Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Supporting Information

Corrole–Protein Interactions in H-NOX and HasA

Christopher M. Lemon,[1,2,3,*] Amos J. Nissley,[4] Naomi R. Latorraca,[1,3] Elizabeth C. Wittenborn,[2] and Michael A. Marletta[1,2,4,*]

[1] Department of Molecular and Cell Biology, [2] California Institute for Quantitative Biosciences (QB3), [3] Miller Institute for Basic Research in Science, and [4] Department of Chemistry University of California, Berkeley, Berkeley, California, 94720, United States

clemon@berkeley.edu
marletta@berkeley.edu

Materials

The following chemicals were used as received: potassium phosphate dibasic trihydrate (K_2HPO_4), sodium acetate, boric acid, deoxyribonuclease I from bovine pancreas (DNase), D-(+)-glucose, ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA), 2′,7′-dichlorofluorescein (fluorescein 27), and antifoam SE-15 from Sigma-Aldrich; Luria Broth (LB), Terrific Broth (TB), ampicillin, tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF), and lysozyme (from egg white) from Research Products International; agarose, dimethyl sulfoxide (DMSO), BugBuster Master Mix, ammonium bicarbonate, and Immobilon Classico Western HRP substrate from EMD Millipore; hemin chloride, Pd(II) mesoporphyrin IX [Pd(MPIX)], and Zn(II) protoporphyrin IX [Zn(PPIX)] from Frontier Scientific; PrimeSTAR Max DNA polymerase and His60 Ni Superflow resin from Takara Bio; mouse anti-His_6 antibody and goat anti-mouse HRP conjugate antibody from Life Technologies; Bacto Agar from BD; benzamidine from Accela; triethanolamine hydrochloride (TEA) from Spectrum Chemical; imidazole from Oakwood Chemical; Gibson Assembly Master Mix from New England BioLabs; fluorescein-5-maleimide from Molecular Probes; nitrogen gas (N_2) from Praxair; isopropyl β-D-1-thiogalactopyranoside (IPTG), sodium chloride (NaCl), and glycerol and from Fisher Scientific; and PD-10 desalting columns (Sephadex G-25 medium) from GE Healthcare. Argon gas (Ar) from Praxair was passed over an Oxiclear gas purifier prior to use. The DH5α, XL1-Blue, and BL21(DE3) CodonPlus RIL *E. coli* competent cells, as well as Gibson-like Master Mix, were obtained from the UC Berkeley Macrolab. The RP523(DE3) competent cells were obtained from lab-generated stocks.[1] The corrole 8,12-bis(2-carboxyethyl)-2,3,7,13,17,18-hexamethylcorrolato(oxo)-phosphorus(V) (**P-1**) was prepared according to literature methods.[2]

WT *Cs* H-NOX Sequence with a C-Terminal His_6 Tag

MKGTIVGTWIKTLRDLYGNDVVDESLKSVGWEPDRVITPLEDIDDDEVRRIFAKVSEKTGKNVNEIWR
EVGRQNIKTFSEWFPSYFAGRRLVNFLMMMDEVHLQLTKMIKGATPPRLIAKPVAKDAIEMEYVSKR
KMYDYFLGLIEGSSKFFKEEISVEEVERGEKDGFSRLKVRIKFKNPVFEYKKNGAAALEHHHHHH

(1) M. A. Herzik, R. Jonnalagadda, J. Kuriyan and M. A. Marletta, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, E4156–E4164.
(2) C. M. Lemon and M. A. Marletta, *Inorg. Chem.*, 2021, **60**, 2716–2729.

WT *Pa* HasA Sequence with an N-Terminal His_6 Tag and Truncation of the Secretion Signal Peptide

MHHHHHHENLYFQGSISISYSTTYSGWTVADYLADWSAYFGDVNHRPGQVVDGSNTGGFNPGPFDGS
QYALKSTASDAAFIAGGDLHYTLFSNPSHTLWGKLDSIALGDTLTGGASSGGYALDSQEVSFSNLGLDS
PIAQGRDGTVHKVVYGLMSGDSSALQGQIDALLKAVDPSLSINSTFDQLAAAGVAHATPAA

Preparation of Expression Plasmids

The pET20b plasmid containing the WT *Cs* H-NOX sequence with a C-terminal His_6 tag[3] and the pCW plasmid containing the WT *Pa* HasA sequence with an N-terminal His_6 tag[2] have been previously reported. Point mutations were generated by amplifying gene fragments containing the desired mutation using PCR and then combined using Gibson assembly. PCR reactions (35 cycles) were performed on a 50 μL scale using PrimeSTAR Max DNA polymerase with a denaturation temperature of 98 °C, an annealing temperature of 55 °C, and an extension/elongation temperature of 72 °C. Each reaction contained 2.5 μL of the primers (10 μM stock solutions), 1 μL of the plasmid template (~30 pM stock solution), 19 μL of deionised H_2O, and 25 μL of PrimeSTAR Max master mix. Reactions were purified on a 1% agarose gel, the desired bands were excised, and the DNA was purified using the Qiagen gel extraction kit. Cloning reactions were performed using NEB Gibson Assembly Master Mix and run on a 20 μL scale, following the manufacturer's protocol. After performing the Gibson reaction at 50 °C for 30 min, 10 μL of the reaction mixture was transformed (heat shock at 42 °C for 42 s) into DH5α, XL1-Blue, or DH10B-$T1^R$ *E. coli* competent cells and then plated on an LB agar plate (1.5% (w/v) agar) supplemented with ampicillin (100 μg/mL). Single colonies were selected and grown in 5 mL of LB supplemented with ampicillin (100 μg/mL) at 37 °C overnight. The plasmid was extracted using the Qiagen mini-prep kit and the sequence was validated (UC Berkeley DNA Sequencing Facility).

Primers for PCR

Linearised pET20b Vector

FWD: 5′–GTATAAGAAAAATGGAGCGGCCGCAC–3′
REV: 5′–GTCCCGACGATTGTCCCCTTCATATG–3′

Linearised pCW Vector

FWD: 5′–CGATGATAAGCTGTCAAACATGAGCAG– 3′
REV: 5′–ATGACCTCCTAAGCATCGATGGATCC–3′

(3) C. W. Hespen, J. J. Bruegger, C. M. Phillips-Piro and M. A. Marletta, *ACS Chem. Biol.*, 2016, **11**, 2337–2346.

WT *Cs* H-NOX Gene for pET20b Vector

FWD: 5′–CATATGAAGGGGACAATCGTCGGGAC–3′

REV: 5′–GTGCGGCCGCTCCATTTTTCTTATAC–3′

WT *Cs* H-NOX Gene for pCW Vector

FWD: 5′–GGATCCATCGATGCTTAGGAGGTCATATGAAGGGGACAATCGTCGGGAC–3′

REV: 5′–CTGCTCATGTTTGACAGCTTATCATCGTCAGTGGTGGTGGTGGTGGTGC–3′

W9F *Cs* H-NOX

FWD: 5′–CGTCGGGACATTTATAAAGACCCTGAGG–3′

REV: 5′–CCTCAGGGTCTTTATAAATGTCCCGACG–3′

H102G *Cs* H-NOX

FWD: 5′–GATGATGGATGAGGTAGGCCTACAGCTTACC–3′

REV: 5′–GGTAAGCTGTAGGCCTACCTCATCCATCATC–3′

H102S *Cs* H-NOX

FWD: 5′–GATGATGGATGAGGTATCCCTACAGCTTACC–3′

REV: 5′–GGTAAGCTGTAGGGATACCTCATCCATCATC–3′

P115A *Cs* H-NOX

FWD: 5′–GGAGCCACTCCTGCGAGGCTTATTG–3′

REV: 5′–CAATAAGCCTCGCAGGAGTGGCTCC–3′

Y140F *Cs* H-NOX

FWD: 5′–GAAAGATGTACGATTTTTTTTTAGGGC–3′

REV: 5′–GCCCTAAAAAAAAATCGTACATCTTTC–3′

WT *Pa* HasA for pCW Vector

FWD: 5′–GGATCCATCGATGCTTAGGAGGTCATATGCATCACCATCACCATCACG-AAAACCTGTACTTCCAAGGGAGCATTTCGATTTCCTACAGCACCAC–3′

REV: 5′–CTGCTCATGTTTGACAGCTTATCATCGTTACGCAGCCGGGGTGGCAT–3′

H32C *Pa* HasA

FWD: 5′–TACTTCGGCGACGTCAACTGTCGTCCG–3′

REV: 5′–CGGACGACAGTTGACGTCGCCGAAGTA–3′

H32S *Pa* HasA

FWD: 5′–TACTTCGGCGACGTCAACTCTCGTCCG–3′

REV: 5′–CGGACGAGAGTTGACGTCGCCGAAGTA–3′

V38C *Pa* HasA

FWD: 5′–CAGGTGTGCGACGGCAGCAATAC–3′

REV: 5′–GTATTGCTGCCGTCGCACACCTG–3′

Y75F *Pa* HasA

FWD: 5′–CGACCTCCACTTCACCCTGTTC–3′

REV: 5′–GAACAGGGTGAAGTGGAGGTCG–3′

S82C *Pa* HasA

FWD: 5′–TGTTCAGCAACCCCTGCCACACCCTGT–3′

REV: 5′–ACAGGGTGTGGCAGGGGTTGCTGAACA–3′

H83A *Pa* HasA

FWD: 5′–TTCAGCAACCCCAGCGCAACCCTGT–3′

REV: 5′–ACAGGGTTGCGCTGGGGTTGCTGAA–3′

H83D *Pa* HasA

FWD: 5′–TTCAGCAACCCCAGCGATACCCTGTGG–3′

REV: 5′–CCACAGGGTATCGCTGGGGTTGCTGAA–3′

H83S *Pa* HasA

FWD: 5′–TTCAGCAACCCCAGCAGCACCCTGT–3′

REV: 5′–ACAGGGTGCTGCTGGGGTTGCTGAA–3′

Protein Expression Test

Plasmids were transformed into the RP523(DE3)[1] strain of *E. coli* then grown on an LB agar plate (1.5% (w/v) agar) supplemented with ampicillin (100 μg/mL) and hemin (20 μg/mL). Ten colonies were selected to test protein expression and were grown in 3 mL of TB supplemented with ampicillin (100 μg/mL) and hemin (20 μg/mL) at 37 °C overnight. Glycerol stocks were prepared for each culture. The cells were then sub-cultured (1:200 dilution) in 50 mL of TB supplemented with ampicillin (100 μg/mL), hemin (20 μg/mL), and 0.2% (w/v) glucose and grown at 37 °C until the OD_{600} reached ~0.6. IPTG was added (100 μM final concentration) and the cultures were then transferred to an 18 °C incubator and grown overnight. Protein expression levels were determined by preparing cell density-matched samples of each culture, as well as a pre-induction control. The cells were pelleted (2300 *g* for 5 minutes) then lysed using BugBuster® and the whole-cell lysate was analysed by SDS-PAGE and Western blot to identify His-tagged proteins. Western blotting was performed using a mouse anti-His_6 primary antibody, a goat anti-mouse–HRP conjugate secondary antibody, and the Classico HRP substrate.

Protein Expression and Purification

Colonies with the highest levels of protein expression, as identified by the expression test, were then selected for large-scale expression. For aerobic expression of the holoprotein, a 5 mL overnight culture of TB supplemented with ampicillin (100 μg/mL) and hemin (20 μg/mL) was inoculated directly from a frozen RP523 glycerol stock and then grown at 37 °C overnight. This culture was then used to inoculate 1 L of TB (containing 100 mg of ampicillin, 20 mg of hemin, and 0.2% (w/v) glucose) and was grown at 37 °C until the OD_{600} reached ~0.6. IPTG was added (100 μM final concentration) and the culture was then transferred to an 18 °C incubator and grown for ~20 hours. Cells were harvested by centrifugation at 4300 *g* for 20 min and the resultant pellets were snap-frozen in liquid nitrogen and stored at –80 °C.

For anaerobic protein expression of the apoprotein, a starter culture was prepared in a serum bottle containing 60 mL of TB, supplemented with ampicillin (100 μg/mL) and 0.2% (w/v) glucose, then inoculated directly from a frozen RP523 glycerol stock. The bottle was sealed with a rubber septum and purged with Ar for 1 hour, and then grown at 37 °C for at least 20 hours. A freshly autoclaved 10 L fermenter[4] was charged with 7 L of hot, sterilised TB and 1 L of hot, sterilised H_2O (to account for the ~1 L volume loss due to overnight purging). After cooling to 37 °C, 2 mL of antifoam and 400 mg of ampicillin were added, and the mixture was sparged with N_2 overnight. The following day, ampicillin (500 mg) and glucose (0.2% (w/v) final concentration) were added to the media, which was then inoculated with the 60 mL anaerobic starter culture. The resultant culture was grown at 37 °C until the OD_{600} reached ~0.45 and was then cooled to 18 °C prior to inducing protein expression. Once the OD_{600} reached 0.6–0.7 and the temperature reached 18 °C, IPTG was added (100 μM final concentration) and the culture was grown at 18 °C for 20–24 hours. Cells were harvested by centrifugation at 4300 *g* for 20 min and the resultant pellets were snap-frozen in liquid nitrogen and stored at –80 °C.

Cell pellets of *Cs* H-NOX were thawed in a water bath at room temperature then resuspended in Buffer A (50 mM K_2HPO_4, 300 mM NaCl, 20 mM imidazole, and 5% (v/v) glycerol at pH 8.0) supplemented with 110 mM benzamidine, 0.4 mM AEBSF, 0.1 mg/mL lysozyme, and 0.3 mg/mL DNase. Similarly, *Pa* HasA cell pellets were resuspended in Buffer B (20 mM K_2HPO_4, 250 mM NaCl, and 30 mM imidazole at pH 7.4) supplemented with 110 mM benzamidine, 0.4 mM AEBSF, 0.1 mg/mL lysozyme, and 0.3 mg/mL DNase. For both proteins, cells were lysed using a high-pressure homogeniser (Avestin Emulsiflex-C5). For the H-NOX variants, the lysate was heated for 1 h in a water bath at 70 °C for the holoprotein or 60 °C for the apoprotein. In the case of Y140F H-NOX, the apoprotein is not

(4) M. B. Winter, J. J. Woodward and M. A. Marletta, in *Cytochrome P450 Protocols, Methods in Molecular Biology*, ed. I. R. Phillips, E. A. Shepard and P. R. Ortiz de Montellano, Humana Press, Totowa, NJ, 2013, vol. 987, pp. 95–106.

particularly heat stable, so the lysate was heated at 50 °C. The lysate was clarified by centrifugation at 42,000 *g* and the supernatant was loaded onto a column containing His60 resin (1 mL for 1 L holoprotein pellets, or 2 mL for ~7 L apoprotein pellets) equilibrated with Buffer A for H-NOX or Buffer B for HasA. The resin was then washed with 20 column volumes (CV) of Buffer A for H-NOX or Buffer B for HasA. For some H-NOX variants with low expression (*e.g.*, apo Y140F), it was found that nucleic acid contamination was significant. In these cases, an additional wash step (20–30 CV) was included using Buffer C (50 mM K_2HPO_4, 1 M NaCl, and 20 mM imidazole at pH 8.0) to remove nucleic acid contamination from the Ni IMAC column. After the high salt wash, an additional wash step (20–30 CV of Buffer A) was utilised to re-equilibrate the column with the NaCl concentration of the elution buffer. Protein was eluted with Buffer D for H-NOX (50 mM K_2HPO_4, 300 mM NaCl, 300 mM imidazole, and 5% (v/v) glycerol at pH 8.0) or Buffer E for HasA (20 mM K_2HPO_4, 250 mM NaCl, and 300 mM imidazole at pH 7.4). The eluent was collected in six, 2–3 mL fractions. Fractions with comparable purity were combined and then dialysed overnight into Buffer F (20 mM K_2HPO_4, 250 mM NaCl, and 1 mM EDTA at pH 7.4) using snakeskin dialysis tubing (3500 MW cutoff). The protein was then filtered with a 0.22 μm syringe filter and stored in sterile tubes at 4 °C. Apoprotein concentrations were determined using the 280 nm absorbance. The 280 nm extinction coefficient and molecular weight of the proteins were calculated based on the primary sequence using the Expasy ProtParam tool.[5] Protein identity was validated by mass spectrometry.

WT H-NOX: ε_{280} = 30940 M^{-1} cm^{-1}, Calc. MW = 23347.88, Found (ESI-MS) = 23349.

W9F H-NOX: ε_{280} = 25440 M^{-1} cm^{-1}, Calc. MW = 23308.84, Found (ESI-MS) = 23310.

H102G H-NOX: ε_{280} = 30940 M^{-1} cm^{-1}, Calc. MW = 23267.79, Found (ESI-MS) = 23269.

H102S H-NOX: ε_{280} = 30940 M^{-1} cm^{-1}, Calc. MW = 23297.82, Found (ESI-MS) = 23299.

P115A H-NOX: ε_{280} = 30940 M^{-1} cm^{-1}, Calc. MW = 23321.84, Found (ESI-MS) = 23323.

Y140F H-NOX: ε_{280} = 29450 M^{-1} cm^{-1}, Calc. MW = 23331.88, Found (ESI-MS) = 23333.

WT HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20565.40, Found (ESI-MS) = 20565.

H32C HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20531.40, Found (ESI-MS) = 20531.

H32S HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20515.34, Found (ESI-MS) = 20516.

(5) https://web.expasy.org/protparam.

V38C HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20569.41, Found (ESI-MS) = 20569.

Y75F HasA: ε_{280} = 28420 M^{-1} cm^{-1}, Calc. MW = 20549.41, Found (ESI-MS) = 20552.

H83A HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20499.34, Found (ESI-MS) = 20500.

H83D HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20543.35, Found (ESI-MS) = 20544.

H83S HasA: ε_{280} = 29910 M^{-1} cm^{-1}, Calc. MW = 20515.34, Found (ESI-MS) = 20516.

Apoprotein Reconstitution

A 1 mg/mL stock solution of corrole **P-1** was prepared in DMSO. An aliquot of the H-NOX or HasA apoprotein (~1 mg) was diluted to 1 mL with Buffer F in a 1.7 mL microcentrifuge tube. Then, 100 μL of the **P-1** corrole solution was slowly added dropwise to the protein; the tube was gently inverted after each addition of two drops of the corrole stock solution. This method yields final concentrations of 175 μM corrole and 43 μM protein for the reconstitution reaction. The reaction mixture was gently rocked at room temperature for 2 hours, protected from ambient light. After incubation, the reaction mixture was centrifuged at 21,000 *g* for 1 minute to pellet any solids. The solution was then applied to a PD-10 column (Sephadex G-25 medium), pre-equilibrated with Buffer G (20 mM TEA, 1 mM EDTA, pH 6.5). The pink eluent was collected in 1 mL fractions. The same procedure was utilised to reconstitute proteins with Pd(MPIX) or Zn(PPIX).

Labelling HasA Cysteine Variants

In a 1.7 mL microcentrifuge tube, 1 mL of ~100 μM HasA cysteine variant (apo or holo) in Buffer F was treated with 50 equivalents of TCEP (stock solution in Buffer F) and the solution was rocked at room temperature for 10 minutes. Then, 15 equivalents of fluorescein-5-maleimide (stock solution in DMSO) were slowly added dropwise to the protein; the tube was gently inverted after each addition of two drops of the dye stock solution. The resultant mixture was rocked at room temperature overnight, protected from ambient light. The reaction mixture was centrifuged at 21,000 *g* for 1 minute to pellet any solids. The solution was then applied to a PD-10 column (Sephadex G-25 medium), pre-equilibrated with Buffer F. The eluent was collected in 1 mL fractions. In order to remove any traces of unbound dye, the protein-containing fractions were concentrated to < 1 mL (using a 5000 MW cutoff filter) and a second PD-10 column was run using Buffer F. The labelling efficiency was calculated[6] to be 42%. It should be noted that incomplete labelling

(6) Y. Kim, S. O. Ho, N. R. Gassman, Y. Korlann, E. V. Landorf, F. R. Collart and S. Weiss, *Bioconjugate Chem.*, 2008, **19**, 786–791.

with a FRET donor is not an issue.[7] The experiments interrogate only the labelled population of the sample, and the presence of unlabelled protein does not impact the results. Reconstitution of fluorescein-labelled HasA variants with **P-1** was performed using the general method described above.

Physical Measurements

UV–vis absorption spectra were acquired using a Cary 300 spectrometer (Agilent) or a NanoDrop 2000C (Thermo Scientific). Steady-state emission spectra were recorded on a Horiba Scientific FluoroMax-4 spectrofluorometer. Time-resolved emission was measured using a PicoQuant FluoTime 300 fluorometer and a 470 nm PicoQuant pulsed diode laser with a PDL 820 driver. Signal was collected at the emission maximum with a 2 nm bandwidth and detected using a TimeHarp 260 time-correlated single photon counting board. Data analysis was performed using PicoQuant FluoFit software. Relative quantum yields of fluorescein-labelled HasA in TEA buffer (η = 1.3329, interpolated from published data)[8] or phosphate buffer (η = 1.334)[9] were calculated using fluorescein 27 in 0.1 M NaOH (η = 1.335) as the reference (Φ_{ref} = 0.87)[10] according to the following equation:

$$\Phi_{sam} = \Phi_{ref}\left(\frac{\nabla_{sam}}{\nabla_{ref}}\right)\left(\frac{\eta_{sam}}{\eta_{ref}}\right)^2 \quad (1)$$

where ∇ is the slope of the plot of integrated fluorescence intensity versus absorbance (constructed with 5 points) and η is the refractive index of the solvent. Time of flight electrospray ionization mass spectrometry (TOF-ESI-MS) of proteins was performed using an Agilent 1260 series liquid chromatography (LC) system with an Agilent 6530C Q-TOF LC-MS system in positive ion mode. Agilent Mass Hunter Workstation software was used to extract total ion chromatograms from the raw data and the Maximum Entropy deconvolution algorithm was used to reconstruct the original masses of the protein sample.

Förster Resonance Energy Transfer (FRET) Analysis

The efficiency of energy transfer (E) from the appended dye to the protein cofactor was determined using Förster analysis:

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (2)$$

where r is the distance between the donor and acceptor and R_0 is the Förster distance, or the distance at which the energy transfer is 50%. The efficiency can be measured

(7) J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer, New York, 2006.
(8) Y.-M. Tseng and A. R. Thompson, *J. Chem. Eng. Data.*, 1964, **9**, 264–267.
(9) F.-C. Chien and S.-J. Chen, *Opt. Lett.*, 2006, **31**, 187–189.
(10) M. Grabolle, M. Spieles, V. Lesnyak, N. Gaponik, A. Eychmüller and U. Resch-Genger, *Anal. Chem.*, 2009, **81**, 6285–6294.

experimentally using time-resolved or steady-state methods (for absorption-matched samples):

$$E = 1 - \frac{\tau_{D-A}}{\tau_D} = 1 - \frac{I_{D-A}}{I_D} \quad (3)$$

where τ represents the fluorescence lifetime and I represents the emission intensity; the subscript is used to denote the donor alone (D) or the donor in the presence of the acceptor (D–A). The Förster distance is a constant quantity for a given donor–acceptor pair and is calculated using the following equation:

$$R_0^6 = \frac{9000(\ln 10)\kappa^2\Phi_D}{128\pi^5 N n^4} \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\,\lambda^4\,\mathrm{d}\lambda \quad (4)$$

where κ^2 is the relative orientation factor of the dipoles, taken to be 2/3 for random donor–acceptor orientations, Φ_D is the quantum yield of the donor, N is Avogadro's number, and n is the refractive index of the solvent. The latter half of the equation represents the spectral overlap integral of donor emission and acceptor absorption, often denoted by the variable J, where $F_D(\lambda)$ is the normalised emission intensity of the donor and $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor at wavelength λ.

Molecular Dynamics Simulations

Simulations of WT (PDB ID: 3TF0) and P115A (PDB ID: 3EEE) *Cs* H-NOX were performed using previously reported crystal structures. To simulate the apoprotein, haem, oxygen, and buffer ions were removed from the structures. Crystallographic water molecules were retained. The Protein Preparation tool in Maestro (Schrödinger, Inc.) was used to prepare each structure. Titratable residues were retained in their dominant protonation states at pH 7. A neutral methylamide group was added to cap the C-terminus of each structure; the buried N-terminus was left uncapped (positively charged), as it appears to form ionic interactions with neighboring side chains.

Six independent ~1.0 μs simulations were performed for WT and P115A apo *Cs* H-NOX. The program Dabble[11] was used to place each structure in a water box with initial dimensions of 92 Å × 92 Å × 92 Å. Sodium and chloride ions were added to neutralise each system to a concentration of 150 mM; each system contained ~23,000 water molecules, 59–60 sodium ions, and 63 chloride ions for a total of ~73,000 atoms. All simulations utilised the CHARMM36m force field for the protein, the CHARMM36 parameter set for ions, and the CHARMM TIP3P water model.[12–14] Simulations were performed using the AMBER 18

(11) R. Betz, Dabble, Version 2.6.3. *Zenodo* DOI: 10.5281/zenodo.836914.
(12) J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell, *Nat. Methods*, 2017, **14**, 71–73.
(13) R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.

software package[15] with the Compute Unified Device Architecture (CUDA) version of Particle-Mesh Ewald Molecular Dynamics (PMEMD) on graphical processing units (GPUs).[16] Systems were heated from 0 K to 100 K in the NVT ensemble over 12.5 ps, and then from 100 K to 310 K in the NPT ensemble over 125 ps at 1 bar with 10.0 kcal mol^{-1} $Å^{-2}$ harmonic restraints applied to non-hydrogen protein atoms. Systems were then equilibrated at 310 K in the NPT ensemble at 1 bar with harmonic restraints on non-hydrogen protein atoms, starting at 5.0 kcal mol^{-1} $Å^{-2}$, and then decreasing in 1.0 kcal mol^{-1} $Å^{-2}$ increments every 2 ns for 10 ns, and then decreasing in 0.1 kcal mol^{-1} $Å^{-2}$ increments every 2 ns for an additional 18 ns. Subsequent production simulations were carried out in the NPT ensemble at 310 K and 1 bar using a Langevin thermostat and a Monte Carlo barostat with isotropic pressure control. Simulations employed a 4 fs time step with hydrogen mass repartitioning.[17] Bond lengths were constrained using SHAKE.[18] Non-bonded interactions were cut off at 10.0 Å. Long-range electrostatic interactions were computed using Particle Mesh Ewald with an Ewald coefficient of approximately 0.28 and a B-spline interpolation order of 4. The width of a grid cell in the FFT grid was chosen to be approximately 1 Å. For each independent simulation, initial atom velocities were assigned randomly and independently. Trajectory snapshots were written out every 200 ps.

The AmberTools19 CPPTRAJ package[19] was employed to process trajectories and Visual Molecular Dynamics (VMD) was used to visualise and analyse simulations.[20] Plots and other analyses used data sampled every 1 ns from each trajectory. Plots involving simulation data were visualised using the PyPlot package in Matplotlib.[21] For analysis of simulation data in aggregate, the first 100 ns of simulations were removed to allow for system relaxation. To calculate the root-mean-square deviation of trajectories from their initial conformation, simulations were aligned on Cα residues. Simulation trajectories and analysis code are available on Zenodo.

---

(14) J. Huang and A. D. Mackerell, *J. Comput. Chem.*, 2013, **34**, 2135–2145.
(15) D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, G. Giambasu, T. Giese, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, R. Krasny, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, J. Swails, R. C. Walker, J. Wang, H. Wei, L. Wilson, R. M. Wolf, X. Wu, L. Xiao, Y. Xiong, D. M. York and P. A. Kollman, *AMBER* 2019, University of California, San Francisco, 2019.
(16) R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.
(17) C. W. Hopkins, S. Le Grand, R. C. Walker and A. E. Roitberg, *J. Chem. Theory Comput.*, 2015, **11**, 1864–1874.
(18) J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
(19) D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
(20) W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
(21) J. D. Hunter, *Comput. Sci. Eng.*, 2007, **9**, 90–95.

Technical Notes

The low solubility of the corrole and propensity for aggregation make it difficult to determine the exact solution concentration of the corrole. The reconstitution conditions were optimized for the WT proteins in a previous report,[2] and these methods have been utilised here for the protein variants. In order to obtain maximal corrole incorporation while avoiding non-specific corrole binding, ~ 4 mole equivalents of **P-1** per mole of protein must be used. Using more corrole results in non-specific binding, which is readily detected by absorption spectroscopy. The excess corrole is difficult to completely remove, even with dialysis and/or additional desalting columns. Using less corrole does not result in complete protein loading (*i.e.*, there is residual apoprotein in the sample), as evidenced by the WT protein as a benchmark. Small errors in the concentration of the corrole stock solution (*e.g.*, balance fluctuations and/or incomplete drying of the solid corrole) can lead to over- or under-loading the protein with corrole. Given the day-to-day variability, the only way to obtain reproducible, quantitative corrole binding is to determine the extent of corrole binding relative to the WT protein as a standard, which was prepared in parallel using the same corrole stock solution.

Since each protein variant incorporates a different amount of corrole and the protein recovery can vary from sample to sample, the concentrations of each sample are different. Moreover, differing amounts of bound corrole in the protein variants necessitate the preparation of samples with different concentrations to record satisfactory absorption spectra. Consequently, the absorption spectra presented in the manuscript have been normalised to the protein absorbance at 280 nm to visualise all of the data in a way that illustrates the differences in corrole incorporation.

In a previous study,[2] it was demonstrated that the corrole conjugates with WT H-NOX and HasA are stable for nearly two months with nominal loss of corrole: 13% for H-NOX after 56 days, 5% for HasA after 44 days. These results clearly indicate that the conjugates with the WT proteins are stable in the absence of excess corrole. In the variants with lower **P-1** affinity, corrole dissociation does occur over time. Thus, spectra are recorded as soon as the sample has been purified on the PD-10 desalting column. The samples for each variant are stable for at least 1.5 hours when stored on ice, enabling the completion of all experiments before any spectral changes are observed (Figure S20). Corrole dissociation is readily detected by absorption spectroscopy; the spectrum of free, unbound corrole is distinct from the protein-bound corrole, as previously reported.[2]

When comparing the protein variants, all samples are prepared at the same time using the same corrole stock solution and treated in an identical manner. The only difference in the samples is a single residue in the heme binding pocket. Therefore, it is not unreasonable to conclude that observed differences in holoprotein recovery are a direct consequence of the

point mutation.

While the HasA constructs have a cleavable His_6 tag, cleavage was not accomplished with TEV protease. Preliminary experiments with a tagless construct indicate that the His_6 tag does not interfere with cofactor binding. Moreover, it was previously demonstrated that **P-1** does not bind histidine in solution,[2] suggesting that non-specific corrole binding to the His_6 tag is unlikely. For these reasons, His-tagged constructs were utilised to easily purify the numerous protein variants. While it is still possible that corrole binding could occur with the His_6 tag, this is a constant for all protein variants; the only difference in the samples is a single residue in the heme binding pocket. If the His_6 tag were the primary corrole binding site, then similar extents of corrole incorporation would be observed for all variants. However, this is not what is observed experimentally.

Haem Binding in H-NOX and HasA Variants

To survey **P-1** binding in *Cs* H-NOX, the following variants were prepared: W9F,[22] H102G,[23] H102S, P115A,[24] and Y140F.[25] Of these variants, only H102S has not been previously expressed and characterised. To circumvent the laborious screens necessary for protein expression in RP523 cells, the H102G and H102S variants were initially expressed in the BL21(DE3) strain of *E. coli* in an attempt to directly obtain the apoprotein. Given the impaired haem binding of the H102G variant (due to loss of the haem-ligating histidine),[23] we expected the same for H102S. To facilitate apoprotein isolation, the media was not supplemented with iron and δ-aminolevulinic acid. The H102G variant was obtained largely as the apo protein (~5% haem incorporation), whereas H102S exhibited a substantial level (~22%) of haem incorporation, albeit with a significantly different absorption spectrum compared to the WT protein (Figure S1). When expressed in RP523 cells, which are used to maximise haem incorporation, both H102G and H102S exhibited relatively high levels (~50%) of haem incorporation. The absorption spectra of these variants are blue shifted relative to the WT protein (Figure S2). This suggests that the high levels of imidazole utilised to elute the protein from the Ni affinity column resulted in recovery of haem binding, as previously observed for the H105G variant of the β H-NOX domain of *Rn* soluble guanylate cyclase (sGC).[26]

To survey **P-1** binding in *Pa* HasA, the following variants were prepared: H32C, H32S, V38C, Y75F, S82C, H83A, H83D, and H83S (Figure S8). Although related variants have been

(22) E. M. Boon, S. H. Huang and M. A. Marletta, *Nat. Chem. Biol.*, 2005, **1**, 53–59.
(23) C. Olea, M. A. Herzik, J. Kuriyan and M. A. Marletta, *Protein Sci.*, 2010, **19**, 881–887.
(24) C. Olea, E. M. Boon, P. Pellicena, J. Kuriyan and M. A. Marletta, *ACS Chem. Biol.*, 2008, **3**, 703–710.
(25) E. M. Boon and M. A. Marletta, *J. Am. Chem. Soc.*, 2006, **128**, 10022–10023.
(26) Y. Zhao, J. P. M. Schelvis, G. T. Babcock and M. A. Marletta, *Biochemistry*, 1998, **37**, 4502–4509.

previously characterised (*e.g.*, Y75A[27] and H32A[28]), only H83A has been previously reported.[27] These variants can be grouped together into four categories on the basis of their haem-bound absorption spectra (Figure S9): (a) WT, V38C, S82C, and H83D, (b) H32C and H32S, (c) Y75F, and (d) H83A and H83S. The absorption spectra are complex and reflect a mixture of high- and low-spin states at room temperature.[27] The similarity of the absorption spectra reflects similarities in the electronic structure of the heme cofactor. Since the only difference in these proteins is a single residue, it is not unreasonable to conclude that the observed spectral similarities (or differences) to the WT protein are directly related to the point mutation. It is not surprising that V38C and S82C are similar to the WT protein, as the modified residues are on the periphery of the haem binding site and are not expected to perturb the electronic structure or ligation state of the haem. The H32C and H32S variants are similar to, but distinct from, the H32A variant of *Pa* HasA,[28] as well as HasA from *Yersinia pestis*,[29] which lacks a native H32 residue. Both H83A and H83S are consistent with the previously reported spectra of H83A.[27]

Surprisingly, H83D is quite similar to the WT protein, more so than the other H83 variants (Figure S9d). The H83D variant was expected to provide an alternative hydrogen bonding partner for Y75. A model of the H83D variant was generated using a suitable rotamer that positions the aspartate side chain to form a hydrogen bond with Y75 (Figure S10). Indeed, the distance between O atoms of Y75 and D83 is 2.8 Å, which is quite similar to the H-bond between Y75 and H83 in the WT protein (2.6 Å). Conversely, a model of the H83S variant positions the side chain such that the distance between the O atoms of Y75 and S83 is 3.5 Å, precluding the formation of a suitable hydrogen bonding interaction. Consequently, H83S and H83A exhibit similar spectra, which are distinct from WT and H83D. This suggests that the H83D variant is indeed capable of interacting with Y75 in a way that recapitulates the native interaction.

Pd(MPIX) Binding in H-NOX Variants and WT HasA

The solubility of Pd(MPIX) in buffer is very low and the porphyrin is prone to aggregation. For all reconstitution reactions, the samples are rocked for two hours so that any precipitated solids may redissolve in solution as the various equilibria shift upon protein binding. High Pd(MPIX) binding is observed for the H-NOX variants (WT, H102G, H102S), but no Pd(MPIX) binding is observed for WT HasA under the same experimental conditions. After separating the protein from the unbound porphyrin, a colorless HasA

(27) R. Kumar, H. Matsumura, S. Lovell, H. Yao, J. C. Rodríguez, K. P. Battaile, P. Moënne-Loccoz and M. Rivera, *Biochemistry*, 2014, **53**, 2112–2125.
(28) E. T. Yukl, G. Jepkorir, A. Y. Alontaga, L. Pautsch, J. C. Rodriguez, M. Rivera and P. Moënne-Loccoz, *Biochemistry*, 2010, **49**, 6646–6654.
(29) R. Kumar, S. Lovell, H. Matsumura, K. P. Battaile, P. Moënne-Loccoz and M. Rivera, *Biochemistry*, 2013, **52**, 2705–2707.

protein solution is recovered that does not contain any Pd(MPIX) detectible by absorption spectroscopy.

It is well known that square planar Pd(II) complexes do not bind axial ligands. This has been demonstrated experimentally, observing that the heme-ligating residue (H102) in H-NOX is not required to bind Pd(MPIX); WT, H102G, and H102S H-NOX all bind the same amount of the porphyrin. Therefore, the nature of the heme binding pocket itself enables Pd(MPIX) binding in H-NOX. This is likely due to the buried hydrophobic heme pocket and the Y–S–R motif. Conversely, HasA lacks these features, as apo HasA exhibits an extended conformation with exposed hydrophobic surfaces rather than a true, buried heme pocket. These differences account for the isolation of stable H-NOX conjugates with Pd(MPIX), while no detectible binding to HasA is observed. Based on these observations, hydrophobic interactions alone are insufficient for Pd(MPIX) binding in HasA. This interpretation is further supported by the lack of **P-1** binding in the H83 variants. Consequently, a specific protein–cofactor interaction is required to enable cofactor binding in HasA, as hydrophobic interactions alone appear to be insufficient.

FRET Experiments with Haem Bound HasA Variants

In the case of labelled holo/haem V38C, fluorescein emission is significantly quenched relative to the corresponding apoprotein. For the emission spectra in Figure S15a, the samples were prepared such that both the apo and holo samples have the same amount of fluorescein absorbance at the excitation wavelength (470 nm), making the emission intensity directly comparable. This takes into account the haem absorbance at 470 nm (Figure S14). Analysis of these data indicates that the FRET efficiency is 88%. These results are corroborated by time-resolved data, which accounts for sample heterogeneity and avoids the need to correct sample absorbance. In the case of fluorescein-labelled apo V38C (donor alone in phosphate buffer at pH 7.4), the data cleanly fits to a monoexponential decay with a 4.28 ns lifetime. Conversely, the holo sample exhibits biexponential kinetics. The lifetime of the long component is consistent with the apoprotein, while the short component has a lifetime of 0.47 ns. The emission lifetimes yield a FRET efficiency of 89%, consistent with the steady-state data. This yields a donor–acceptor distance (r) of 3.01 nm.

This long distance suggests that the protein exhibits an open, apo-like conformation. Based on the holo HasA structure, the Fe to Cα of V38 distance is 1.03 nm (Figure 5a). At most, the dye, linker, and side chain would add 1.36 nm (based on a 3D structural model generated in GaussView, Figure S16), to give an expected maximum distance of 2.39 nm; this is substantially shorter than the FRET-determined distance of 3.01 nm. Instead, the FRET-determined distance is similar to the O atom of the Y75 side chain to Cα of V38 distance of 2.54 nm observed in the apo HasA crystal structure (Figure 5b). The difference between the FRET-determined distance and the apo crystal structure distance is 0.47 nm. This

deviation reasonably accounts for the cysteine side chain, the maleimide, and the fluorescein dye. It has been previously observed that V30, V37, and V38 form a hydrophobic network and are involved in key interaction events during H32 loop closure, as identified by MD simulations.[30] Consequently, it is not surprising that labelling residue 38 with a charged dye perturbs the specific interactions necessary for H32 loop closure, resulting in an extended conformation.

In order to validate the method and ensure that the observations described above are not due to some artifact or misinterpretation of the data, the S82C variant was prepared. In this case, the fluorescein label is directly adjacent to the haem binding pocket and resides on the Y75 loop, avoiding the complication of a potential apo-like conformation. The steady-state data, correcting for haem absorbance at the excitation wavelength, gives a FRET efficiency of 96% (Figure S15b). This is corroborated by time-resolved data, where the fluorescence lifetime of the fluorescein dye is reduced to 94 ps, corresponding to a FRET efficiency of 98%. The FRET efficiency in the S82C variant is significantly higher than in the V38C variant (89%), indicating that the dye is closer to the haem. The donor–acceptor distance obtained from the FRET analysis is 2.23 nm. Based on the holo HasA structure, the Fe to Cα of S82 distance is 0.88 nm. Adding the estimated length of the dye, the longest expected D–A distance would be 2.24 nm, which corresponds well to the FRET distance of 2.23 nm.

(30) G. Jepkorir, J. C. Rodríguez, H. Rui, W. Im, S. Lovell, K. P. Battaile, A. Y. Alontaga, E. T. Yukl, P. Moënne-Loccoz and M. Rivera, *J. Am. Chem. Soc.*, 2010, **132**, 9857–9872.

Table S1. FRET Analysis of HasA Cysteine Variants

Variant	Cofactor	Buffer	pH	Φ_D[a]	J (M^{-1} cm^3)[b]	R_0 (nm)[c]	τ_D (ns)	τ_{D-A} (ns)	Eτ[d]	$r\tau$ (nm)[e]	E_I[f]	r_I (nm)[g]
H32C	**P-1**	TEA	6.5	0.28	6.82×10^{-14}	3.91	4.11	1.06	0.74	3.28	0.68	3.44
V38C	**P-1**	TEA	6.5	0.28	6.82×10^{-14}	3.91	4.11	0.91	0.78	3.16	0.75	3.25
V38C	Haem	Phosphate	7.4	0.66	4.88×10^{-14}	4.26	4.28	0.47	0.89	3.01	0.88	3.06
S82C	Haem	Phosphate	7.4	0.66	4.88×10^{-14}	4.26	4.28	0.09	0.98	2.23	0.96	2.51

[a] Fluorescence quantum yield of fluorescein-labelled apo V38C, determined relative to fluorescein 27 and calculated using Eq. 1; [b] Calculated using the integral in Eq. 4; [c] Calculated using Eq. 4; [d] FRET efficiency determined from time-resolved data, calculated using Eq. 3; [e] D–A distance determined from time-resolved data, calculated using Eq. 2; [f] FRET efficiency determined from steady-state data, calculated using Eq. 3; [g] D–A distance determined from steady-state data, calculated using Eq. 2.

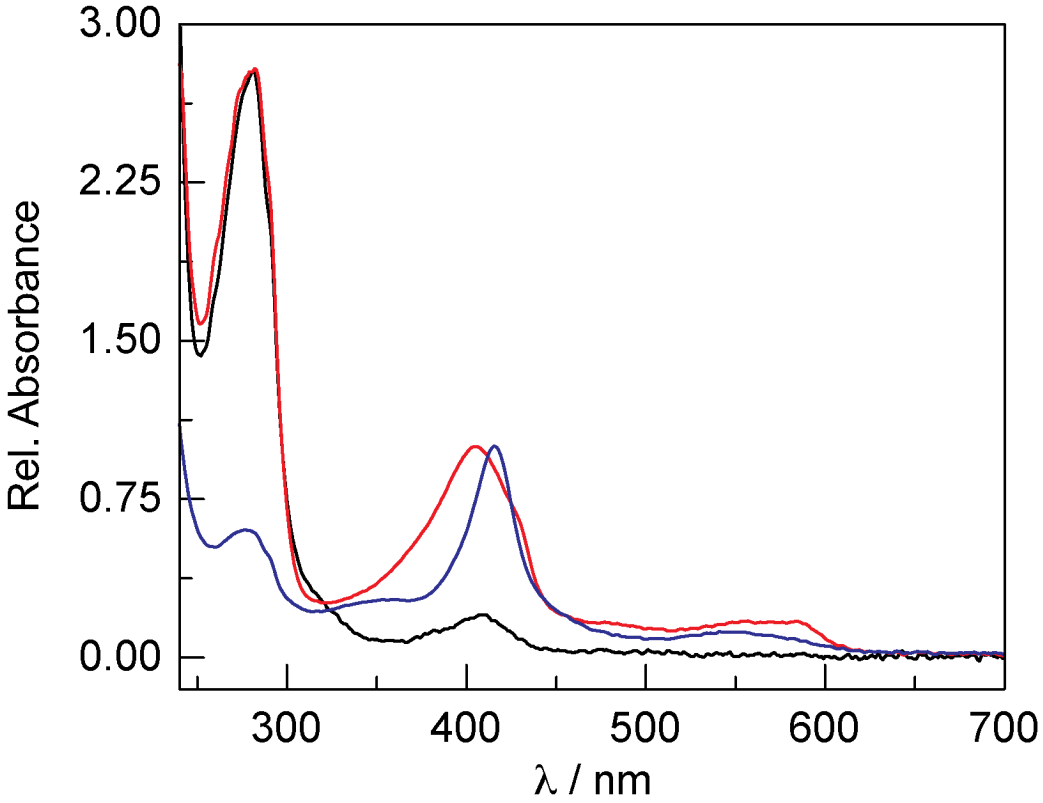

Figure S1. Absorption spectra of H102 variants of *Cs* H-NOX when expressed in the BL21(DE3) strain of *E. coli*. As expected, the H102G variant (—) cannot efficiently bind haem, giving protein with only ~5% haem incorporation relative to the WT protein (when expressed in the RP523(DE3) strain of *E. coli*). Conversely, the H102S variant (—) was purified with a haem loading of ~22%. However, the haem absorption profile is significantly different from the WT protein (—). This is likely due to haem oxidation, since the absence of the haem-ligating histidine likely prevents stabilisation of the O_2 adduct, resulting in haem oxidation.

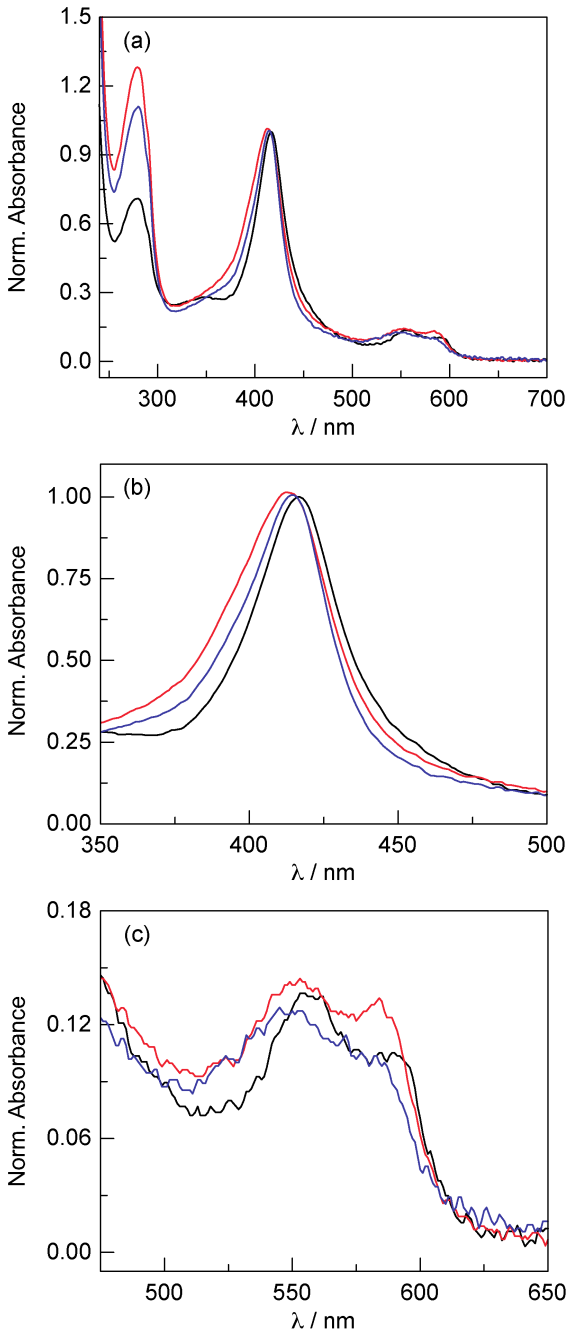

Figure S2. Normalised absorption spectra of H102 variants of *Cs* H-NOX when expressed in the RP523(DE3) strain of *E. coli*. (a) Surprisingly, both the H102G (—) and H102S (—) variants bind a significant amount of haem, exhibiting 49% and 56% haem incorporation, respectively, relative to the WT protein (—). However, the absorption profiles are blue-shifted relative to the WT protein. The (b) Soret and (c) Q band regions have been plotted separately to better illustrate the spectral shift. This result is consistent with imidazole ligation at the haem, as previously observed for the H105G variant of the β H-NOX domain of *Rn* sGC. Haem binding is rescued in these variants due to the presence of excess imidazole (300 mM) in the elution buffer from the Ni IMAC column.

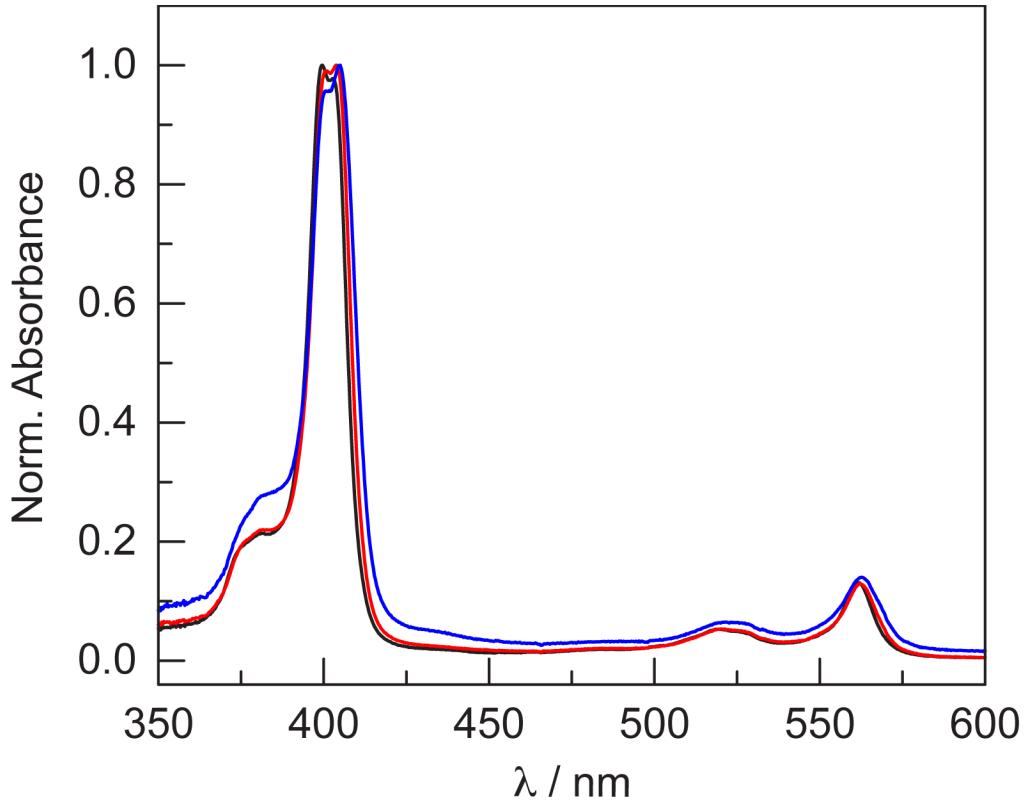

Figure S3. Absorption spectra of WT (—), H102G (—), and H102S (—) H-NOX in TEA buffer at pH 6.5. The corrole absorbance is the same with or without H102, indicating that the histidine does not coordinate to the phosphorus center and the complex remains five-coordinate. A nominal red-shift is observed that likely reflects changes in the dielectric constant of the haem binding pocket in these variants.

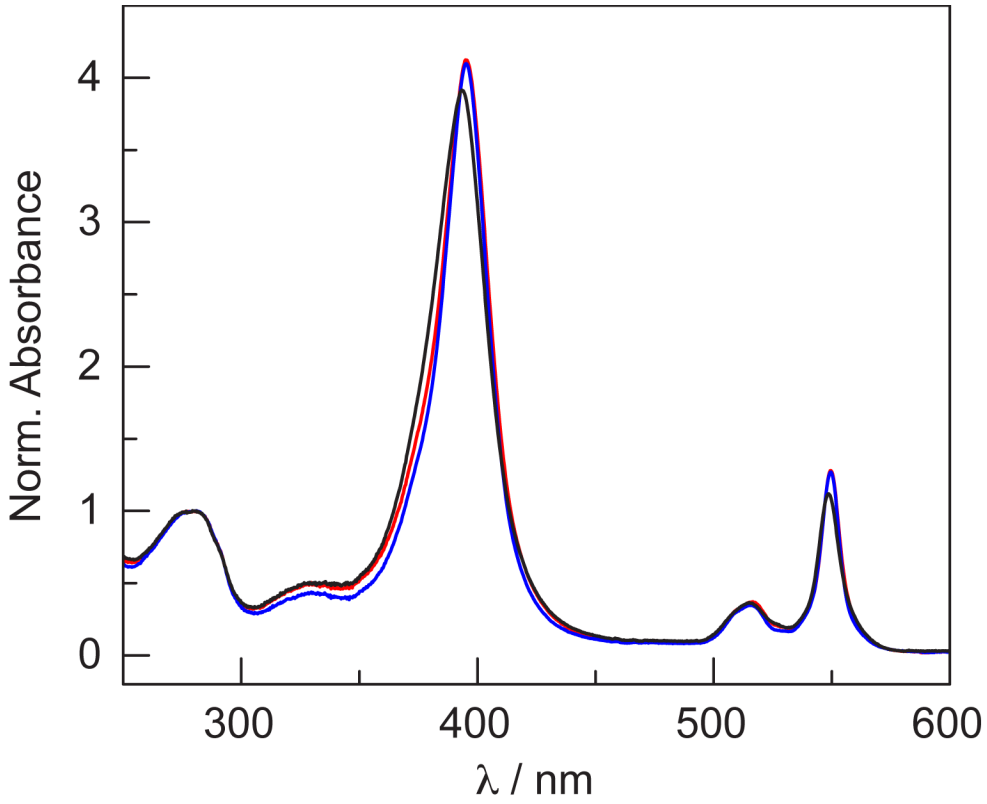

Figure S4. Normalised absorption spectra of WT (—), H102G (—), and H102S (—) H-NOX reconstituted with Pd(MPIX) in TEA buffer at pH 6.5. The spectra of H102G and H102S are nominally red-shifted relative to the WT protein; this is also observed for **P-1** binding in these variants (see Figure S3). The H102 variants exhibit porphyrin incorporation that is identical to the WT protein, indicating that apo H102G and H102S are properly folded.

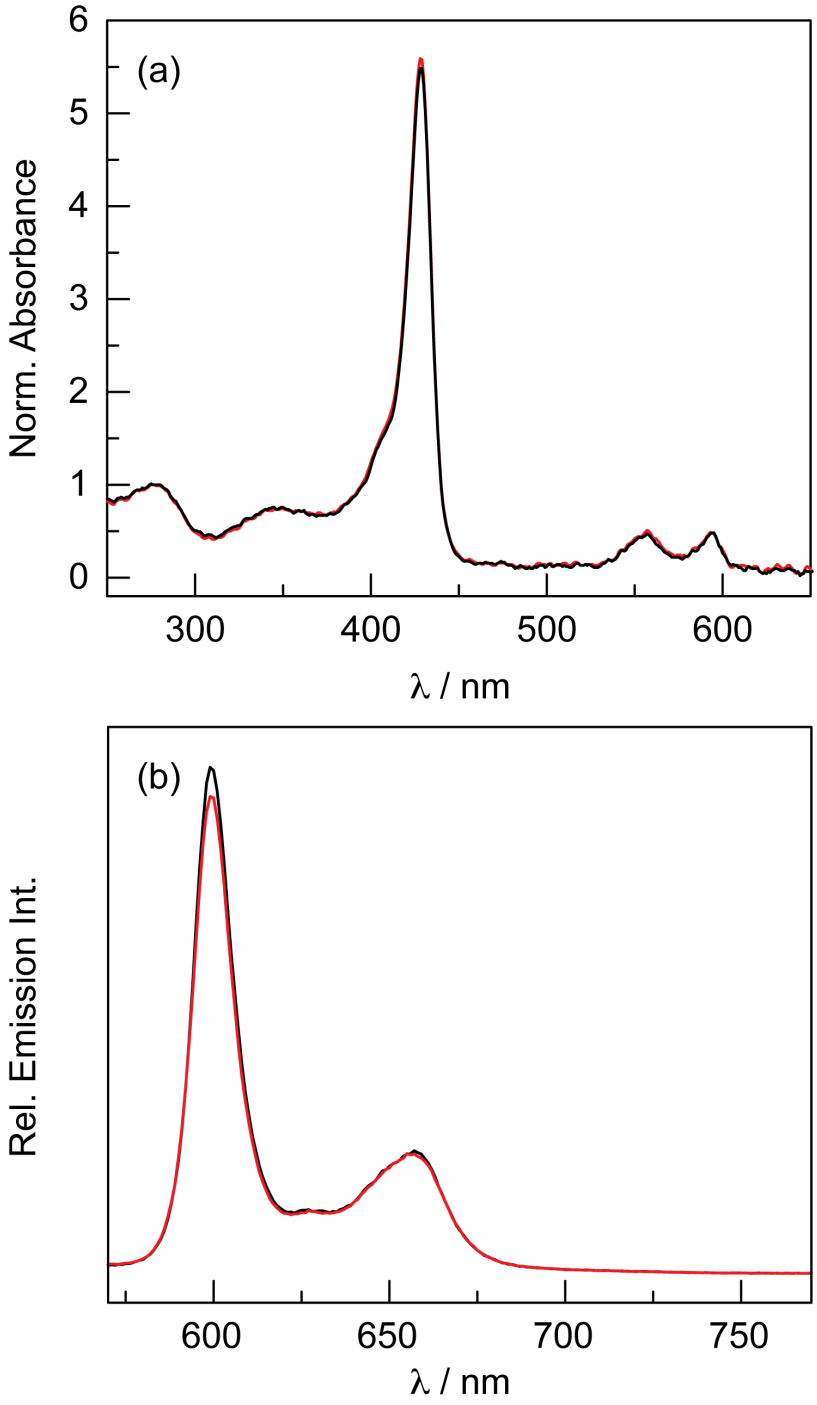

Figure S5. (a) Normalised absorption spectra of WT (—) and P115A (—) H-NOX reconstituted with Zn(PPIX) in TEA buffer at pH 6.5. Porphyrin incorporation is identical to the WT protein, indicating that apo P115A is properly folded. (b) Absorbance-matched emission spectra (A(560) = 0.0823 ± 0.0002) of WT (—) and P115A (—) H-NOX reconstituted with Zn(PPIX) in TEA buffer at pH 6.5. Emission from the P115A variant is 96% that of the WT protein.

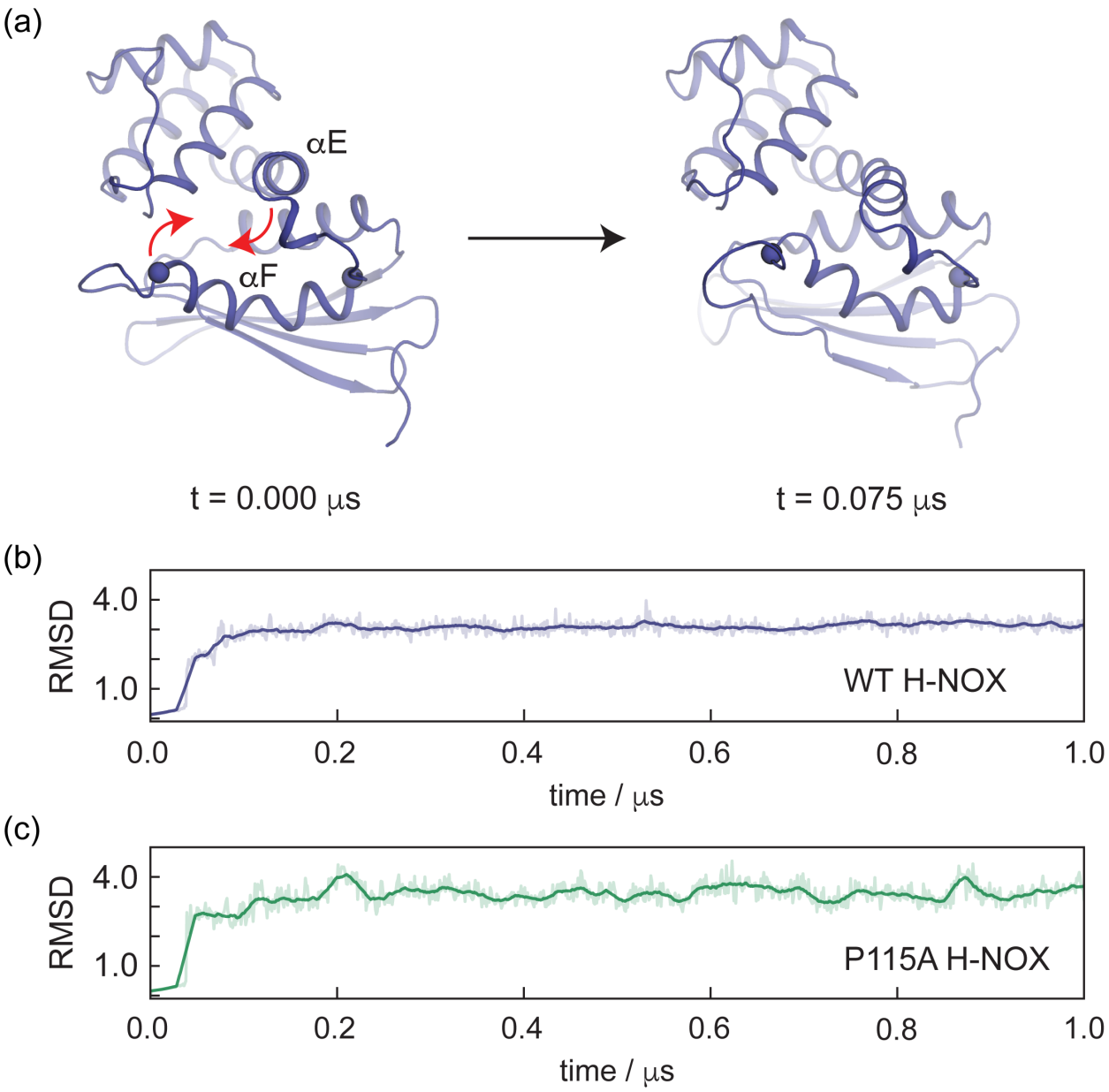

Figure S6. (a) Initial rearrangement of the WT apo H-NOX structure in MD simulations. The αF helix (bounded by spheres) moves upward and the αE helix collapses downward (red arrows) to fill the vacancy in the haem cavity. Snapshots reflect the simulation at 0 and 75 ns. The structures are aligned on residues 1–90. (b, c) Plots of the Cα root-mean-square deviation (RMSD) from the initial structure over the course of the simulations for (b) WT and (c) P115A H-NOX. The RMSD plateaus within the first 100 ns of simulation, rearranging to fill the cavity occupied by haem in the crystal structure of the holoproteins. The RMSD traces correspond to Simulation 1 for both WT and P115A H-NOX (see Figure S7).

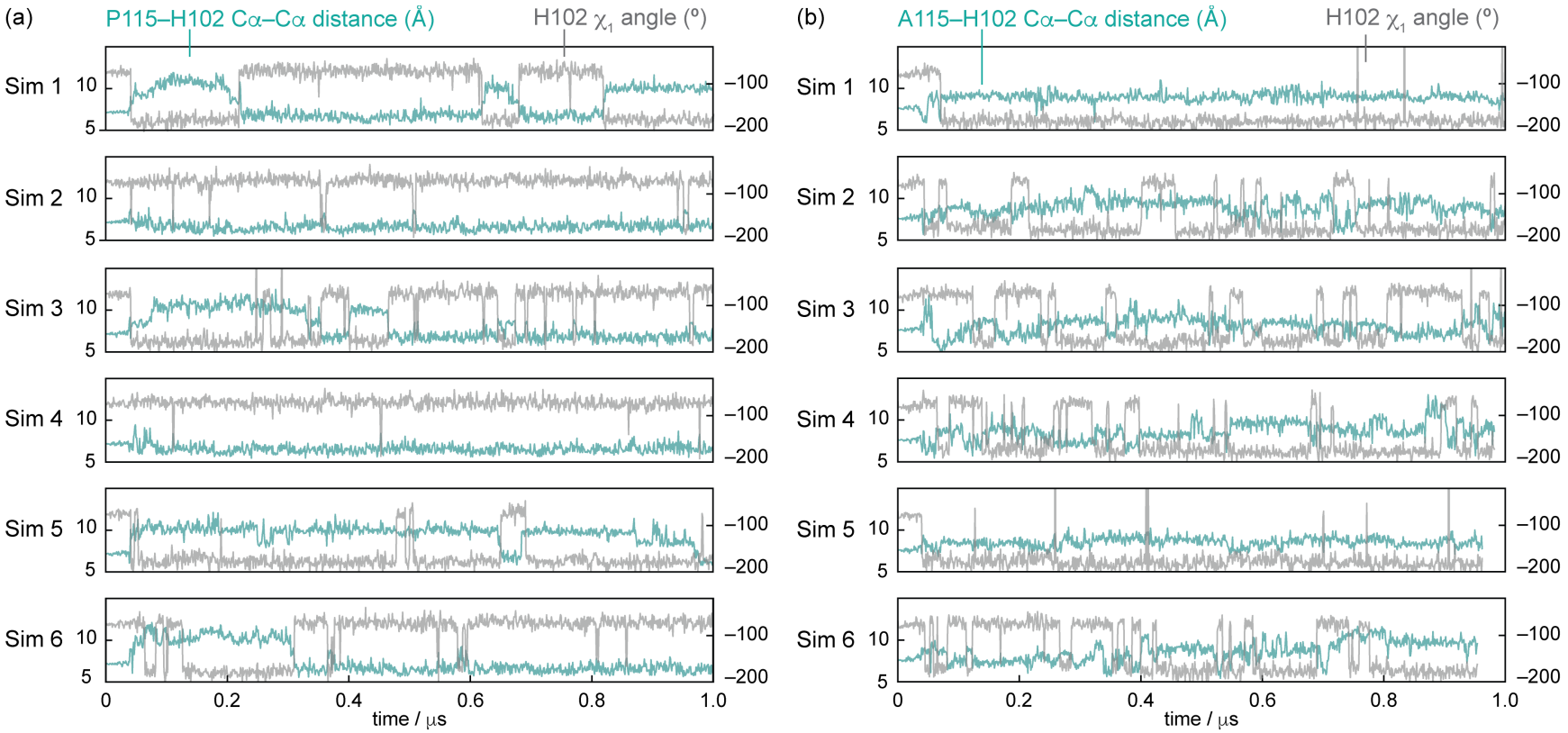

Figure S7. (a) Traces representing the Cα–Cα distance between H102 and P115 (turquoise) and the H102 χ_1 torsion angle (gray) over the course of six, 1 μs simulations of WT apo H-NOX. Simulation 1 is the representative trace illustrated in the main text (Figure 3c). (b) Traces representing the Cα–Cα distance between H102 and A115 (turquoise) and the H102 χ_1 torsion angle (grey) over the course of six, ~1 μs simulations of P115A apo H-NOX. Simulation 2 is the representative trace illustrated in the main text (Figure 3c).

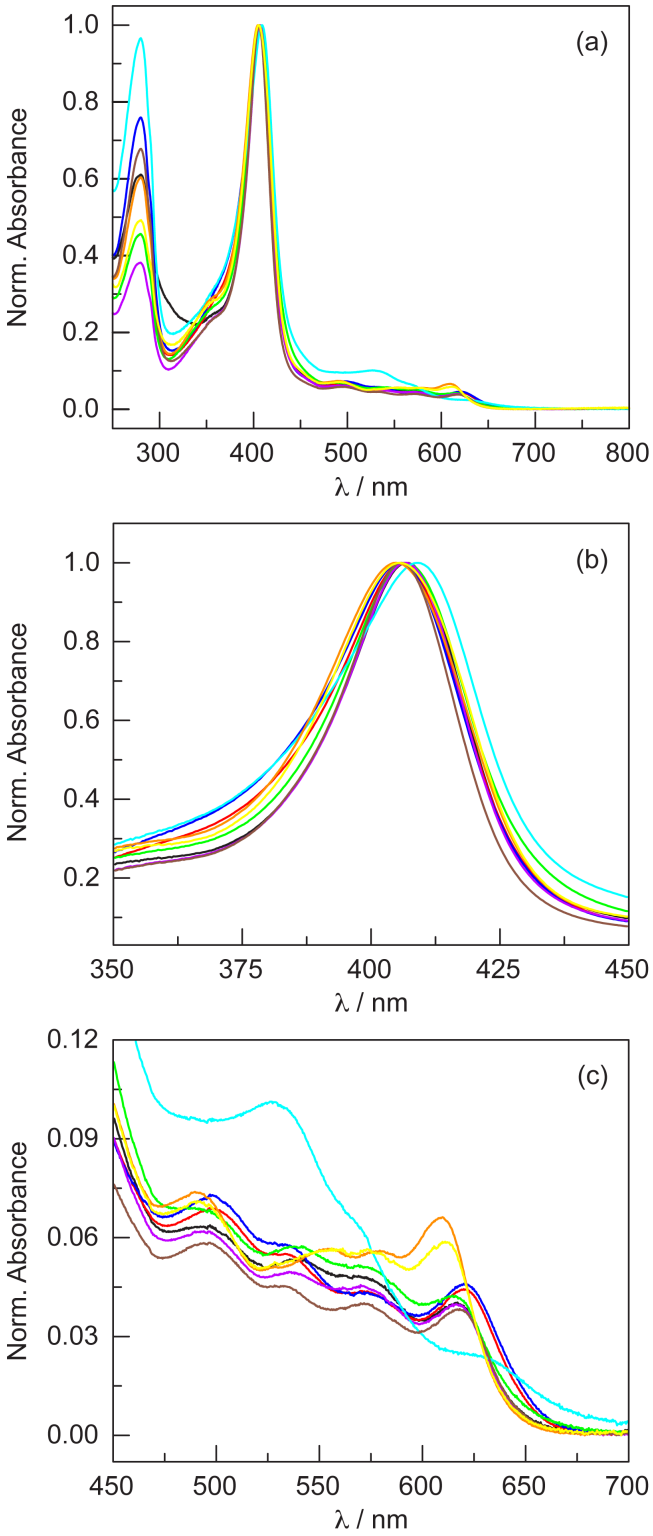

Figure S8. (a) Absorption spectra of holo/haem *Pa* HasA variants in phosphate buffer at pH 7.4: WT (—), H32C (—), H32S (—), V38C (—), Y75F (—), S82C (—), H83A (—), H83D (—), and H83S (—). The (b) Soret and (c) Q band regions have been plotted separately to better illustrate differences in these variants.

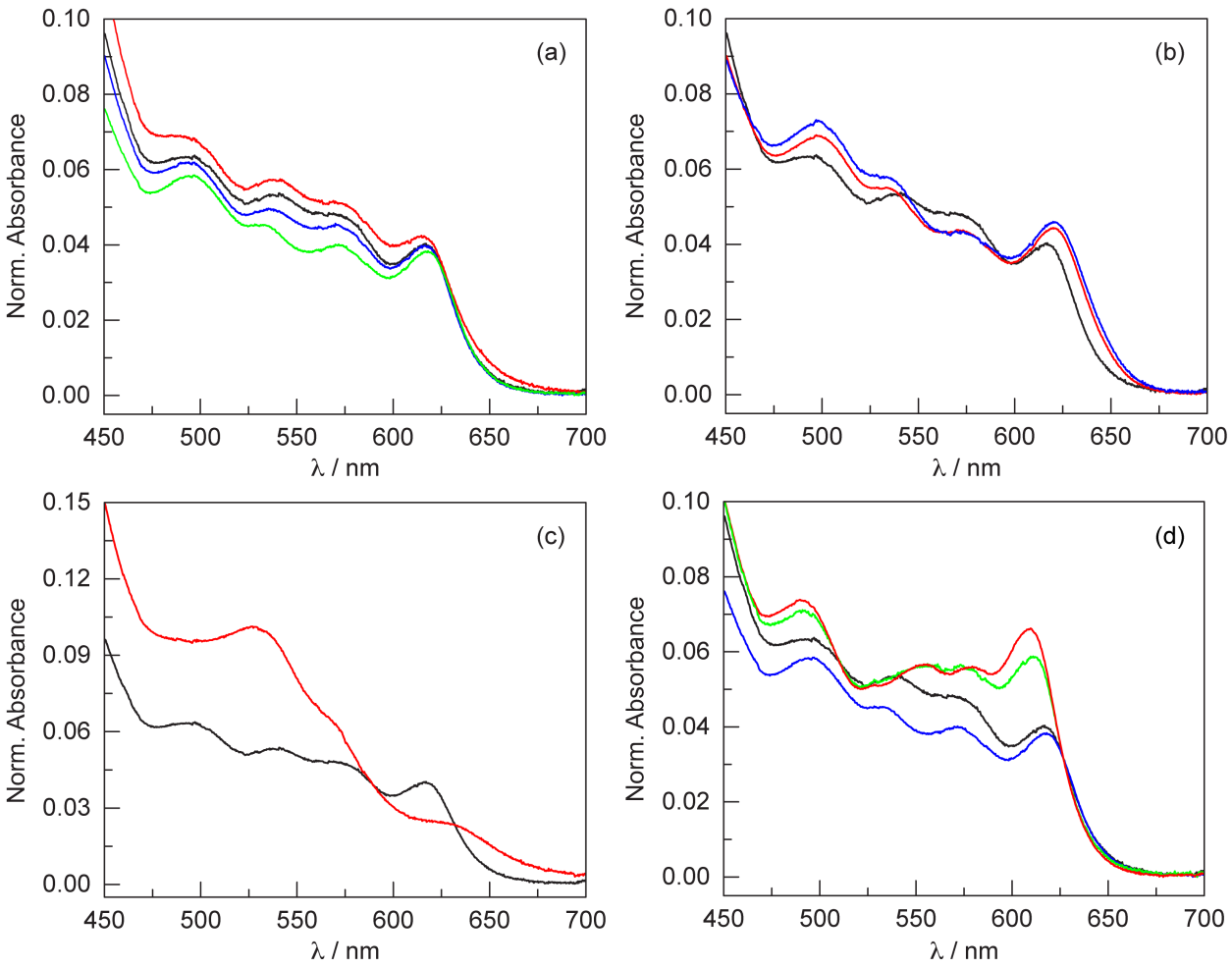

Figure S9. Comparison of Q band absorption of holo/haem *Pa* HasA variants in phosphate buffer at pH 7.4. (a) WT (—), V38C (—), S82C (—), and H83D (—) all exhibit similar spectra. (b) Both H32C (—) and H32S (—) have similar spectra that are distinct from the WT protein (—). (c) The Y75F (—) variant has an absorption spectrum that is very different from the WT protein (—) due to the loss of the haem-ligating residue. (d) H83D (—) is distinct from the other H83 variants, H83A (—) and H83S (—), and closely resembles the WT spectrum (—).

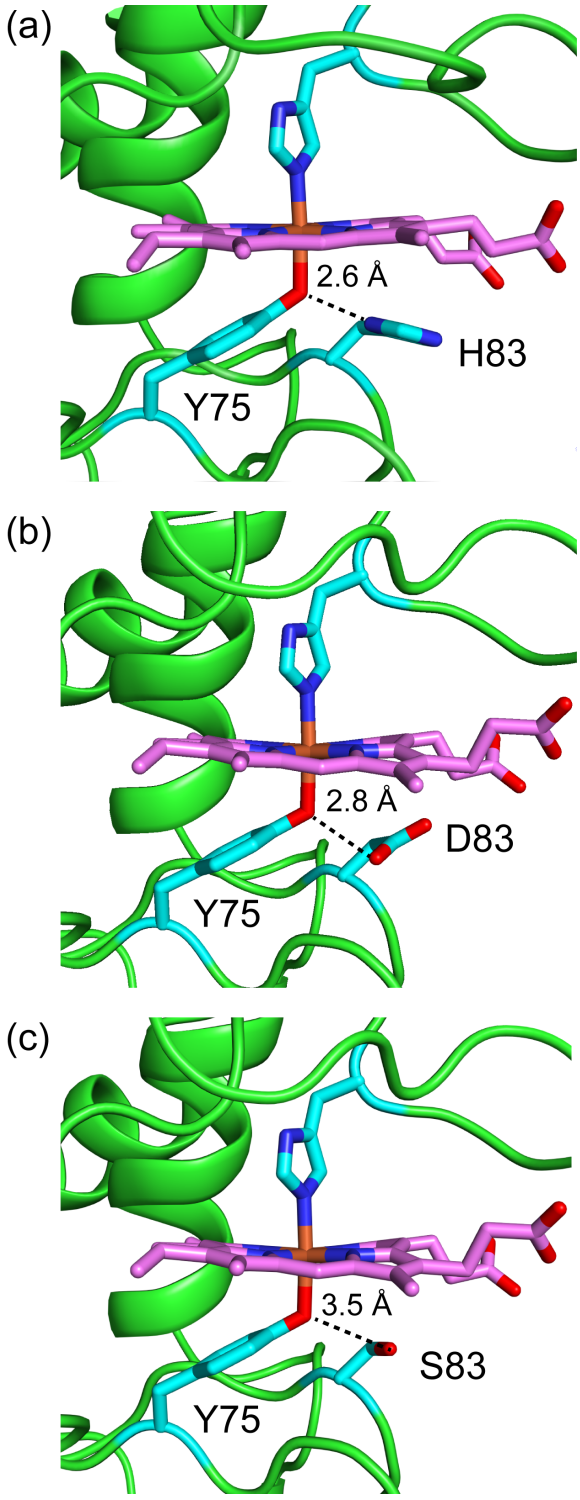

Figure S10. (a) WT HasA (PDB ID: 3ELL) showing the 2.6 Å hydrogen bonding interaction between Y75 and H83. (b) Structural model of the H83D variant that could provide a reasonable (2.8 Å) hydrogen bond between Y75 and D83. (c). Structural model of the H83S variant that exhibits a long 3.5 Å distance between Y75 and S83, precluding a hydrogen bond between these residues.

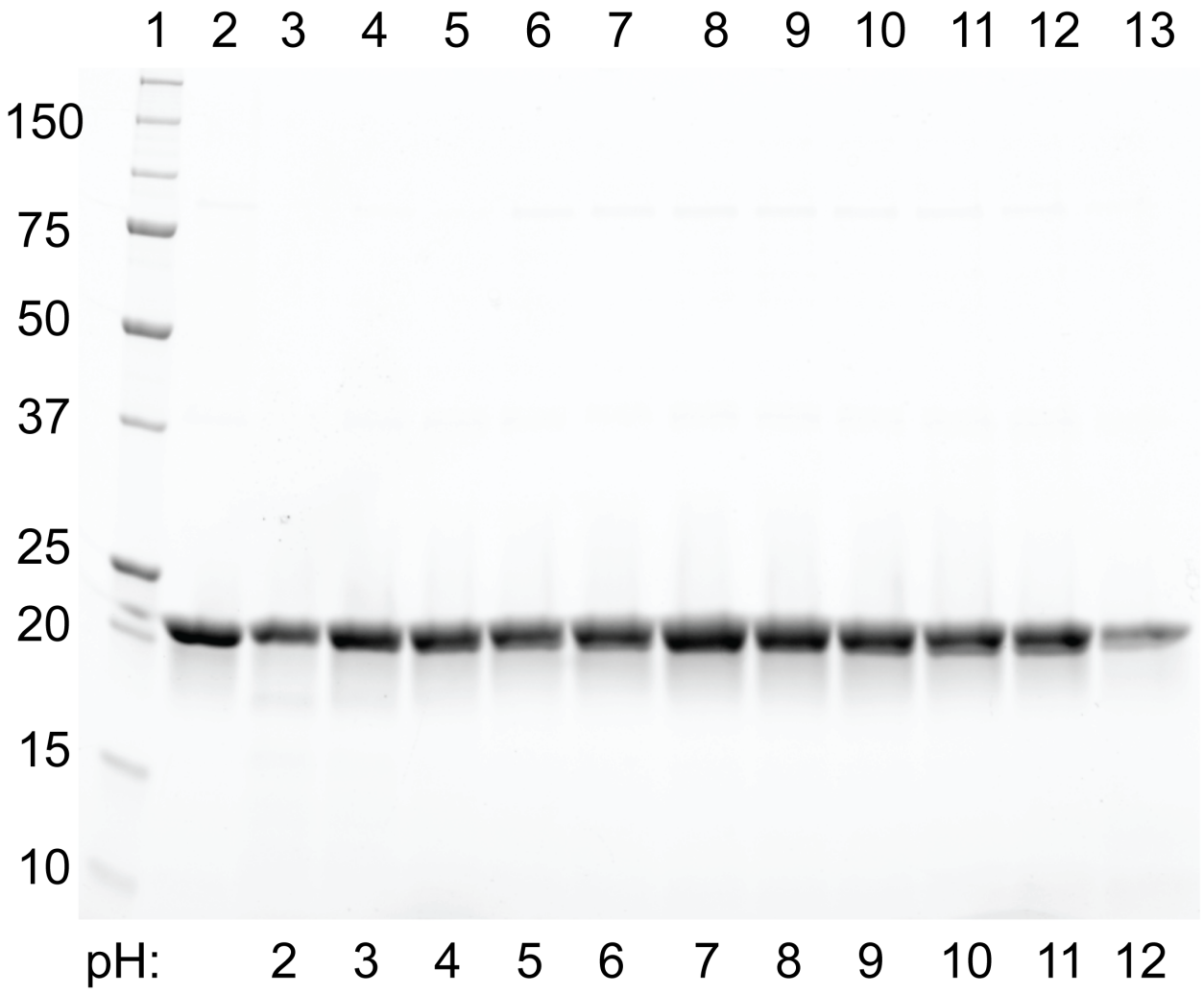

Figure S11. SDS-PAGE analysis of apo HasA stability from pH 2 to 12. For each lane, the pH values are indicated at the bottom. A ~1 mg aliquot of HasA was added to 1 mL of universal buffer (20 mM borate, 20 mM phosphate, 20 mM acetate) at a given pH and incubated at room temperature for 2 hours. The samples were centrifuged for 1 minute and the soluble protein was analysed by SDS-PAGE. Lane 1: protein ladder (molecular weights indicated in kDa). Lane 2: control sample of HasA incubated in phosphate buffer at pH 7.4 (same buffer as the protein aliquot). Lanes 3 to 13: sample of HasA incubated in universal buffer over the pH 2 to 12 range. Some protein loss is observed at the pH extremes of 2 and 12.

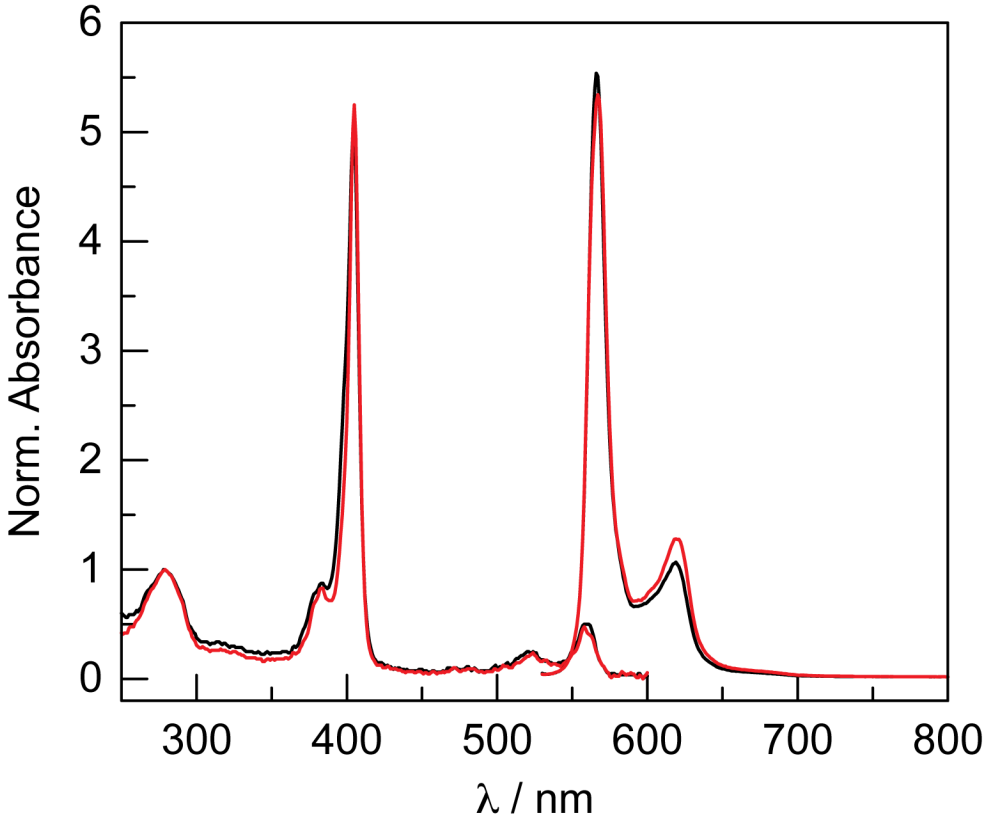

Figure S12. Absorption and emission spectra (λ_{ex} = 520 nm) of WT HasA reconstituted with **P-1** at pH 6.5 in TEA buffer (—) or pH 5.0 in universal buffer (—). For the emission spectra, the samples were absorbance-matched at the excitation wavelength: A(520) = 0.0339 ± 0.0003.

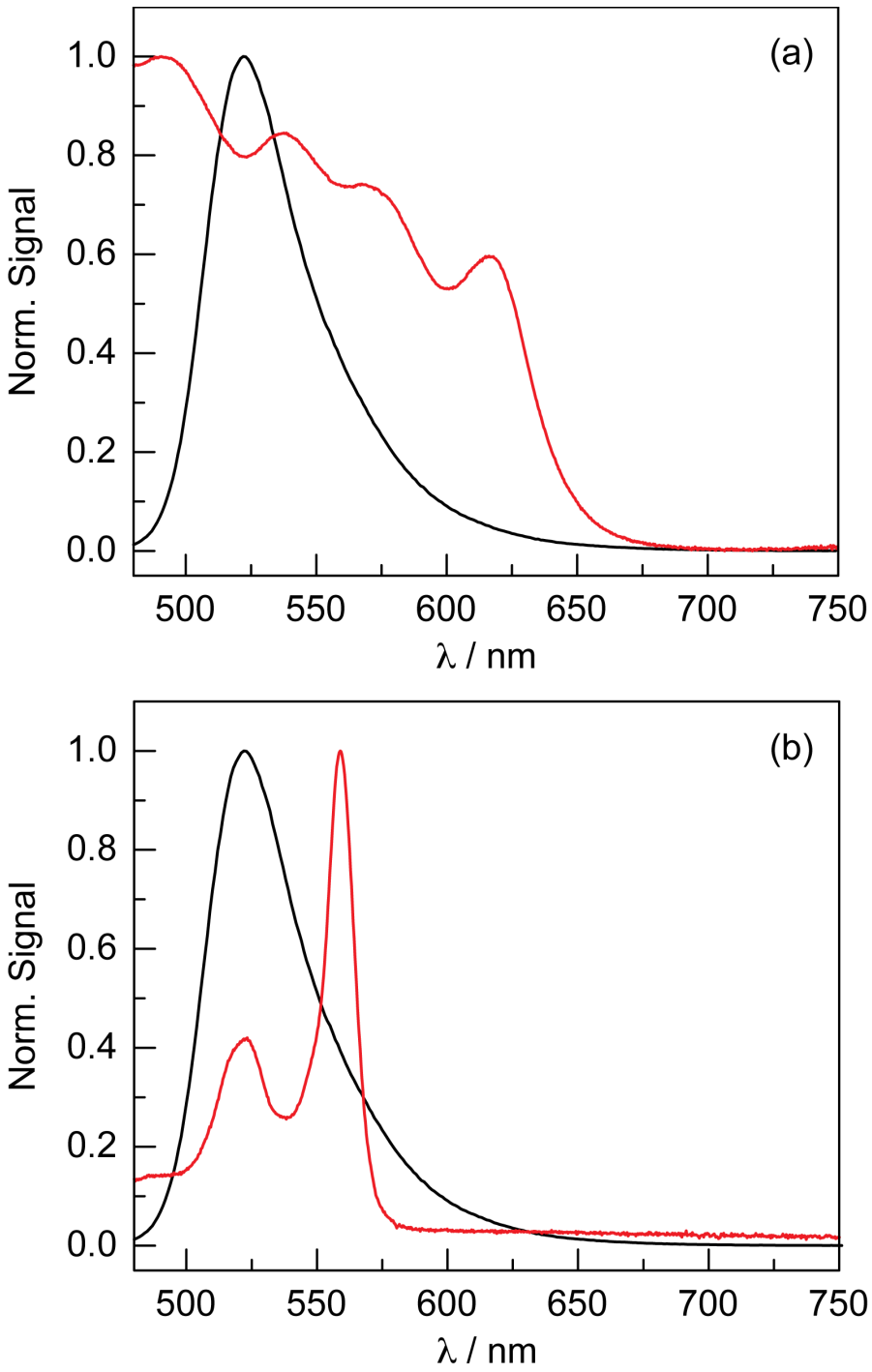

Figure S13. (a) Spectral overlap between fluorescein emission (—) and holo/haem HasA absorption (—). (b) Spectral overlap between fluorescein emission (—) and **P-1** HasA absorption (—).

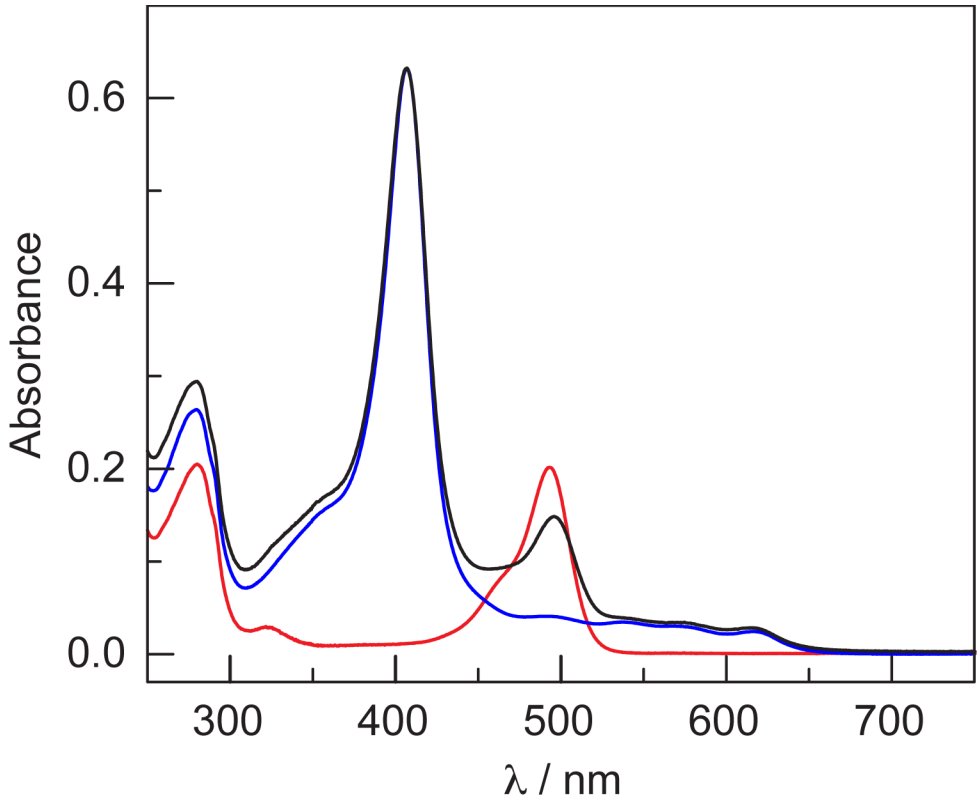

Figure S14. Absorption spectra of fluorescein-labelled apo V38C (—), fluorescein-labelled holo/haem V38C (—), and holo/haem V38C (—) in phosphate buffer at pH 7.4. Haem absorbance accounts for 44% of the 470 nm absorbance in the labelled holoprotein.

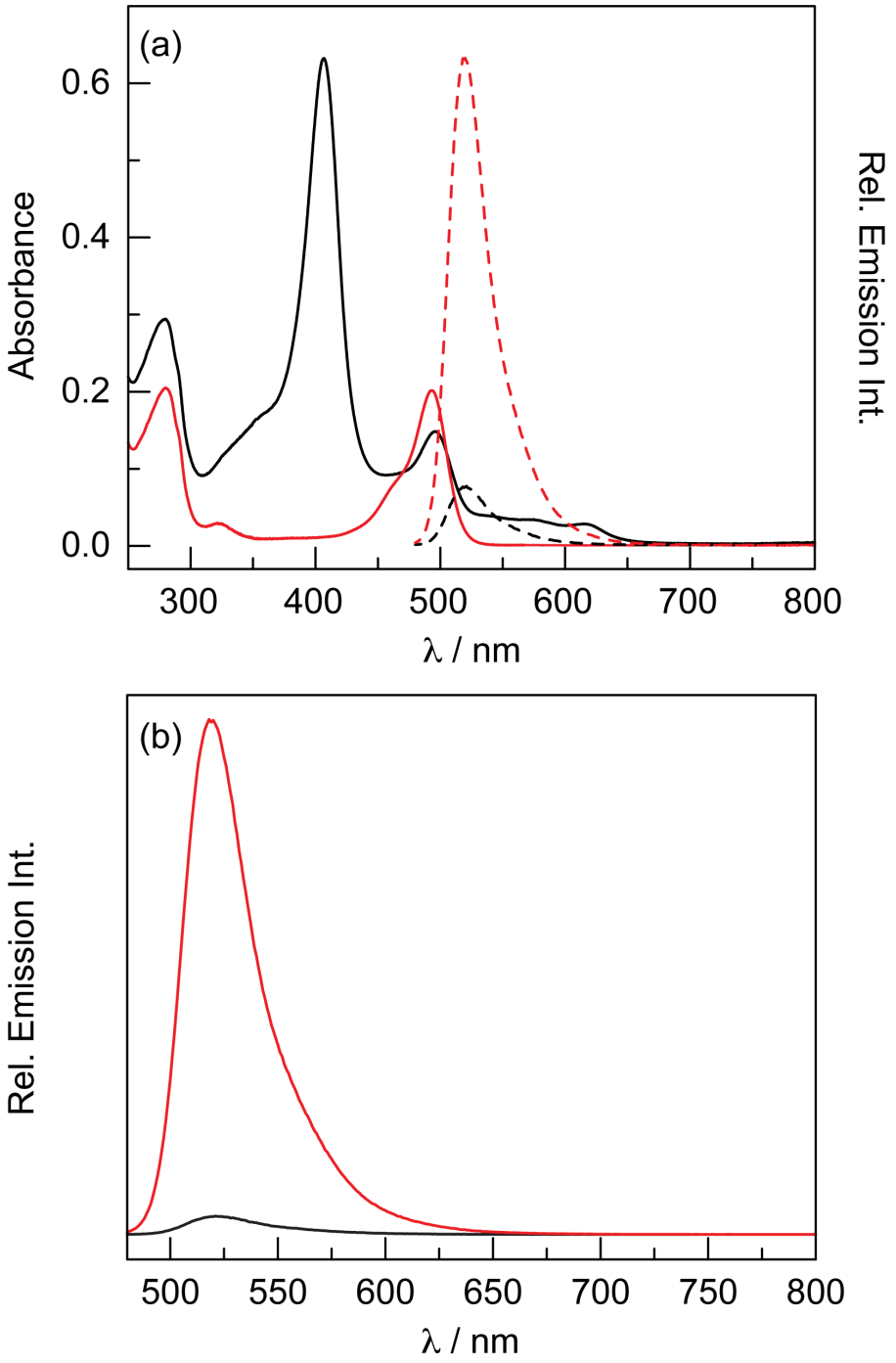

Figure S15. (a) Absorption (solid lines) and emission spectra (λ_{ex} = 470 nm, dashed lines) for fluorescein-labelled V38C HasA. Samples of apo (▬) (A(470) = 0.0478) and holo/haem (▬) (A(470) = 0.0855) V38C HasA in phosphate buffer at pH 7.4. (b) Fluorescein-labelled S82C HasA utilised for FRET analysis. The samples for the emission spectra were prepared to match fluorescein absorption at the excitation wavelength (470 nm), correcting for haem absorbance in the holo sample: apo (▬) V38C HasA (A(470) = 0.0441) and holo/haem (▬) S82C HasA (A(470) = 0.0788) in phosphate buffer at pH 7.4.

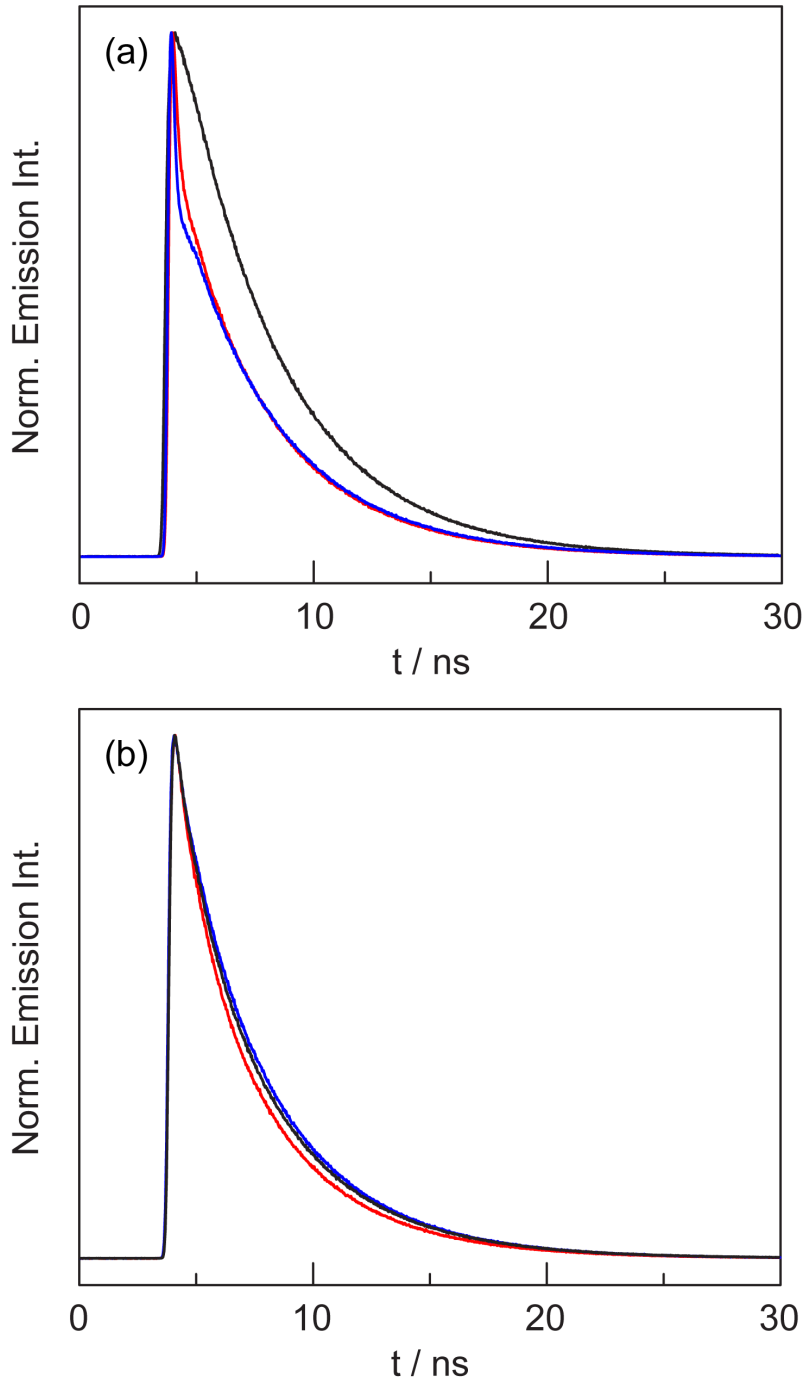

Figure S16. (a) Fluorescence decay traces showing the time-resolved fluorescein emission for labelled apo H32C (—), holo/haem V38C (—), and holo/haem S82C (—) HasA in phosphate buffer at pH 7.4. (b) Fluorescence decay traces showing the time-resolved fluorescein emission for labelled apo H32C (—), **P-1**-substituted H32C (—), and **P-1**-substituted V38C (—) HasA in phosphate buffer at pH 6.5.

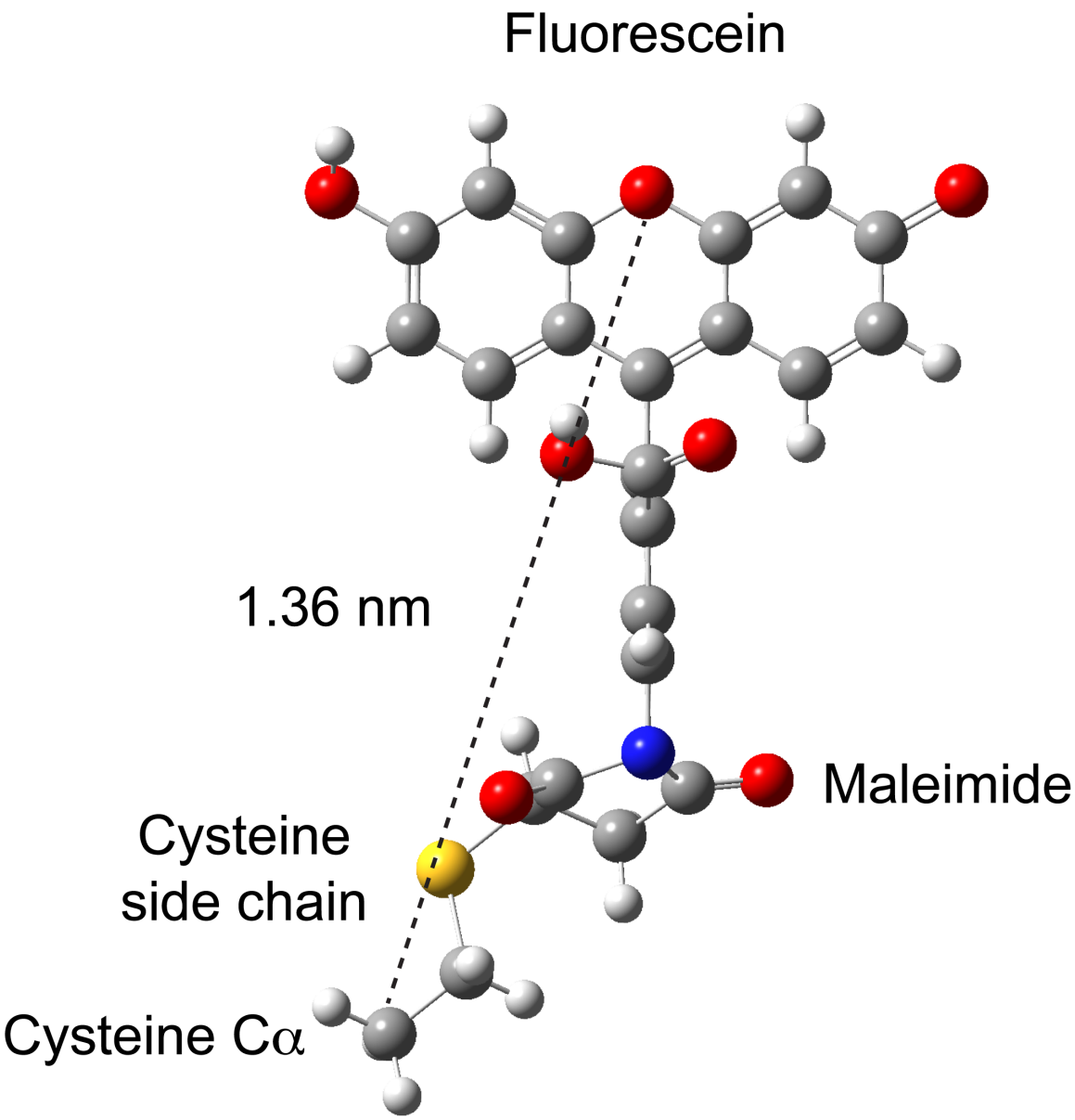

Figure S17. Three-dimensional structural model of the fluorescein maleimide conjugate generated in GaussView. This model includes the dye, the maleimide, the cysteine side chain, and the Cα of the labelled cysteine residue. The length of this structure is 1.36 nm and represents the maximum distance that could reasonably account for the dye and side chain of the labelled residue.

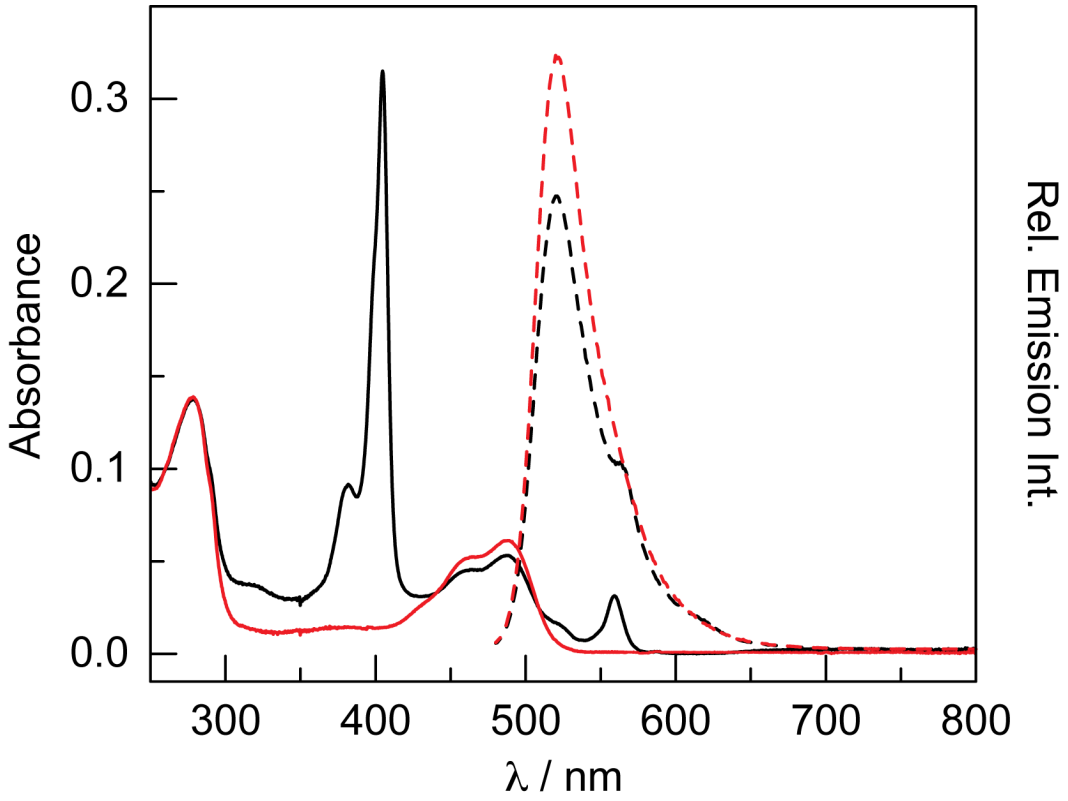

Figure S18. Fluorescein-labelled H32C HasA utilised for FRET analysis. The absorption spectra are shown with solid lines and the emission spectra (λ_{ex} = 470 nm) are indicated with dashed lines. The emission spectra are for absorption-matched samples (A(470) = 0.0523 ± 0.0003) of apo (—) and **P-1** (—) H32C HasA in TEA buffer at pH 6.5.

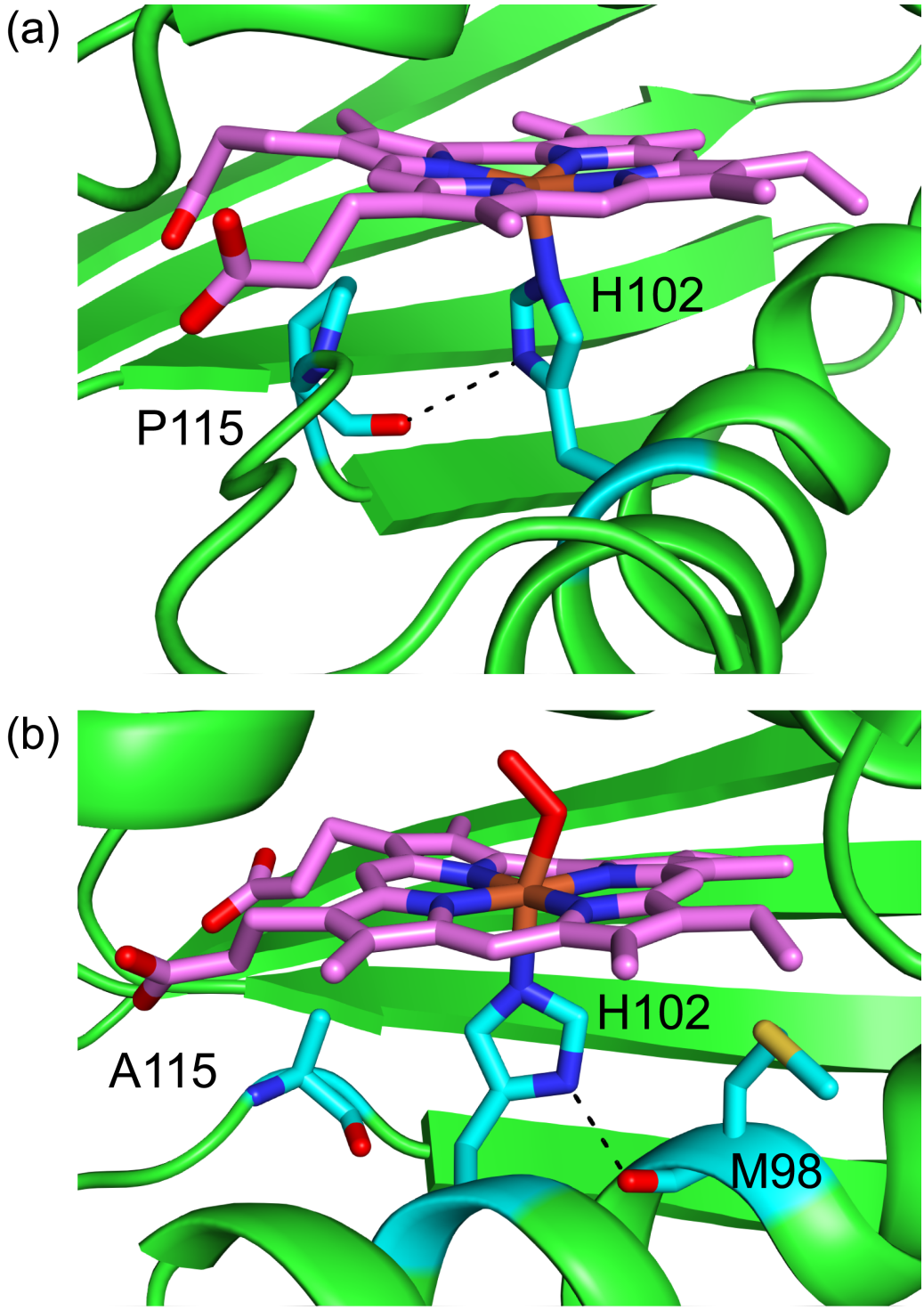

Figure S19. (a) In the unliganded WT *Cs* H-NOX structure (PDB ID: 5JRU), H102 adopts a rotamer, distinct from the O_2-bound structure, which could hydrogen bond to the backbone carbonyl of P115. (b) In the P115A structure (PDB ID: 3EEE), the backbone of the A115 loop adopts a different conformation. This rotamer of H102 hydrogen bonds with the backbone carbonyl of M98, which is similar to the WT O_2-bound structure.

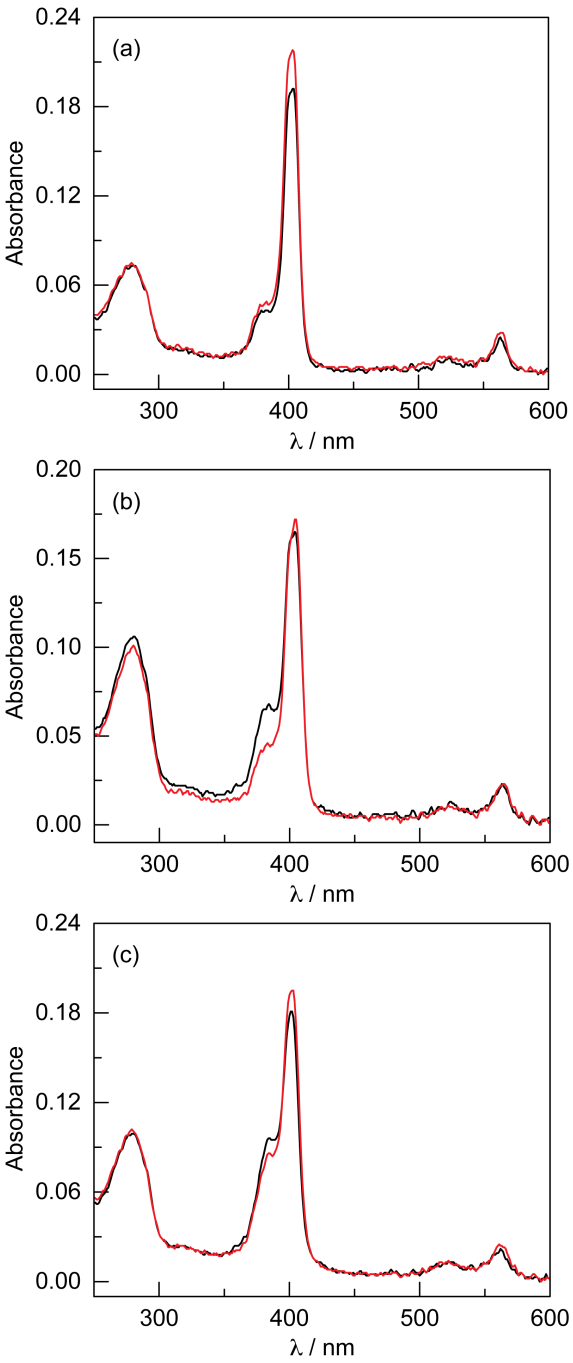

Figure S20. (a) Absorption spectrum of a freshly prepared sample of H102G H-NOX substituted with **P-1** (—). The spectrum was remeasured 1.5 hours later (—). (b) Absorption spectrum of a freshly prepared sample of H102S H-NOX substituted with **P-1** (—). The spectrum was remeasured 1.5 hours later (—). (c) Absorption spectrum of a freshly prepared sample of P115A H-NOX substituted with **P-1** (—). The spectrum was remeasured 1.5 hours later (—). The nominal spectral differences are due to variability in the NanoDrop spectrometer. Indeed, it is not possible for the amount of corrole in the sample to increase over time.