Electronic Supplementary Information

Highly malleable haem-binding site of the haemoprotein HasA permits stable accommodation of bulky tetraphenylporphycenes

Erika Sakakibara,[a] Yuma Shisaka,[a] Hiroki Onoda,[a] Daiki Koga,[b] Ning Xu,[b] Toshikazu Ono,[b] Yoshio Hisaeda,[b] Hiroshi Sugimoto,[c,d] Yoshitsugu Shiro,[e] Yoshihito Watanabe,[f] Osami Shoji*[a,c]

Materials and Methods

All chemical reagents were purchased from commercial sources (e.g. Wako, TCI, and Aldrich) and used without further purification until otherwise noticed. Reactions heated by microwave irradiation were performed in a Monowave400 microwave (Anton-Paar) with internal temperature control.

Construction of an expression system for TEV protease

The synthetic gene of TEV protease coded in the plasmid pRK793[1] (MBP-TEV cleavage site-His_6-TEV protease/S219V; MBP-TEV) was purchased from GENEWIZ. The codon of gene sequence was optimised for expression in *Escherichia coli* cells. The MBP-TEV gene fragment was initially cloned into the plasmid pET11a using *Nde*I and *Bam*HI restriction sites. However, a large amount of TEV protease was inactivated due to self-cleavage for removal of MBP domain. To address this issue, a part of TEV recognition site was removed by inverse PCR using the following primers: 5'-CACCACCATCACCATCATCACGGCGAAAGCC-3' and 5'-ACCACGGCCCTCGATGCCGCGT-AAGTTGTT-3'). The resulting plasmid pET11a-MBP-TEV was then introduced into *E. coli* BL21(DE3) cells. The eliminated bases in TEV recognition site are shown in the bottom gene sequence as bold font. Recognition sites of *Eco*RI and *Bam*HI are underlined.

Optimised gene sequence of MBP-TEV cleavage site-His_6-TEV protease/S219V (MBP-TEV)

CATATGAAGATCGAAGAAGGCAAACTGGTTATTTGGATTAACGGCGATAAAGGCTATAACGGCCTGG
CCGAAGTGGGTAAAAAGTTTGAGAAAGACACGGGTATTAAAGTGACCGTTGAGCATCCGGACAAACT
GGAAGAAAAGTTCCCTCAGGTGGCCGCCACCGGTGATGGTCCGGATATCATCTTTTGGGCGCATGAC
CGTTTTGGTGGTTATGCGCAGAGCGGCCTGCTGGCGGAAATCACCCCGGATAAGGCGTTCCAGGACA
AACTGTATCCTTTCACGTGGGATGCGGTGCGCTACAACGGCAAGCTGATCGCCTATCCGATTGCCGT
GGAGGCGCTGAGCCTGATTTACAACAAAGACCTGTTACCTAACCCGCCGAAAACCTGGGAGGAAATT
CCGGCCCTGGATAAGGAACTGAAGGCCAAAGGCAAATCGGCCCTGATGTTTAACCTGCAGGAGCCGT
ATTTTACGTGGCCGCTGATTGCGGCGGACGGCGGTTACGCCTTCAAATACGAAAATGGCAAGTACGA
CATTAAGGATGTTGGCGTGGATAACGCGGGTGCGAAAGCGGGCCTGACCTTCCTGGTGGATTTAATC

AAAAACAAACACATGAATGCCGACACCGACTACTCGATCGCGGAGGCGGCCTTTAACAAGGGCGAGA
CCGCGATGACCATTAATGGCCCGTGGGCCTGGAGCAATATCGACACCTCGAAAGTGAACTACGGTGT
TACGGTGCTGCCGACCTTCAAAGGCCAACCTAGCAAGCCGTTCGTGGGTGTGCTGAGCGCGGGTATT
AACGCCGCGTCGCCGAATAAGGAATTAGCGAAAGAATTTCTGGAAAATTACCTGCTGACGGACGAGG
GCCTGGAAGCGGTTAACAAAGATAAGCCGCTGGGCGCCGTGGCGCTGAAATCGTATGAGGAAGAGCT
GGCCAAAGATCCTCGCATCGCCGCCACCATGGAGAACGCGCAGAAGGGCGAGATTATGCCTAACATC
CCGCAAATGAGCGCCTTCTGGTACGCGGTGCGCACCGCGGTGATTAATGCCGCCAGCGGCCGCCAGA
CCGTTGATGAAGCCCTGAAAGACGCCCAGACGAACAGCTCGAGCAACAACAACAATAACAATAATAA
TAACAACTTACGCGGCATCGAGGGCCGTGGT**GAAAATCTGTACTTCCAGGGT**CACCACCATCACCAT
CATCACGGCGAAAGCCTGTTCAAAGGTCCTCGCGACTACAACCCGATTTCGAGCACCATCTGCCATC
TGACCAACGAGAGCGATGGTCACACGACGAGCCTGTATGGCATCGGCTTTGGCCCGTTTATCATCAC
GAACAAGCACCTGTTCCGCCGCAACAACGGTACGCTGCTGGTGCAGAGCCTGCACGGCGTGTTCAAG
GTGAAAAACACCACGACGCTGCAACAACACTTAATCGACGGTCGCGACATGATCATCATCCGTATGC
CGAAAGATTTCCCGCCGTTCCCGCAAAAGCTGAAATTTCGCGAACCGCAGCGTGAGGAACGTATTTG
TCTGGTGACCACCAATTTCCAGACCAAGAGCATGAGCAGCATGGTGAGCGATACCAGCTGTACCTTT
CCGAGCAGCGATGGCATTTTCTGGAAACACTGGATTCAAACCAAGGATGGCCAGTGTGGCAGCCCGC
TGGTTAGCACGCGTGACGGCTTCATTGTGGGCATTCACAGCGCCAGCAATTTCACGAACACCAATAA
TTACTTCACCTCGGTGCCGAAGAACTTTATGGAACTGCTGACCAACCAGGAGGCCCAGCAGTGGGTG
AGCGGTTGGCGTCTGAACGCGGACTCGGTGTTATGGGGCGGCCATAAAGTGTTCATGGTTAAACCGG
AAGAGCCGTTCCAACCGGTGAAAGAGGCGACCCAGCTGATGAACCGTCGCCGTCGCCGCTAATAAGG
ATCC

Expression and Purification of TEV protease

A single colony of *E. coli* transformed with pET11a-MBP-TEV was inoculated in 3 mL of a Luria–Bertani (LB) medium including ampicillin (0.1 mg/mL) and cultivated overnight with shaking. An aliquot of bacterial culture was then transferred to 100 mL of LB medium including 0.1 mg/mL ampicillin and cultured overnight. *E. coli* culture was finally scaled up by 1 L of LB medium using a 5 L baffled flask. After 2 h incubation at 37 °C, isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM. *E. coli* was further incubated for 22 h at 20 °C. Bacterial cells were harvested by centrifugation and stored at –80 °C. The frozen cells were suspended in a TEV buffer (50 mM sodium phosphate, 10%(v/v) glycerol, 200 mM NaCl, pH 8.0) and disrupted by ultrasonicator on an ice bath. Cell debris was removed by centrifugation at 38000 g for 30 min, and the supernatant containing TEV protease was loaded onto a Ni-affinity column (HiTrap Chelating HP; GE Healthcare) equilibrated with a TEV buffer. The bound proteins were then eluted with 12 column volumes of a linear gradient from 0% to 100% TEV buffer including 250 mM imidazole. The eluates were dialysed overnight against a TEV buffer and further purified via gel filtration chromatography

(HiPrep 16/60 Sephacryl S-100 HR; GE Healthcare) equilibrated with a TEV buffer. Eluted fractions were stored at –80 °C until use.

Construction of an expression system for His-tagged HasA including TEV protease cleavage site

The plasmid pQE30t-*hasAp*(T), which is the plasmid pQE30t inserted *hasAp*(T) gene,[2] was used as a template of polymerase chain reaction (PCR). To introduce the recognition site of TEV protease (Glu-Asn-Leu-Tyr-Phe-Gln-Gly) between N-terminal $(His)_6$-tag and truncated HasA, PCR was performed using the following primers: Forward primer; 5'-GAGCGGATAACAATTTCACACACAAGAATTC-ATTAAGAGG-3' (underlined: *Eco*RI recognition site for cloning into the plasmid pQE30t-*hasAp*(T)), Reverse primer; 5'-CATGGA**TCC**TTGGAAGTACAGGTTCTCCGAGTGATGGTGATGGTGAT-GCGATCC-3' (Underlined: *Bam*HI recognition site for cloning into the plasmid pQE30t-*hasAp*(T), **bold font**: TEV protease recognition site). The obtained PCR product was cloned into pQE30t-*hasAp*(T) using Ligation high Ver. 2 (TOYOBO). The resulting plasmid (pQE30tev-*hasAp*(T)) was then introduced into *Escherichia coli* strain M15.

Expression and Purification of Apo-HasA

E. coli cells expressing His-tagged HasA were suspended in an H buffer (20 mM sodium phosphate, 15 mM 2-mercaptoethanol, 10 mM imidazole, 500 mM NaCl, pH 7.4) and disrupted by ultrasonicator on an ice bath. Cell debris was removed by centrifugation at 38000 g for 30 min, and the supernatant containing His-tagged HasA was loaded onto a Ni-affinity column (cOmplete His-Tag Purification Resin; Roche). After washing the column with H buffer, the bound proteins were eluted with H buffer containing 200 mM imidazole. The eluates were treated with TEV protease[3] to remove His-tag on HasA and dialysed overnight against a phosphate-buffered saline (PBS; 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) solution. This sample was purified via a Ni-affinity column equilibrated with PBS solution. Purified HasA solution was dropped in to acetone containing 0.2% (v/v) HCl to remove haem. The precipitate of HasA was collected by centrifugation and dissolved in 100 mM Tris-HCl (pH 7.5) including 7 M urea. Denatured HasA solution was dialysed overnight against a PBS solution. Resulting solution was concentrated using Amicon Ultra (Merck Millipore, 3 kDa cutoff) and purified via a gel filtration column (HiPrep 16/60 Sephacryl S-200 HR; GE Healthcare) equilibrated with a PBS solution to remove misfolded HasA. The concentration of HasA without haem (apo-HasA) was determined by the absorption at 280 nm (ε_{280} = 28.6 mM^{-1} cm^{-1}).[4] Purified apo-HasA solution was frozen by liquid nitrogen and stored at –80 °C until use.

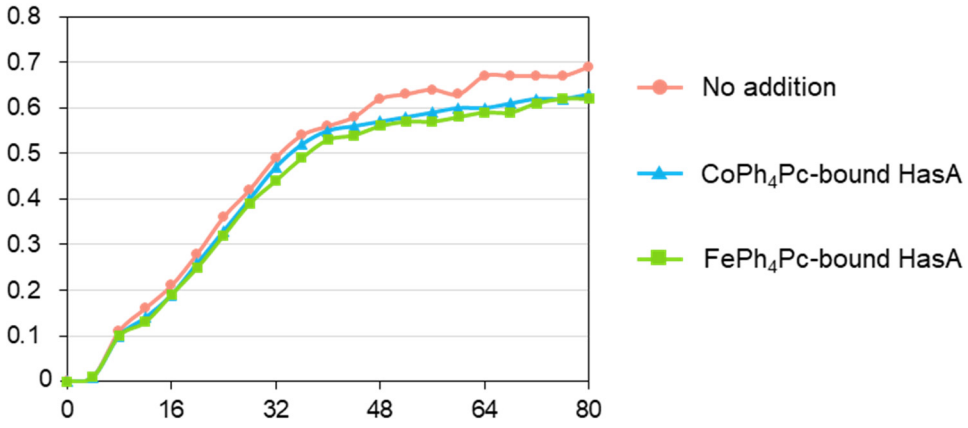

Figure S1 Growth curve of *P. aeruginosa* in M9-based medium without EDTA. *P. aeruginosa* was grown in the presence or absence of HasA with metallo-Ph_4Pcs in the case of using M9-based medium without EDTA. This result indicated that metallo-Ph_4Pcs showed no cytotoxicity.

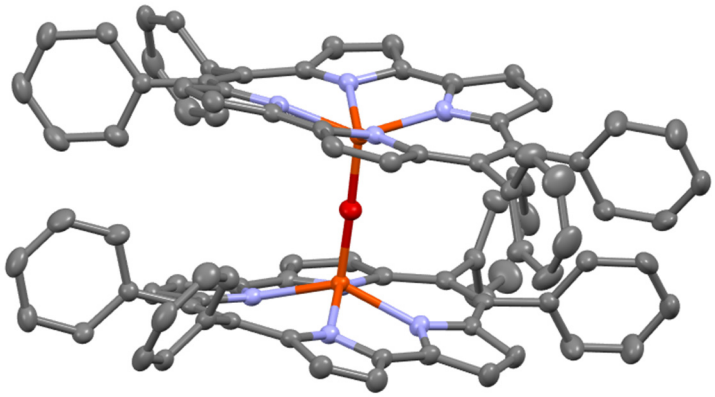

Figure S2. Crystal structure of (Fe-Ph_4Pc) μ-oxodimer. The thermal ellipsoids represent for 50% probability. Orange = Fe, Red = O, purple = N, and Grey = C.

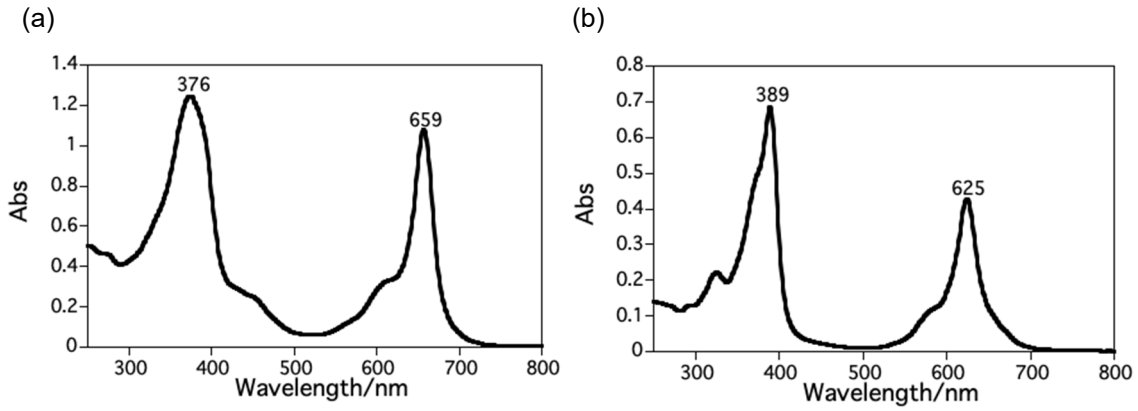

Figure S3. UV-vis spectra of (a) (Fe-Ph_4Pc) μ-oxodimer and (b) Co-Ph_4Pc in CH_2Cl_2 at room temperature.

Table S1

Metal complex coordinating HasA (PDB ID)	Co-Ph_4Pc HasA (6JLG)
Data collection	
Wavelength (Å)	1.0
Space group	*C* 1 2 1
Cell dimensions	
a, *b*, *c* (Å)	237.58, 92.22, 93.19
α, β, γ (°)	90.00, 100.96, 90.00
Resolution range [a] (Å)	46.83 - 2.50 (2.56 - 2.50)
No. of total observed reflections [a]	470308 (33264)
No. of unique reflections [a]	68363 (4603)
R_{merge} [a,b]	0.231 (1.796)
Completeness [a] (%)	99.7 (99.9)
$I/\sigma(I)$ [a]	7.0 (1.5)
$CC_{1/2}$ [a]	0.991 (0.512)
Redundancy [a]	6.9 (7.2)
Refinement statistics	
Resolution range (Å)	45.79 - 2.50
No. of monomer/asymmetric unit	8
R_{work}/R_{free} [c,d] (%)	21.29/24.06
RMSD bond lengths [e] (Å)	0.0032
RMSD bond angles [e] (°)	1.1759
No. of non-hydrogen atoms	11180
Average *B*-factor (Å^2)	54.00

[a] The values in parentheses are for the highest resolution sell.

[b] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of the *i* oservations.

[c] $R_{work} = \Sigma_{hkl}|F_{obs}(hkl) - F_{calc}(hkl)|\Sigma_{hkl}F_{obs}(hkl)$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[d] R_{free} was calculated with 5% of the reflections that were not included in the refinement.

[e] r. m. s. d. = root mean square deviation

Supplementary References

1. R. B. Kapust, J. Tozser, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland and D. S. Waugh, *Protein Eng., Des. Sel.*, 2001, **14**, 993-1000.
2. C. Shirataki, O. Shoji, M. Terada, S.-i. Ozaki, H. Sugimoto, Y. Shiro and Y. Watanabe, *Angew. Chem., Int. Ed.*, 2014, **53**, 2862-2866.
3. J. E. Tropea, S. Cherry and D. S. Waugh, *Methods Mol. Biol.*, 2009, **498**, 297-307.
4. G. Jepkorir, J. C. Rodriguez, H. A. Rui, W. Im, S. Lovell, K. P. Battaile, A. Y. Alontaga, E. T. Yukl, P. Moenne-Loccoz and M. Rivera, *J. Am. Chem. Soc.*, 2010, **132**, 9857-9872.