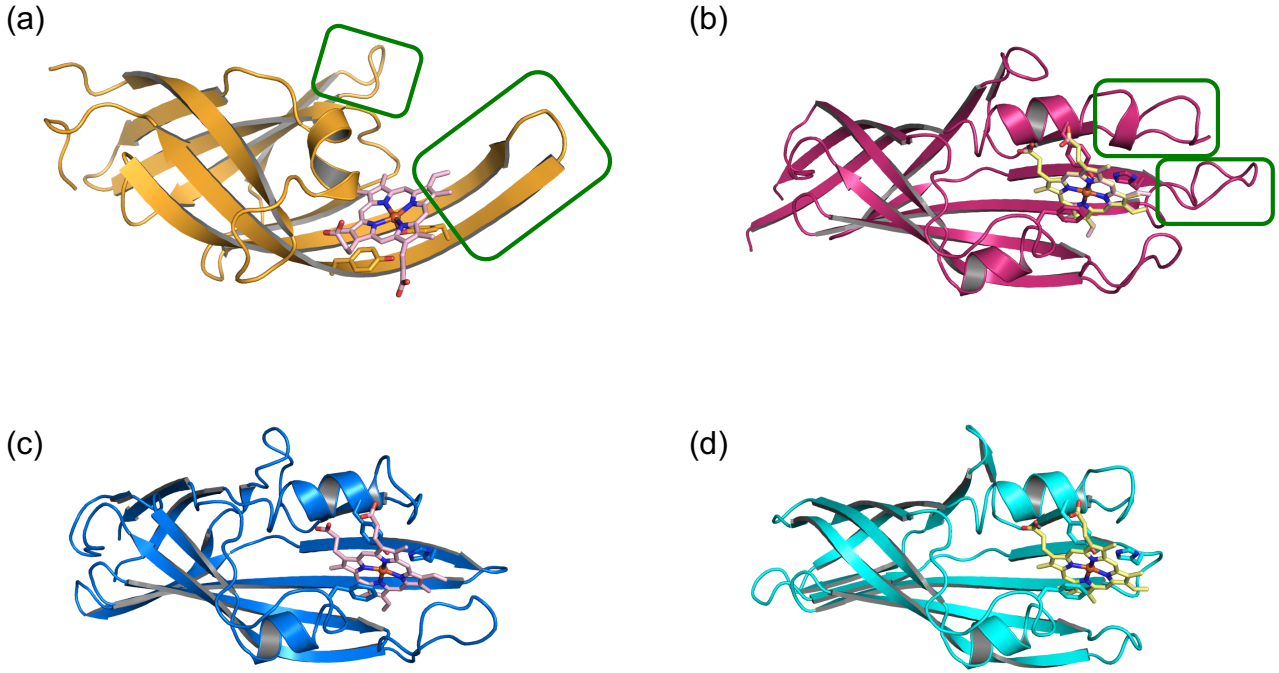


Fig. S16. The structures of (a) IsdC, (b) HtaB, (c) HtaA-N, and (d) HtaA-C. The loops interacting with IsdE during heme transfer are shown in green boxes in (a). (b) The loops1-2 and 7-8 in HtaB, which are shown in green boxes, are longer compared with HtaA-N and HtaA-C. These loops may interact with HmuT during heme transfer, as is the case of IsdC/IsdE.