The Hemophore HasA from *Yersinia pestis* ($HasA_{yp}$) Coordinates Hemin with a Single Residue, Tyr75, and with Minimal Conformational Change

Ritesh Kumar[‡], Scott Lovell[£], Hirotoshi Matsumura[€], Kevin P. Battaile[¶], Pierre Moënne-Loccoz[€] and Mario Rivera[*,†]

[‡]Center for Bioinformatics, University of Kansas, 2030 Becker Drive, Lawrence, Kansas 66047.
[£]Del Shankel Structural Biology Center, University of Kansas, 2034 Becker Drive, Lawrence, Kansas 66047.
[¶]IMCA-CAT, Hauptman Woodward Medical Research Institute, 9700 S. Cass Avenue, Bldg. 435A, Argonne, Illinois 60439.
[€] Division, of Environmental & Biomolecular Systems, Institute of Environmental Health, Oregon Health and Science University, 20000 Northwest Walker Road, Beaverton, Oregon 97006-8921.
[†]Department of Chemistry, University of Kansas, Multidisciplinary Research Building, 2030 Becker Drive, Lawrence, Kansas 66047.

Supporting Information

Contents

Experimental Methods

1. Cloning of the $HasA_{yp}$ Gene. A gene coding for full-length *Yersinia pestis* HasA ($HasA_{yp}$) was synthesized and sub-cloned into the pET11a expression vector by GenScript (Piscataway, NJ). Silent mutations to introduce codons favored by *Escherichia coli* were engineered into the $HasA_{yp}$ gene (*1*). Restriction sites *Nde*I and *Bam*HI were introduced at the 5' and 3' ends of the gene, respectively (Figure S7), to facilitate sub-cloning into the pET11a vector. The recombinant plasmid harboring the $HasA_{yp}$ gene was then transformed into the *E. coli* BL21-GOLD (DE3) competent cells (Stratagene, La Jolla, CA). A gene coding for truncated $HasA_{yp}$ (full-length – 12 residues) was constructed by appropriately placing a stop codon in the recombinant pET11a plasmid harboring the gene coding for full-length $HasA_{yp}$ using the QuickChange mutagenesis kit from Stratagene (La Jolla, CA). The oligonucleotides were synthesized by Integrated DNA Technologies, Inc. and were used without further purification. The primers used were 5'-GCCCCGATGGTTGATACGGTTTAAGTTGTTGATTGTCACGATATG-3' and 5'-CATATCGTGACAATCAACAACTTAAACCGTATCAACCATCGGGGC-3'; the underlined codons represent mismatches that introduce a stop codon after Val 193 (see Figure S7). The mutation was confirmed by sequencing (ACGT, Inc., IL) and the recombinant DNA plasmid was transformed into *Escherichia coli* BL21-GOLD (DE3) competent cells for subsequent protein expression.

2. Protein Expression and Purification. $HasA_{yp}$ was expressed using previously described methods and was isolated using a protocol reported for the purification of apo-$HasA_p$ (*2*) with a few modifications: supernatant from cell lysis was loaded onto a Q-sepharose Fast Flow column (2.6 cm i.d. x 15 cm, GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 7.6) at 4 °C. The column was then washed with 2 column volumes of the same buffer and the protein eluted

with the same buffer using a linear NaCl gradient (0-400 mM). Fractions containing the $HasA_{yp}$ protein, as determined by SDS-PAGE, were pooled and dialyzed against 50 mM sodium phosphate buffer (pH 7.0) containing 900 mM ammonium sulfate before loading onto a Phenyl Sepharose 6 Fast Flow (GE Healthcare) column (2.6 cm i.d. x 12 cm) pre-equilibrated with the same buffer at 25 °C. The protein was eluted in stepwise manner: (a) 50 mM sodium phosphate/750 mM ammonium sulfate (pH 7.0) was used to elute weakly bound proteins, including holo-$HasA_{yp}$ and (b) a linear gradient of sodium phosphate (50 - 20 mM)/ammonium sulfate (750-0 mM), pH 7.0, was used to elute apo-$HasA_{yp}$. Fractions containing apoprotein were pooled and concentrated to a volume of ~ 1 mL using 10 kDa molecular weight cut-off (MWCO) Amicon ultracentrifuge filters (Millipore, MA) before loading onto a Sephadex G-75 (2.6 cm i.d. x 90 cm, GE Healthcare) size exclusion column pre-equilibrated and eluted with sodium phosphate (μ = 0.1, pH 7.8) buffer.

Apo-$HasA_{yp}$ purified to homogeneity was analyzed by electrospray ionization mass spectrometry. The mass spectra indicated that the full-length protein is cleaved at the C-terminus during purification, with the shortest protein corresponding to full-length – 12 residues, not including the initiator Met (Figure S8-A). A stop codon was introduced after Val-193 (see above) to express truncated apo-$HasA_{yp}$ consisting of full-length – 12 residues (FL-12). The mass spectrum (Figure S8-B) shows that FL-12, devoid of the initial initiator Met, can be purified to homogeneity. To prepare holo-$HasA_{yp}$, a solution of apo-$HasA_{yp}$ (FL-12) in sodium phosphate buffer (μ = 0.1, pH 7.8) was titrated with a solution of hemin (1 mM) in dimethyl sulfoxide. The resultant solution was incubated overnight at 4 °C, concentrated to ~ 1 mL using 10 kDa MWCO Amicon ultracentrifuge filters and loaded onto a Sephadex G-75 (2.6 cm i.d., x 90 cm) size exclusion column pre-equilibrated and eluted with sodium phosphate (μ = 0.1, pH 7.8) buffer. The homogeneity of the protein was assessed by SDS-PAGE and mass spectrometry.

The UV-Vis spectrum of $HasA_{yp}$ (Figure S9) exhibits a Soret maximum at 403 nm (ε_{403} = 116.8 mM^{-1} cm^{-1}) and bands at 498, 535 and 620 nm.

3. Crystallization. The experiments were carried out with apo- and holo-$HasA_{yp}$ (FL-12) using the sitting drop vapor diffusion method and Compact Jr. Crystallization plates (Emerald BioSystems). Equal volumes (0.75 μL) of protein and crystallization solution were mixed and equilibrated against 75 μL of reservoir solution.

Apo-$HasA_{yp}$. Crystals showing tetragonal morphology were obtained within one week from a solution of apo-$HasA_{yp}$ (63 mg/mL) in sodium phosphate buffer (μ = 0.1, pH 7.8) and the Index HT screen (Hampton Research) condition H2 (0.2 M potassium sodium tartarate tetrahydrate, 20% (w/v) PEG 3350). These crystals were cryo-protected by transferring to a solution containing 80% crystallization solution and 20% (v/v) PEG 400 for approximately 30 s before freezing in liquid nitrogen. Crystals with hexagonal morphology were obtained upon mixing the same protein solution with the Index HT screen (Hampton Research) condition C9 (1.1 M sodium malonate pH 7.0, 0.1 HEPES pH 7.0 and 0.5% (v/v) Jeffamine ED-2001 pH 7.0). These crystals were transferred to a solution containing 80% crystallization solution and 20% (v/v) PEG 400 for approximately 30 s before freezing in liquid nitrogen.

Holo-$HasA_{yp}$. A cluster of thin needles grew within 3-4 days from 32 mg/mL holo-$HasA_{yp}$ in sodium phosphate buffer (μ = 0.1, pH 7.8) and the Wizard III screen (Emerald BioSystems) condition 3 (20% (w/v) PEG 3350, 200 mM magnesium formate dihydrate). These crystals diffracted to low resolution. Therefore, refinement of the crystallization conditions was carried out using the pHat Buffer (Emerald Biosystems) screen. Thick, needle like crystals were obtained within 1-2 days using pHat Buffer Screen condition C10 (0.5 M CHES/ NaOH, pH 8.8). The crystals were cryo-protected by transferring to a solution containing 80%

crystallization solution and 20% DMSO for approximately 30 s prior to freezing in liquid nitrogen.

4. X-ray Data Collection, Structure Solution and Refinement. Data for apo-$HasA_{yp}^{tet}$, apo-$HasA_{yp}^{hex}$ and holo-$HasA_{yp}$ were collected at 100 K at the Advanced Photon Source (APS) IMCA-CAT, beamline 17-ID using a Dectris Pilatus 6M pixel array detector. The diffraction data were integrated with XDS (*3*) via the XDSAPP (*4*) interface and the Laue class and data scaling were performed with Aimless (*5*). Structure solution of the apo-protein was obtained by molecular replacement with Phaser (*6*) via the Phenix (*7*) interface using the holo-$HasA_{sm}$ structure (PDB: 1DK0) as a search model. Automated model building and structure refinement was conducted using Phenix and manual model building was carried out with Coot (*8*). The final apo-$HasA_{yp}^{tet}$ structure was used as the search model for subsequent molecular replacement searches for the apo-$HasA_{yp}^{hex}$ and holo-$HasA_{yp}$ data. Crystals of apo-$HasA_{yp}^{tet}$ were indexed in a tetragonal *P* lattice with the most probable Laue class 4/*mmm* and space groups $P4_12_12$ or $P4_32_12$. Molecular replacement searches were conducted in all space groups with 422 point symmetry and the top solution was obtained in $P4_12_12$ with one molecule in the asymmetric unit.The model was refined with anisotropic atomic displacement parameters for all atoms. A sodium ion was modeled near the C-terminus which is coordinated to the backbone carbonyls of L172, A175 and F169. For apo-$HasA_{yp}^{hex}$ the crystals were indexed in a hexagonal lattice *P* with the most probable Laue class 6/*m* and space group $P6_3$. Indeed, the top molecular replacement solution was obtained in $P6_3$ although space groups with 6 point symmetry were tested. The electron density maps contained features consistent with a tris molecule (from the buffer) on a crystallographic 3-fold as well as four malonate and three polyethylene glycol (PEG) molecules (from crystallization solution). TLS refinement (*9*) was incorporated in the latter stages of structure refinement. Holo-$HasA_{yp}$ crystals were indexed in a *C*-centered lattice (space group *C*2). Molecular replacement was conducted in this space group searching for 10 molecules in the

asymmetric unit which were positioned by Phaser. Following subsequent rounds of refinement, large difference electron density peaks (F_o-F_c) greater than 3σ were observed on the distal sites of the heme molecules. These sites were originally assigned as water molecules, however residual positive electron density above the 3σ contour level was observed around these water molecules following refinement. Chloride ions were ultimately assigned at these sites based on the electron density, the coordination environment (positively charged Arg residues) and expected non-bonded distances to the heme Fe-atoms (*10*). The average distance between the Cl^- and heme-Fe atoms is 2.82 Å. Table S1 summarizes the crystallographic details for the three refined structures.

5. Spontaneous cleaving of the C-terminus in FL-12. Mass spectral analysis of apo-FL-12 stored at -20 °C for several months showed that FL-12 had been cleaved predominantly at Ser183, producing a protein that corresponds to full-length minus 21 residues (FL-21). Additional investigation showed that apo-FL-12 is nearly quantitatively converted to apo-FL-21 upon incubation at 4 °C for a week. In comparison, holo-FL-12 is stable to C-terminal cleavage even if the protein is incubated at 4 °C for two weeks. These observations, which indicate that the apo-FL-12 protein has a propensity to undergo relatively slow additional cleavage of its C terminal domain to produce apo-FL-21, prompted us to carry out mass spectrometric analysis of protein in the single crystals obtained from the wells that produced the samples used for X-ray diffraction. Consistent with the observations in solution, it was found that apo-FL-12 obtained from the single crystals was cleaved at Ser185 (FL-21), whereas holo-FL-12 obtained from the corresponding single crystals remained as the FL-12 protein. These observations are in agreement with reports indicating that $HasA_p$ secreted to the extracellular environment undergo proteolytic cleavage of their C-termini that remove the last 15-21 amino acids (*11*).

6. Spectroscopic and kinetic analyses of $HasA_{yp}$. EPR spectra of holo-$HasA_{yp}$ were measured on a Bruker E500 X-band EPR spectrometer with a dual mode cavity and a helium-flow cryostat (ESR 900, Oxford Instruments, Inc.). Resonance Raman (RR) spectra were recorded with a 413 nm excitation from a krypton laser (Innova 302, Coherent) using a McPherson 2061/207 spectrograph and a Kaiser Optic supernotch filter to attenuate the Rayleigh scattering. The samples were kept at 110 K with a liquid-nitrogen coldfinger and probed in a backscattering geometry. Frequency calibration was executed with aspirin and is accurate to ± 1cm^{-1}. Although fluoride binding to holo-$HasA_{yp}$ is supported by subtle but significant changes in the electronic absorption spectra, attempts to detect iron(III)-chloride, iron(III)-fluoride, or iron(III)-OH stretching frequencies in the RR spectra were unsuccessful. Stopped-flow experiments were performed as previously (*12*), with an SX20 apparatus (Applied Photophysics) equilibrated at 4 °C. Hemin was dissolved in 10 mM NaOH while the apoprotein was diluted in 200 mM HEPES pH 7.8; excess solutions were recovered from the stopped-flow to confirm the concentration of the reactants.

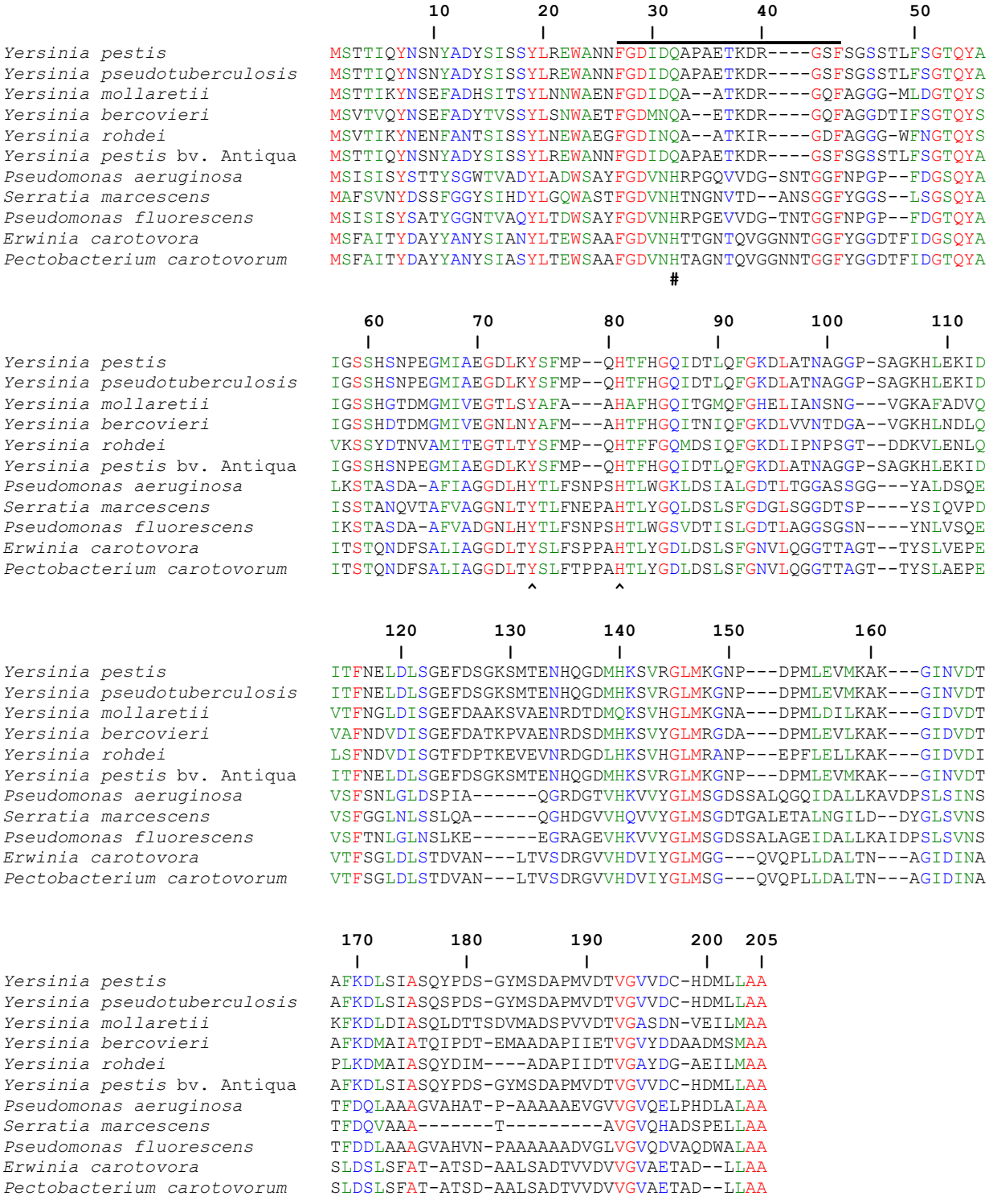


Figure S1. Full-length HasA sequences aligned against the sequence of HasA from *Yersinia pestis*. Residue 32 (H or Q) is highlighted by (#) and proximal loop residues Tyr75 and His83 are highlighted by (^). A horizontal line above the sequence highlights residues comprising the distal loop. Conserved residues across the alignment are in red, conservative substitutions in green and semi-conservative substitutions in blue; sequence numbering is according to the *Yersinia pestis* sequence.

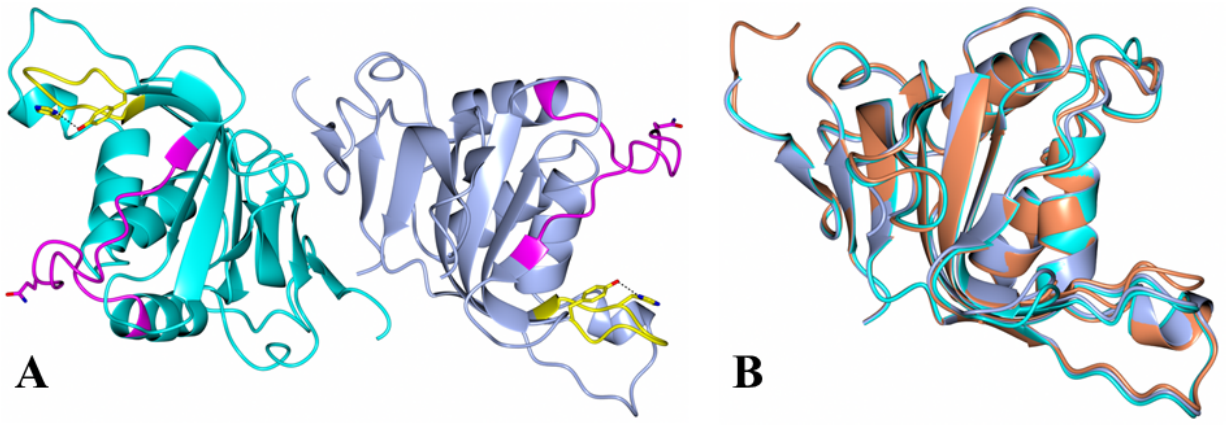


Figure S2. (A) Crystal structure of apo-$HasA_{yp}^{hex}$. Molecule A is shown in light blue and molecule B in cyan. In both molecules the Tyr75 loop is in yellow with Tyr75 and His81 rendered in sticks and the Gln32 loop is in magenta with Gln32 in sticks. (**B**) Superimposed crystal structures of apo-$HasA_{yp}^{tet}$ (coral), apo-$HasA_{yp}^{hex}$ chain A (blue) and apo-$HasA_{yp}^{hex}$ chain B (cyan). Residues Ser2 to Asp180 from the three structures were used to calculate RMSD values using the secondary structure matching algorithm in the program Superpose (*13*) via the CCP4 (*14*) interface.

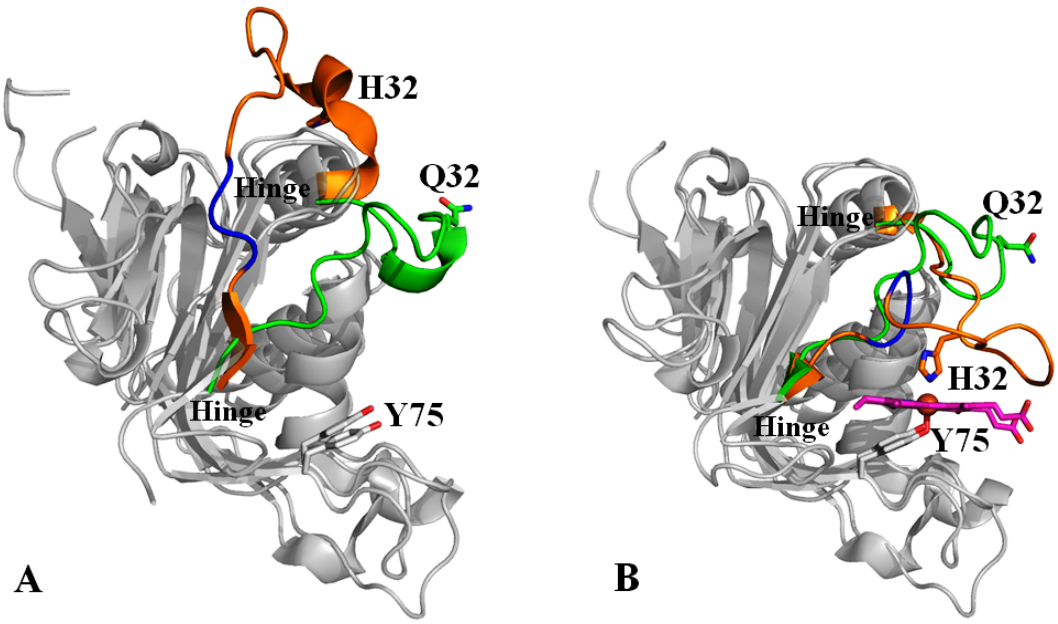
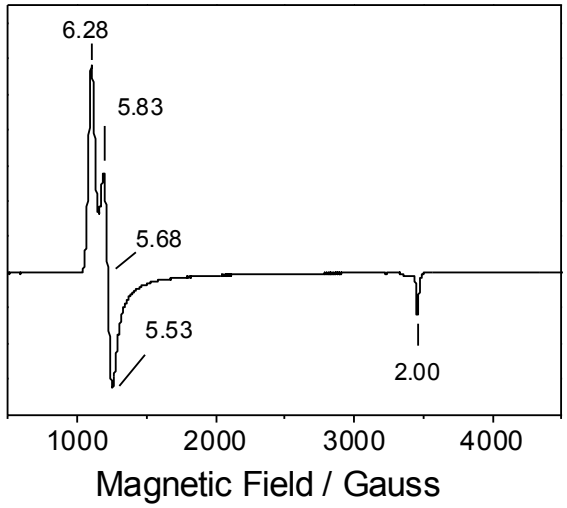
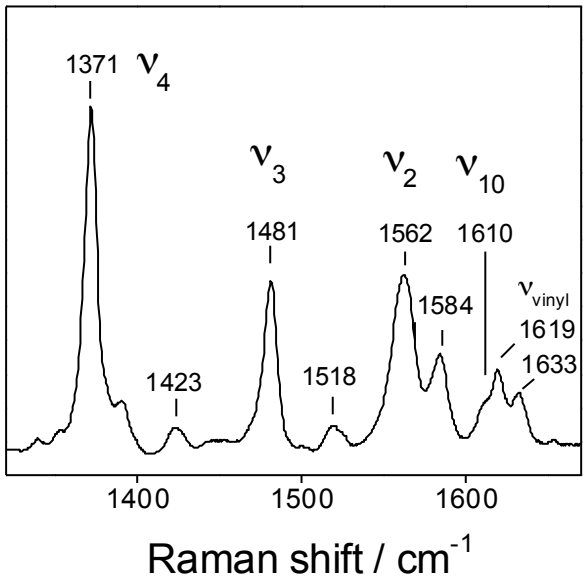


Figure S3. Superimposed crystal structures of: (A) apo-$HasA_p$ (PDB: 3MOK) and apo-$HasA_{yp.}$ (B) holo-$HasA_p$ (PDB: 3ELL) and holo-$HasA_{yp}$. The Q32 loop of $HasA_{yp}$ is shown in green, and the H32 loop of $HasA_p$ in orange, with the three extra residues ($S^{41}N^{42}T^{43}$) highlighted in blue.

Figure S4:



X-band EPR spectrum of holo-HasA_{yp} at 300 μM. Experimental conditions: 1 mW microwave power, 10 G modulation amplitude, sample temperature 6.7 K; *g* values of rhombic high-spin ferric heme signals are given.



RR spectrum of holo-HasA_{yp} at 300 μM. Experimental conditions: 413 nm excitation (20 mW), 7 cm^{-1} resolution, sample temperature 110 K; prominent porphyrin skeletal modes supporting the 6-coordinate high-spin configuration and the two vinyl stretching modes are labeled.

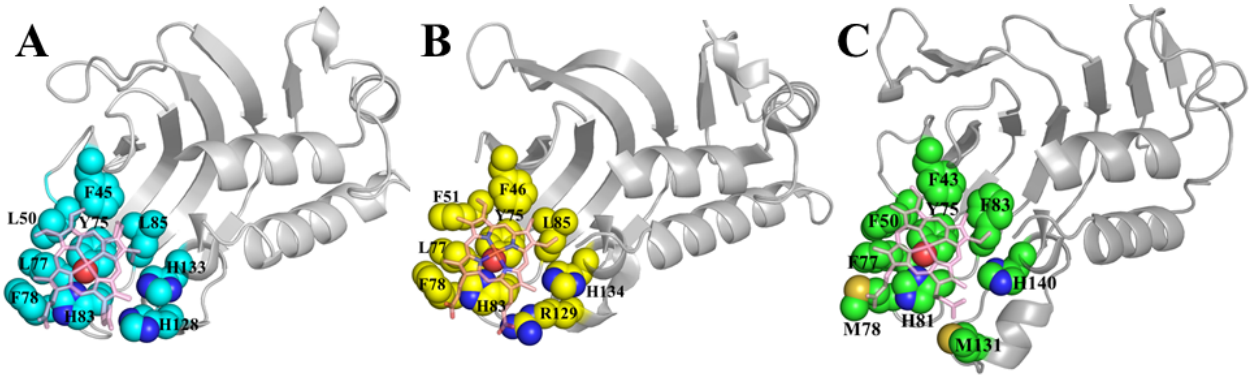


Figure S5. Top view of the hemin-binding pocket. Residues that form a hydrophobic (sticky) platform for macrocycle capture onto the Y75 loop and residues that interact with hemin propionates are highlighted in sphere rendering. (A) holo-$HasA_{sm}$ (PDB: 1B2Y) (B) holo-$HasA_{p}$ (PDB: 3ELL) and holo-$HasA_{yp}$.

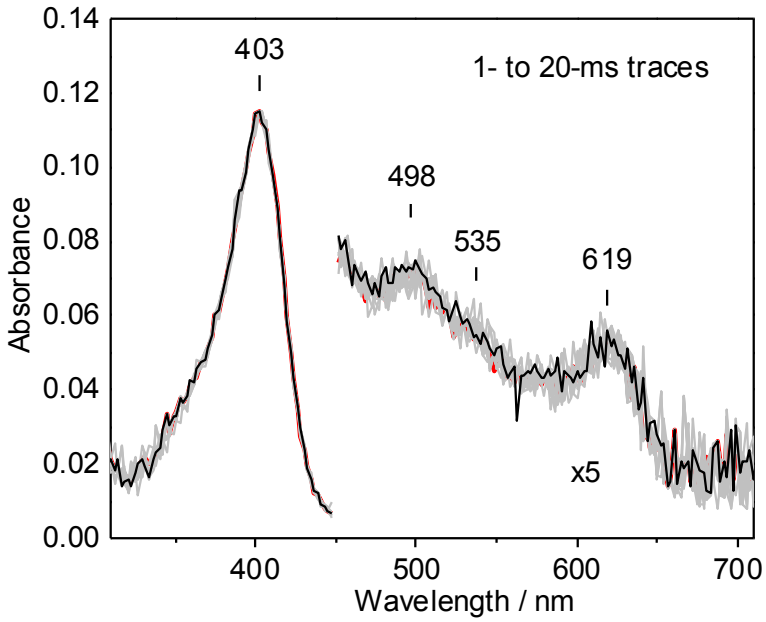


Figure S6: Stopped-flow absorption spectra obtained upon reacting 5 μM hemin with 30 μM apo-$HasA_{yp}$ at 4 °C (red trace, 1 ms; grey traces, 2 to 19 ms; back trace, 20 ms). The capture of hemin and formation of holo-$HasA_{yp}$ is completed within the millisecond mixing time of the instrument.

***Nde* I**

5'-**CATATG** TCT ACC ACC ATC CAA TAC AAC AGC AAC TAT GCG GAC TAC TCT ATC TCC TCG
M S T T I Q Y N S N Y A D Y S I S S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

TAC CTG CGT GAA TGG GCA AAC AAC TTC GGC GAT ATT GAT CAG GCG CCG GCC GAA ACG
Y L R E W A N N F G D I D Q A P A E T
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37

AAA GAC CGT GGC TCA TTT TCG GGT AGC TCT ACC CTG TTC AGC GGC ACG CAA TAT GCA
K D R G S F S G S S T L F S G T Q Y A
38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

ATT GGT AGT TCC CAT TCT AAC CCG GAA GGC ATG ATC GCT GAA GGT GAT CTG AAA TAC
I G S S H S N P E G M I A E G D L K Y
57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

TCT TTT ATG CCG CAG CAT ACC TTC CAC GGC CAG ATC GAT ACG CTG CAA TTT GGT AAA
S F M P Q H T F H G Q I D T L Q F G K
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94

GAC CTG GCA ACC AAT GCT GGC GGT CCG AGT GCA GGC AAA CAC CTG GAA AAA ATT GAT
D L A T N A G G P S A G K H L E K I D
95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113

ATC ACG TTT AAC GAA CTG GAT CTG AGC GGC GAA TTC GAC AGC GGT AAA TCT ATG ACC
I T F N E L D L S G E F D S G K S M T
114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132

GAA AAT CAT CAG GGT GAT ATG CAC AAA AGC GTC CGC GGC CTG ATG AAA GGT AAC CCG
E N H Q G D M H K S V R G L M K G N P
133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151

GAC CCG ATG CTG GAA GTG ATG AAA GCC AAA GGC ATT AAT GTT GAT ACC GCG TTC AAA
D P M L E V M K A K G I N V D T A F K
152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170

GAC CTG AGT ATC GCC TCC CAA TAT CCG GAT TCA GGT TAC ATG TCG GAC GCC CCG ATG
D L S I A S Q Y P D S G Y M S D A P M
171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189

***Bam* HI**

GTT GAT ACG GTT **GGT** GTT GTT GAT TGT CAC GAT ATG CTG CTG GCT GCC TAA **GGATCC**-3'
V D T V **G** V V D C H D M L L A A -
190 191 192 193 **194** 195 196 197 198 199 200 201 202 203 204 205
^

Figure S7. DNA and corresponding amino acid sequence of full-length $HasA_{yp}$. *Nde*I and *Bam*HI restriction endonuclease sites were introduced at the 5' and 3' end for sub-cloning. The carrot symbol (^) indicates the place where a stop codon was introduced to express full length minus 12 residues (FL – 12) $HasA_{yp}$.

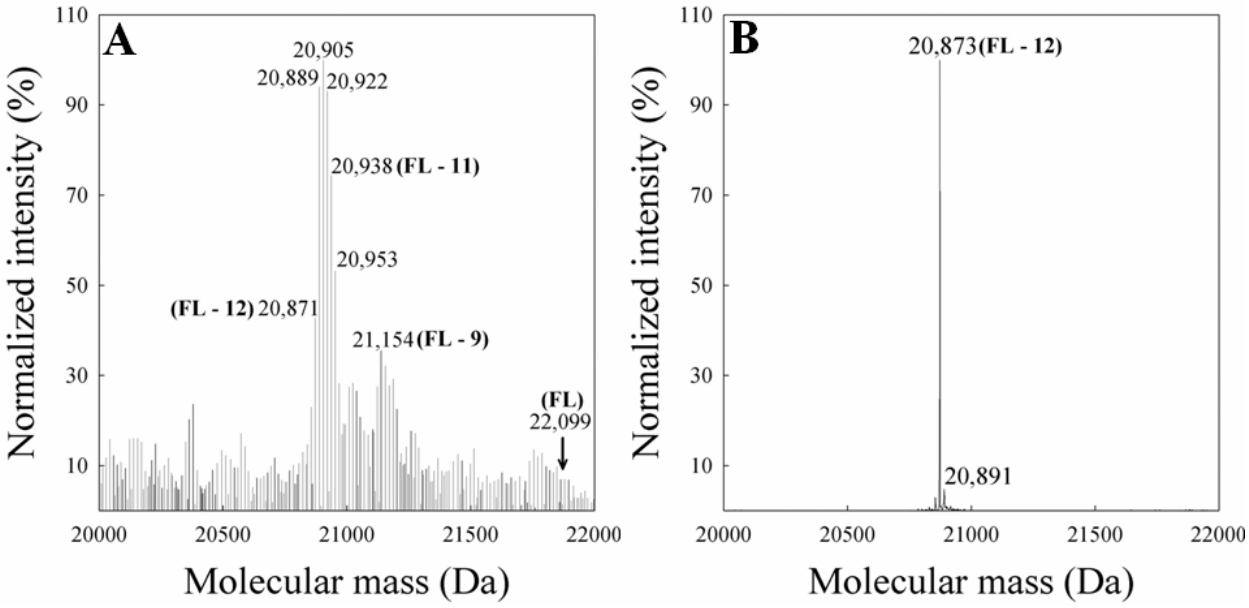


Figure S8. (A) Mass spectroscopic analysis of protein obtained upon expression and purification of full-length apo-$HasA_{yp}$ (expected MW = 22,099 Da) indicated that the protein is cleaved at the C-terminus, with the shortest protein corresponding to full-length – 12 (FL-12) residues, not including the initiator methionine. (B) Mass spectroscopic analysis of protein obtained upon expression and purification of FL-12 apo-$HasA_{yp}$ showed that the truncated protein devoid of the initiator methionine (expected MW = 20,874 Da) can be purified to homogeneity (experimental MW = 20,873).

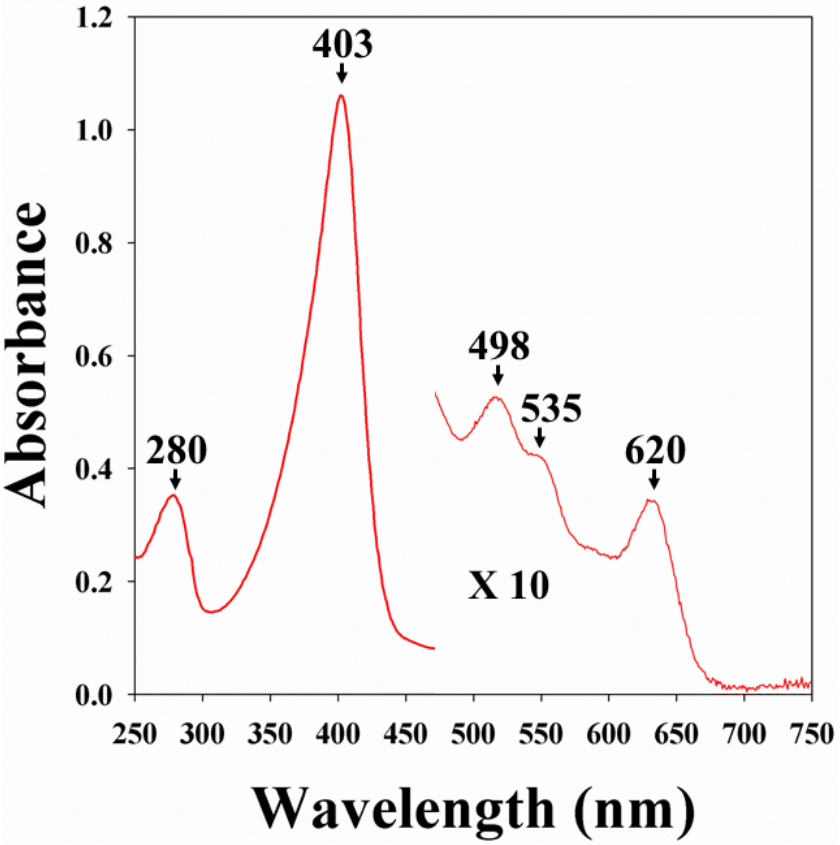


Figure S9. Electronic absorption spectrum of heme-reconstituted holo-$HasA_{yp}$.

Table S1. Crystallographic data for $HasA_{yp}$ structures

	Apo-$HasA_{yp}^{tet}$	**Apo-$HasA_{yp}^{hex}$**	**Holo-$HasA_{yp}$**
Data Collection			
Unit-cell parameters (Å, °)	*a*=70.34, *c*=74.84	*a*=75.07, *c*=129.30	*a*=139.40, *b*=69.56, *c*=232.30, *β*=100.1
Space group	$P4_12_12$	$P6_3$	$C2$
Resolution (Å)[1]	41.42-1.10 (1.12-1.10)	45.83-1.60 (1.63-1.60)	47.05-2.20 (2.24-2.20)
Wavelength (Å)	1.0000	1.0000	1.0000
Temperature (K)	100	100	100
Observed reflections	638,353	552,152	379,425
Unique reflections	76,552	54,313	109,337
$\langle I/\sigma(I)\rangle$[1]	21.8 (2.9)	18.2 (2.1)	9.9 (2.0)
Completeness (%)[1]	100 (100)	100 (100)	98.2 (97.7)
Multiplicity[1]	8.3 (8.2)	10.2 (10.3)	3.5 (3.6)
R_{merge} (%)[1, 2]	4.2 (75.5)	7.6 (1.257)	8.7 (66.8)
R_{meas} (%)[1, 4]	4.5 (80.5)	8.1 (1.323)	10.3 (78.4)
R_{pim} (%)[1, 4]	1.5 (27.6)	2.5 (41.2)	5.5 (40.8)
$CC_{1/2}$ [1, 5]	1.000 (0.802)	0.999 (0.688)	0.996 (0.686)
Refinement			
Resolution (Å)	33.04-1.10	37.54-1.60	38.62-2.20
Reflections (working/test)	72,634 / 3,844	51,507 / 2,761	103,819 / 5,497
R_{factor} / R_{free} (%)[3]	13.9 / 14.6	16.1 / 19.0	16.6 / 22.3
No. of atoms (Protein/Heme/Water)	1,458/-/220	2,754 / - / 259	13,624 / 430 / 720
Model Quality			
R.m.s deviations			
Bond lengths (Å)	0.010	0.009	0.012
Bond angles (°)	1.214	1.046	1.158
Average *B*-factor (Å^2)			
All Atoms	16.9	25.9	31.8
Protein	15.2	25.1	31.6
Heme	-	-	35.4
Water	28.5	32.4	33.7
Coordinate error, maximum likelihood (Å)	0.08	0.16	0.29
Ramachandran Plot			
Most favored (%)	98.45	98.35	98.08
Additionally allowed (%)	1.55	1.38	1.92

1) Values in parenthesis are for the highest resolution shell.
2) $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i*th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.
3) $R_{factor} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$; Rfree is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.
4) R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}(*5, 15*). R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}(*16, 17*).
5) $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data (*18, 19*).

REFERENCES

1. Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms, *Mol Biol Evol 2*, 13-34.
2. Jepkorir, G., Rodriguez, J. C., Rui, H., Im, W., Lovell, S., Battaile, K. P., Alontaga, A. Y., Yukl, E. T., Moenne-Loccoz, P., and Rivera, M. (2010) Structural, NMR spectroscopic, and computational investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa, *J Am Chem Soc 132*, 9857-9872.
3. Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.
4. Krug, M., Weiss, M. S., Heinemann, U., and Mueller, U. (2012) XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS, *Journal of Applied Crystallography 45*, 568-572.
5. Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics, *Acta Crystallogr D Biol Crystallogr 67*, 282-292.
6. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *Phaser* crystallographic software, *J. Appl. Cryst. 40*, 658-674.
7. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.
8. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.
9. Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion, *Acta Crystallogr D Biol Crystallogr 62*, 439-450.
10. Wong, S. G., Abdulqadir, R., Le Brun, N. E., Moore, G. R., and Mauk, A. G. (2012) Fe-Haem Bound to Escherichia coli Bacterioferritin Accelerates Iron Core Formation by an Electron Transfer Mechanism, *Biochm. J. 444*, 553-560.
11. Arevalo-Ferro, C., Hentzer, M., Reil, G., Gorg, A., Kjelleberg, S., Givskov, M., Riedel, K., and Eberl, L. (2003) Identification of Quorum-Sensing Regulated Proteins in the Opportunistic Pathogen *Pseudomonas aeruginosa* by Proteomics, *Environ. Microbiol. 5*, 1350-1369.
12. Yukl, E. T., Jepkorir, G., Alontaga, A. Y., Pautsch, L., Rodriguez, J. C., Rivera, M., and Moënne-Loccoz, P. (2010) Kinetic and Spectroscopic Studies of Hemin Acquisition in the Hemophore HasAp from Pseudomonas aeruginosa, *Biochemistry 49*, 6646-6654.
13. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr D Biol Crystallogr 60*, 2256-2268.
14. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D Biol Crystallogr 67*, 235-242.
15. Evans, P. (2006) Scaling and assessment of data quality, *Acta Crystallogr D Biol Crystallogr 62*, 72-82.
16. Diederichs, K., and Karplus, P. A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography, *Nat Struct Biol 4*, 269-275.

17. Weiss, M. S. (2001) Global indicators of x-ray data quality, *Journal of Applied Crystallography 34*, 130-135.
18. Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality, *Science 336*, 1030-1033.
19. Evans, P. (2012) Biochemistry. Resolving some old problems in protein crystallography, *Science 336*, 986-987.