Electronic Supplementary Information

# Core-shell protein clusters comprising haemoglobin and recombinant feline serum albumin as an artificial

## O<sub>2</sub> carrier for cats

Kyoko Yokomaku,<sup>a</sup> Motofusa Akiyama,<sup>a</sup> Yoshitsugu Morita,<sup>a</sup> Kiyohito Kihira<sup>b</sup> and Teruyuki Komatsu\*<sup>a</sup>

- <sup>a</sup> Department of Applied Chemistry, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan.
- <sup>b</sup> JEM Utilization Center, Human Spaceflight Technology Directorate, Japan Aerospace Exploration Agency (JAXA), 2-1-1 Sengen, Tsukuba-shi, Ibaraki 305-8505, Japan.

Corresponding Author: Prof. Dr. Teruyuki Komatsu

E-mail: komatsu@kc.chuo-u.ac.jp

#### Experimental

#### Purification of native FSA from feline plasma

Feline plasma was provided from Kyoritsu Seiyaku Corp. The frozen sample was thawed slowly in a refrigerator overnight at 4 °C and was centrifuged (10,000 × g, 30 min, 4 °C) to remove the cryoprecipitate. The supernatant was filtered through a track-etched polycarbonate membrane (Isopore membrane, 25 mm diameter, 50 nm pore; Millipore Corp.) to remove endogenous retroviruses. The filtrate was then brought to 50% saturation with ammonium sulfate. After leaving for 30 min at 4 °C, the solution was centrifuged (10,000 × g, 30 min, 2 °C) and the supernatant was filtered using a membrane filter (DISMIC-25CS, 0.2 µm pore; Toyo Roshi Kaisha Ltd.). The obtained solution was dialyzed against deionized water at 4 °C with subsequent addition of 11% volume of 500 mM sodium phosphate (pH 6.8). The resultant solution in 50 mM sodium phosphate (pH 7.0) was filtered using a membrane filter (C020A047A, 0.2 µm pore; Toyo Roshi Kaisha Ltd.).

The sample was applied to affinity chromatography (Toyopearl AF-Blue HC-650M; Tosoh Corp.). After washing with 50 mM sodium phosphate (pH 7.0), FSA was eluted with 50 mM sodium phosphate (pH 7.4) containing 3 M NaCl. The eluent was dialyzed against deionized water at 4 °C. Thereafter, 25% volume of 100 mM Tris-HCl (pH 8.0) was added, and the resulting 20 mM Tris-HCl solution (pH 8.0) of FSA was filtered using a membrane filter (C020A047A, 0.2 µm pore). Then the sample was subjected to anion exchange chromatography (Q Sepharose Fast Flow; GE Healthcare UK Ltd.) with 20 mM Tris-HCl (pH 8.0) as the running buffer. After washing with 20 mM Tris-HCl (pH 8.0) containing 100 mM NaCl, elution of FSA was performed with 20 mM Tris-HCl (pH 8.0) containing 300 mM NaCl. The eluent was dialyzed against deionized water at 4 °C, followed by addition of 11% volume of 10× phosphate-buffered saline (PBS, pH 7.4). At the last, the FSA solution was concentrated to 30 mL, and was sterilized with a membrane filter (DISMIC-25CS, 0.2 µm pore). All the purification processes were confirmed by SDS-PAGE analysis. The concentration of FSA was measured using a protein assay kit (Pierce 660 nm; Thermo Fisher Scientific K.K.). The cysteinyl thiol assay of CSA was performed by reaction with 4,4'-dithiopyridine (4,4'-DTP).<sup>48</sup>



Fig. S1. MALDI-TOF MS spectra of rFSA and FSA.



**Fig. S2.** Superposition of crystal structures of rFSA (light blue) and rHSA (light green, PDB ID: 1E78).<sup>36</sup>



**Fig. S3.** Surface electrostatic potential representations of (A) rFSA and (B) rHSA. Blue and red respectively represent positive charge and negative charge density. Calculations were carried out using Adaptive Poisson-Boltzmann Solver (APBS) and PyMOL. The pdb files were converted to pqr files for APBS electrostatics calculations by PDB2PQR service.<sup>S1</sup> (PDB ID of rHSA: 1E78)<sup>36</sup>



**Fig. S4.** Time evolution of the RMSD of C $\alpha$  atoms from the starting structure during 10 ns MD simulations of rFSA and rHSA. RMSD values reached plateau at about 4 ns. (PDB ID of rHSA: 1AO6)<sup>37</sup>



Fig. S5. (A) Native-PAGE and (B) IEF of Hb-rFSA<sub>3</sub> cluster.



**Fig. S6.** (A) Immunological reactivity of Hb-rFSA<sub>3</sub> cluster against anti-HbA antibody. The result of quick chaser occult blood test kit. (B) Immunological reactivity of Hb-rFSA<sub>3</sub> cluster against anti-FSA antibody. The relation between FSA unit concentration of the sample and absorption intensity of the reactant solution at 694 nm.

**Table S1** Comparison of amino acid sequences of FSA and HSA. The first row represents FSA sequence (colors correspond to the subdomain colors in Figure 2) and the second row represents HSA sequence (black). The yellow-marked amino acids are different kind pairs between FSA and HSA. The homology of these proteins is 81.9%.

	1	11	21	31	41
FSA	EAH <mark>O</mark> SEIAHR	F <mark>N</mark> DLGEE <mark>H</mark> FR	<mark>g</mark> lvlvaf <mark>s</mark> oy	LOOCPFEDHV	KLVNEVTEFA
HSA	DAHKSEVAHR	F <mark>K</mark> DLGEE <mark>N</mark> FK	ALVLIAFAQY	LQQCPFEDHV	KLVNEVTEFA
	51	61	71	81	91
	k <mark>g</mark> cvad <mark>o</mark> sa <mark>a</mark>	NC <mark>E</mark> KSLH <mark>E</mark> LL	GDKLCTVA <mark>S</mark> L	R <mark>DK</mark> YGEMADC	C <mark>e</mark> k <mark>k</mark> eperne
	K <mark>T</mark> CVAD <mark>E</mark> SAE	NC <mark>D</mark> KSLH <mark>T</mark> LF	GDKLCTVATL	R <mark>ET</mark> YGEMADC	C <mark>AKQ</mark> EPERNE
	101	111	121	131	141
	CFLQHKDDNP	GFGQLV <mark>T</mark> PE <mark>A</mark>	D <mark>a</mark> mctafh <mark>e</mark> n	E <mark>OR</mark> FL <mark>G</mark> KYLY	EIARRHPYFY
	CFLQHKDDNP	NLPRLVRPEV	D <mark>V</mark> MCTAFH <mark>D</mark> N	E <mark>ET</mark> FL <mark>K</mark> KYLY	EIARRHPYFY
	151	161	171	181	191
	APELL <mark>YY</mark> A <mark>EE</mark>	YK <mark>GV</mark> FTECC <mark>E</mark>	AADKAACL <mark>T</mark> P	K <mark>V</mark> D <mark>A</mark> LR <mark>EKVL</mark>	ASSAK <mark>E</mark> RLKC
	APELL <mark>FF</mark> A <mark>KR</mark>	YK <mark>AA</mark> FTECC <mark>Q</mark>	aadkaacl <mark>l</mark> p	K <mark>l</mark> d <mark>e</mark> lr <mark>degk</mark>	ASSAK <mark>Q</mark> RLKC
	201	211	221	231	241
	ASLQKFGERA	FKAW <mark>S</mark> VARLS	Q <mark>k</mark> fpkaefae	<mark>i</mark> sklvtdl <mark>a</mark> k	<mark>I</mark> H <mark>K</mark> ECCHGDL
	ASLQKFGERA	FKAW <mark>A</mark> VARLS	Q <mark>r</mark> fpkaefae	<mark>v</mark> sklvtdl <mark>t</mark> k	VH <mark>T</mark> ECCHGDL
	251	261	271	281	291
	LECADDRADL	AKYICENQDS	IS <mark>T</mark> KLKECC <mark>G</mark>	KP <mark>V</mark> LEKSHCI	<mark>sever</mark> de <mark>l</mark> pa
	LECADDRADL	AKYICENQDS	IS <mark>S</mark> KLKECC <mark>E</mark>	KP <mark>L</mark> LEKSHCI	<mark>a</mark> eve <mark>n</mark> de <mark>m</mark> pa
	301	311	321	331	341
	301 DLP <mark>P</mark> LA <mark>V</mark> DFV	311 E <mark>D</mark> K <mark>E</mark> VCKNY <mark>O</mark>	321 EAKDVFLG <mark>T</mark> F	331 LYEY <mark>S</mark> RRHP <mark>E</mark>	341 YSV <mark>S</mark> LLLRLA
	301 DLP <mark>P</mark> LA <mark>V</mark> DFV DLP <mark>S</mark> LA <mark>A</mark> DFV	311 E <mark>DKE</mark> VCKNY <mark>O</mark> E <mark>S</mark> K <mark>D</mark> VCKNY <mark>A</mark>	321 EAKDVFLG <mark>T</mark> F EAKDVFLG <mark>M</mark> F	331 LYEY <mark>S</mark> RRHP <mark>E</mark> LYEY <mark>A</mark> RRHPD	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA
	301 DLP <mark>P</mark> LA <mark>V</mark> DFV DLP <mark>S</mark> LA <mark>A</mark> DFV 351	311 EDKEVCKNY <mark>O</mark> E <mark>S</mark> KDVCKNYA 361	321 EAKDVFLG <mark>T</mark> F EAKDVFLG <mark>M</mark> F 371	331 LYEY <mark>S</mark> RRHP <mark>E</mark> LYEY <mark>A</mark> RRHP <mark>D</mark> 381	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391
	301 DLPFLAVDFV DLPSLAADFV 351 KEYENTLEKC	311 ECKEVCKNYC ESKDVCKNYA 361 CA <mark>TD</mark> DP <mark>PA</mark> CY	321 EAKDVFLGTF EAKDVFLGMF 371 AEVFDEFKPL	331 LYEY <mark>S</mark> RRHP <mark>E</mark> LYEY <mark>A</mark> RRHPD 381 VEEP <mark>HNLV</mark> KT	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391 NCELFE <mark>K</mark> LGE
	301 DLPPLAVDFV DLPSLAADFV 351 Koyentlekc KTyentlekc	311 EKEVCKNYO ESKDVCKNYA 361 CATODPPACY CAAADPHECY	321 EAKDVFLG <mark>1</mark> F EAKDVFLGMF 371 AFVFDEFKPL AKVFDEFKPL	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391 NCELFE <mark>K</mark> LGE NCELFE <mark>Q</mark> LGE
	301 DLPELAVDFV DLPSLAADFV 351 KEYENTLEKC KTYETTLEKC 401	311 ESKOVCKNY ESKOVCKNYA 361 CATODPPACY CAAADPHECY 411	321 EAKDVFLGTF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421	331 LYEYSRRHPS LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391 NCELFEKLGE NCELFEQLGE 441
	301 DLPELAVDFV DLPELAVDFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR	311 EKEVCKNY ESKDVCKNYA 361 CATDDPPACY CAAADPHECY 411 YTKKVPQVST	321 EAKDVFLGTF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH	341 YSV <mark>9</mark> LLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE
	301 DLPFLAVDFV DLPSLAADFV 351 KUYENTLEKC KTYETTLEKC 401 YGFQNALLVR YKFQNALLVR	311 EKEVCKNY ESKDVCKNY 361 CALDDPFACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST	321 EAKDVFLGIF EAKDVFLGMF 371 AFVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRNL	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE
	301 DLPELANDFV DLPELANDFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR 451	311 EKOVCKNYQ ESKDVCKNYA 361 CATODPPACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST 461	321 EAKDVFLGFF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRNL 471	331 LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481	341 YSV <mark>S</mark> LLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491
	301 DLPELANDFV DLPSLAADFV 351 KEYEDTLEKC KTYETTLEKC 401 YGFQNALLVR YKFQNALLVR 451 DYLSVVLNRL	311 EKEVCKNY ESKDVCKNYA 361 CATODPPACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST 461 CVLHEKTPVS	321 EAKDVFLGTF EAKDVFLGMF 371 AVFDEFKPL 421 PTLVEVSRSL PTLVEVSRNL 471 ERVTKCCTES	331 LYEYSRRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481 LVNRRPCFSA	341 YSV <mark>S</mark> LLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK
	301 DLPFLAVDFV DLPSLAADFV 351 KEYENTLEKC KTYENTLEKC 401 YGFQNALLVR YKFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL	311 EKEVCKNY ESKDVCKNY 361 CATODPFACY CAAADPHECY 411 YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS	321 EAKDVFLGIF EAKDVFLGMF 371 AFVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRNL 471 ERVTKCCTES DRVTKCCTES	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481 LVNRRPCFSA LVNRRPCFSA	341 YSV <mark>S</mark> LLLRLA YSV <mark>V</mark> LLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK
	301 DLPELAVDFV DLPSLAADFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR YKFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL 501	311 EKEVCKNY ESKDVCKNYA 361 CATODPPACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS 511	321 EAKDVFLG F EAKDVFLG F 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSR L PTLVEVSR L 471 ERVTKCCTES DRVTKCCTES 521	331 LYEYSRRHPS LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481 LVNRRPCFSA LVNRRPCFSA 531	341 YSVSLLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK 541
	301 DLPFLAVDFV DLPSLAADFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL 501 EFSAETFTFH	311 EKOVCKNYQ ESKDVCKNYA 361 CATDDPPACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS 511 ADCTLPEAE	321 EAKDVFLGTF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRSL 471 ERVTKCCTES DRVTKCCTES 521	331 LYEYSRRHPS LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481 LVNRRPCFSA LVNRRPCFSA 531 ELKHKPKAT	341 YSVSLLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK 541 EEQLKTVMOD
	301 DLPELAVDFV DLPSLAADFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR YKFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL 501 EFSAETFTFH EFNAETFTFH	311 EKEVCKNY ESKDVCKNYA 361 CATODPPACY CAAADPHECY 411 YTKKVPQVST YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS 511 ADLCTLFEAE ADICTLSEKE	321 EAKDVFLGTF EAKDVFLGMF 371 AVFDEFKPL 421 PTLVEVSRSL PTLVEVSRSL 471 ERVTKCCTES DRVTKCCTES 521 QIKKQSALV RQIKKQTALV	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCKH 481 LVNRRPCFSA LVNRRPCFSA 531 ELKHKPKAT ELVKHKPKAT	341 YSVSLLLRLA YSVSLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK 541 SEQLKIVMGD KEQLKAVMDD
	301 DLPTLAVDFV DLPSLAADFV 351 KTYENTLEKC KTYENTLEKC 401 YGFQNALLVR YKFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL 501 EFSAETFTFH EFNAETFTFH	311 EKEVCKNYQ ESKDVCKNYA 361 CATODPPACY CAAADPHECY 411 YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS 511 ADUCTLFEAE ADICTLSEKE 561	321 EAKDVFLGIF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRSL PTLVEVSRNL 471 ERVTKCCTES DRVTKCCTES 521 KQIKKQSALV RQIKKQTALV 571	331 LYEYSRRHPE LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCCH 481 LVNRRPCFSA LVNRRPCFSA 531 ELIKHKPKAT ELVKHKPKAT 581	341 YSVSLLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK 541 EEQLKTVMGD KEQLKAVMDD
	301 DLPELAVDFV DLPSLAADFV 351 KEYENTLEKC KTYETTLEKC 401 YGFQNALLVR 451 DYLSVVLNRL DYLSVVLNQL 501 EFSAETFTFH EFNAETFTFH 551 F3SFVDKCCA	311 EKEVCKNY ESKDVCKNYA 361 CATDDPPACY CAAADPHECY 411 YTKKVPQVST 461 CVLHEKTPVS CVLHEKTPVS 511 ADLCTLEAE ADICTLEEAE ADICTLEEAE	321 EAKDVFLGIF EAKDVFLGMF 371 AVFDEFKPL AKVFDEFKPL 421 PTLVEVSRSL PTLVEVSRSL 9TLVEVSRNL 471 ERVTKCCTES DRVTKCCTES 521 SQIKKQSALV RQIKKQTALV 571 EGRKLVAAQ	331 LYEYSRRHPS LYEYARRHPD 381 VEEPHNLVKT VEEPQNLIKQ 431 GKVGSKCCTH GKVGSKCCCH 481 LVNRRPCFSA LVNRRPCFSA 531 ELIKHKPKAT ELVKHKPKAT 581 AAL	341 YSVSLLLRLA YSVVLLLRLA 391 NCELFEKLGE NCELFEQLGE 441 PEAERLSCAE PEAKRMPCAE 491 LQVDETYVPK LEVDETYVPK 541 EQLKIVMOD KEQLKAVMDD

Wavelength (Å)	1.0000
Resolution range (Å)	46.8-3.402 (3.524-3.402)
Space group	C 1 2 1
Cell dimensions	<i>a</i> = 106.908 Å, <i>b</i> = 49.145 Å, <i>c</i> = 123.869 Å
	$\alpha = 90^{\circ}, \beta = 110.171^{\circ}, \gamma = 90^{\circ}$
Total reflections	53018
Unique reflections	15776 (2644)
Multiplicity	3.4 (3.1)
Completeness (%)	96.8 (94.0)
Mean I/sigma (I)	3.72 (0.82)
B-factor ( $Å^2$ )	114.12
R-merge	0.252 (1.434)
Refinement	
Resolution range (Å)	46.8-3.2 (3.524-3.402)
No. reflections	8299 (788)
R-work	0.2859 (0.5042)
R-free	0.2967 (0.4432)
No. non-hydrogen atoms	4391
Macromolecules	4391
RMS (bonds)	0.002
RMS (angles)	0.49
Ramachandran favored (%)	93.46
Ramachandran outliers (%)	1.24
Clashscore	4.91
Average B-factor	115.07

Table S2 X-ray crystallography data collection and refinement statistics of rFSA

Statistics for the highest-resolution shell are shown in parentheses.

		$\lambda_{\max}$ (nm)	
Hemoproteins	oxy	deoxy	carbonyl
Hb-rFSA <sub>3</sub>	414, 541, 577	429, 556	419, 538, 569
Hb-FSA3	414, 541, 577	429, 556	419, 538, 569
Hb-HSA <sub>3</sub> <sup>a</sup>	413, 541, 577	430, 556	420, 538, 569
Hb <sup><i>a</i></sup>	414, 541, 577	430, 555	420, 538, 569
HbA <sup>b</sup>	415, 541, 577	430, 555	419, 540, 569

Table S3 UV-visible absorption spectral data of Hb-rFSA $_3$  and Hb-FSA $_3$  clusters in PBS solution (pH 7.4) at 25 °C

<sup>*a*</sup> From ref. 25. <sup>*b*</sup> Ref. 42.

### Reference

T. K. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, 32, W665–W667.