Supplementary Information

Table of Contents

1. Catalytic cycle of cyt P450 monooxygenases and structure of the thiolate heme active site.

The established catalytic cycle of cyt P450 enzymes, including the intermediates that have been identified by various spectroscopies [1, 2], is shown in Fig. 1A. Cyt P450s are heme-containing proteins with a cysteine, in its thiolate (deprotonated) form, as the axial ligand to the heme cofactor and performing the reductive activation of O_2, that results in oxygen insertion in the substrate. [3] The strong electron-donating properties of the thiolate ligand [4] allow chemical reactivities that are not observed in His-coordinated peroxidases, such as the regio- and stereoselective hydroxylation reactions [1], yet proceeding via a common high-valent intermediate, the [Fe^{IV}=O $Por^{\bullet+}$] (known as Compound I), that reacts with the substrates. [5]

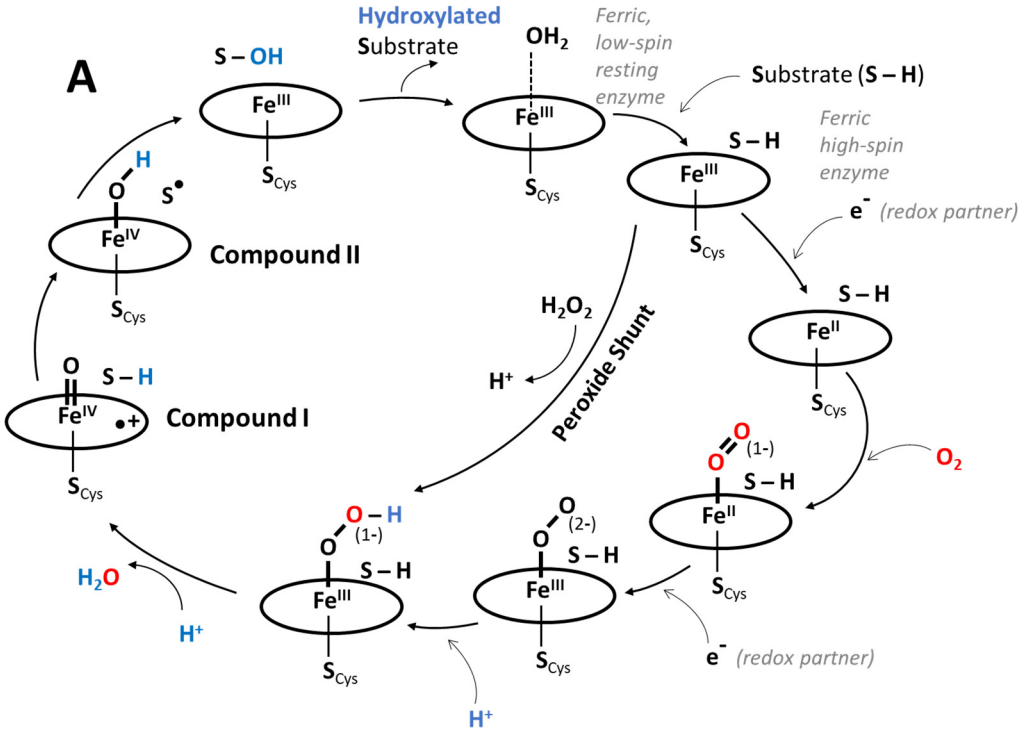


Figure S1: (**A**) Catalytic cycle of cyt P450 monooxygenases for the hydroxylation reaction with substrates. (**B**) Active site structure of cyt P450 (CYP119) from *sulfolobus solfataricus* (PDB code: 1IO8), showing the Cys and water axial ligands to the heme in the ferric resting state.

The conserved P450 structural core [6] consists of a four-helix bundle composed of three parallel helices labeled D, L, and I and one antiparallel helix E. The heme moiety is confined between the distal I helix and proximal L helix and bound to a conserved Cys in the so-called Cys ligand loop (Fig. S1B).

In the ferric resting state, the heme is hexacoordinated with a structural water as the 6th ligand, thus a low-spin species (Fig. S1B). Upon binding of the substrate (represented by SH in Fig. 1A) in the heme distal side, the structural water is displaced and the heme becomes pentacoordinated (high spin state) and with a concomitant decrease of the oxidation-reduction potential, enabling the 1e$^-$ reduction to the ferrous state by a natural redox partner. This ferrous intermediate binds O_2 and forms the oxy-ferrous intermediate (or ferric-superoxide). Another 1e$^-$ reduction by the redox partner, yields the ferric-peroxo species which upon protonation on the distal oxygen, converts to the ferric-hydroperoxo intermediate. Another protonation of the distal oxygen results in the heterolytic O-O bond cleavage, with the formation of the oxoferryl-porphyrin cation radical intermediate ([Fe^{IV}=O $Por^{\bullet+}$]), Compound I, and water release. Hydrogen abstraction from the substrate (SH) by Compound I leads to the formation of Compound II [$Fe^{IV}-OH$] and a substrate radical ($S^{\bullet}$), that further recombine to the ferric enzyme and the hydroxylated substrate (S-OH). Upon dissociation of the hydroxylated substrate, the initial resting ferric enzyme is restored.

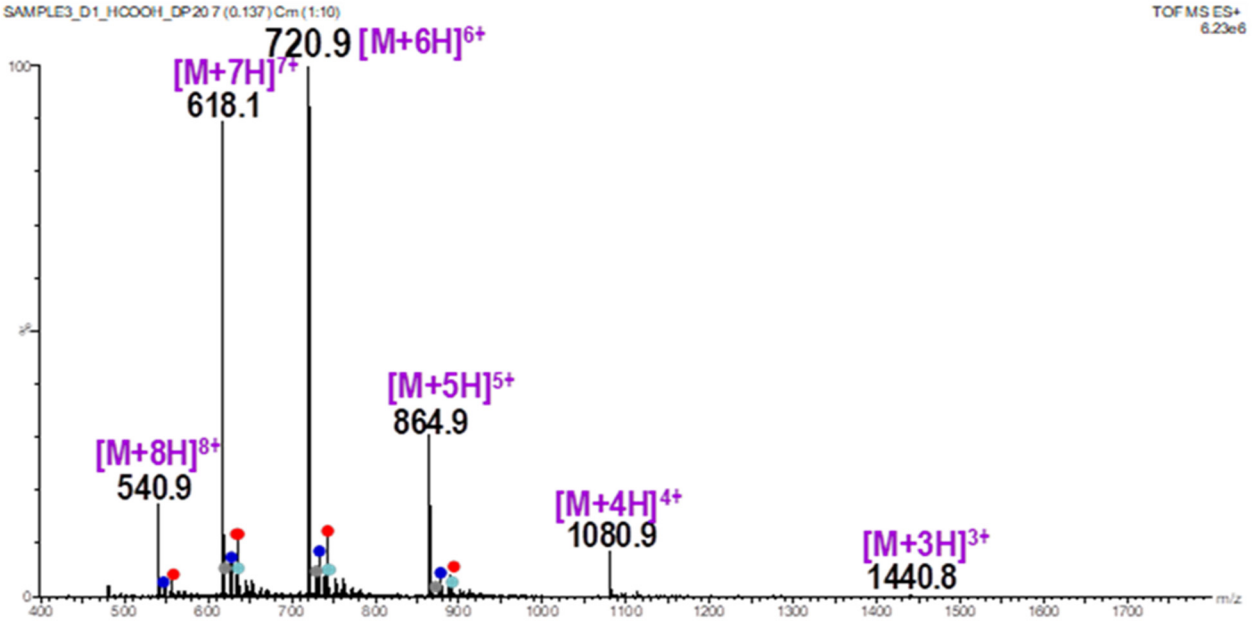
Cyt P450s in the ferric high-spin state can also turn over hydrogen peroxide (H_2O_2) in a peroxidase-like reaction, and via the so-called peroxide shunt pathway (Fig. 1A) with directly formation of the Compound I intermediate.

2. Synthesis and Purification of GRAND Peptides

All reagents were purchased from Sigma-Aldrich or Fisher Scientific and used as supplied unless otherwise indicated. All GRAND peptides were synthesized on a Biotage Initiator$^+$ Alstra peptide synthesizer, applying standard Fmoc/tBu-based protections strategies on Rink Amide MBHA resin with HBTU as coupling reagent. Peptides were cleaved from the resin using a mixture of 90% trifluoroacetic acid (TFA), 5% thioanisole, 3% ethanedithiol, and 2% anisole for 3 h. Cleaved peptides were filtered and precipitated in ice-cold diethyl ether, recovered by vacuum filtration, dissolved in 50% acetonitrile in ddH_2O, and lyophilized to dryness. Peptides were dissolved in 10% acetic acid in water and subsequently purified by reversed-phase HPLC (Agilent 1260 Infinity with Phenomenex Jupiter C18 column, 300 Å, 15 µm, 250 × 21.2 mm; solvent A: 0.1% TFA in H_2O; solvent B: 0.1% TFA in acetonitrile/water (9:1); linear gradient 0–80% of solvent B over 40 min; flow rate: 10 ml/min or a Waters 600 with DeltaPAK C18 column, 300 Å, 15 µm, 300 x 30 mm;

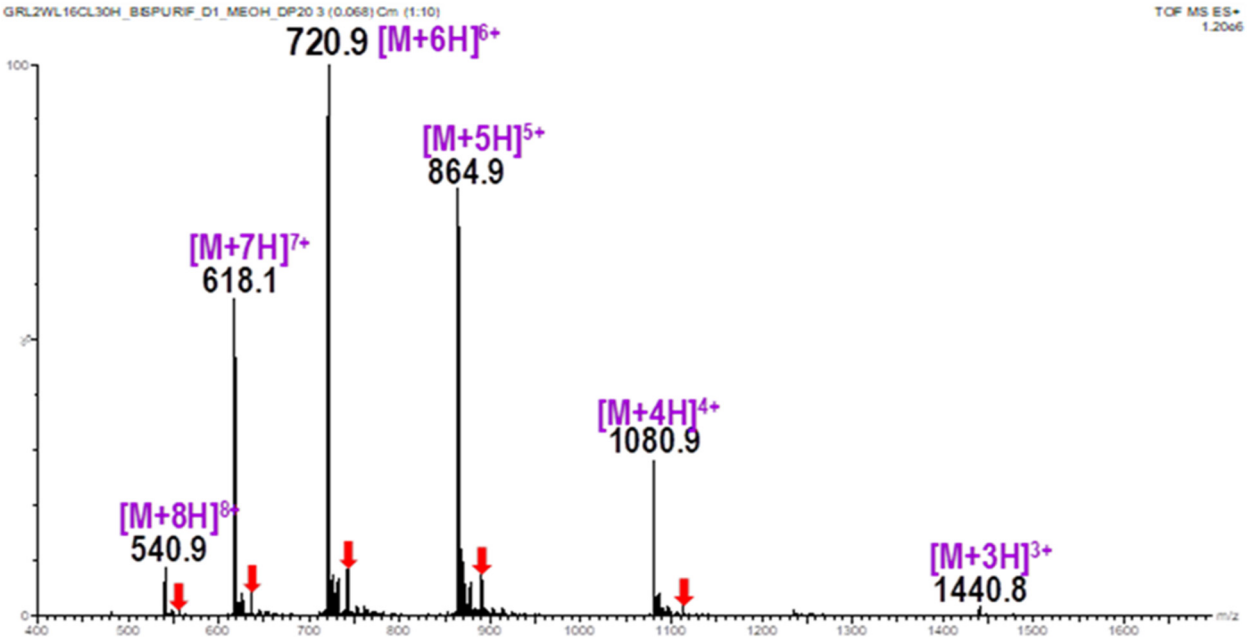
solvent A: 0.1% TFA in H_2O; solvent B: 0.1% TFA in acetonitrile/water (9:1); linear gradient 20-80% of solvent B over 35 min; flow rate: 20 ml/min). Purity was confirmed by reversed-phase analytical HPLC (Agilent 1100 with Phenomenex Jupiter Proteo column, 90 Å, 4 µm, 250 × 4.6 mm; solvent A: 0.1% TFA in H_2O; solvent B: 0.1% TFA in acetonitrile/water (9:1); linear gradient 30–60% of solvent B over 30 min; flow rate: 1 ml/min). The identity of the purified peptides was confirmed by electrospray mass spectrometry (Synapt G2 HDMS, Waters).

A typical ESI-MS spectrum of the peptides obtained with the described protocol is shown below. As illustrated by the ESI-MS spectrum of a GRW-L16CL30H sample, the purification protocol results in pure peptides that have been successfully used for metal binding studies with Cu, Hg or Zn for example. [7, 8, 9] Yet, in the case of our heme binding studies, the presence of low proportion of peptide impurities (shown in gray, red, blue and green dots) proved to prevent from obtaining mini-heme proteins with a pH-dependent heme-binding pattern that would be homogeneous and reproducible.



1st purification: The predominant species with a mass of 4319.4 Da corresponds to the GRW-L16CL30H peptide. The four other contributions (peptide impurities) with masses of 4376.4 Da (gray dots), 4390.5 Da (blue dots), 4432.4 Da (green dots) and 4447.6 Da (red dots).

Accordingly, a systematic second purification (focused on the main pure peak and using a linear gradient 35-45% of solvent B over 10 min, flow rate 10 ml/min) was performed in order to remove most of these minor impurities (see ESI-MS spectrum below), and resulted in the peptides that we used for the heme binding work. The latter had a reproducible pH-dependent heme-binding pattern, as shown by the spectroscopic studies (Fig. 3).



2nd purification: The predominant species with mass of 4319.4 Da corresponds to the GRW-L16CL30H peptide, and one minor contribution with mass of 4447.6 Da (red arrows) is left.

3. Heme complexation with GRAND peptides

Concentrated stock solutions of peptides (*ca.* 15 mM) and heme (*ca*. 5 mM) were prepared by dissolving pure lyophilized peptide in distilled water and lyophilized commercial hemin (Ferriprotoporphyrin IX chloride, BioXtra from porcine, ≥ 97.0% (HPLC), Merck) in 30 mM sodium hydroxide (pellets, semiconductor grade, 99.99% trace metals basis, Sigma-Aldrich). Concentrations were determined using the tryptophan absorbance at 280 nm (ε_{280} = 5500 M^{-1} cm^{-1}) for the GRAND peptides and the Soret band at 385 nm for heme in NaOH (ε_{385} = 58400 M^{-1} cm^{-1}), respectively. Our optimized protocol to obtain reproducible complexation of the GRAND peptides with hemin is as follows: dilute peptide from the stock solution into 250 mM TRIS-maleate buffer at pH 7.0 to the desired final concentration, then add hemin from the stock solution to a final ratio of 1:3 hemin to monomeric peptide ratio (in molar equivalents) and mix for a few seconds. The final pH of the obtained mini-heme protein using such conditions is pH 6.8. Complexation happens readily as indicated by the immediate color change of the sample (to olive-green color for the GRW-L16CL30H and GRW-L30H scaffolds and to brown color for the GRW-L16C scaffold) as compared to the grass-green color of hemin in TRIS-maleate buffer at pH 6.8 and in the absence of peptide. No changes in the absorption spectrum of the resulting mini-heme proteins were observed with longer incubation times, including overnight at 4° C. Mini-heme protein samples at higher pHs were obtained by

titration using small aliquots of concentrated NaOH and gradually increasing the pH, to allow the ligand switch His-to-Cys effect and preventing Cys oxidation at *ca.* pH ≥ 9.0.

The hemin titration into the GRAND peptides clearly showed the complexation of one heme per peptide dimer (see Fig. 2), and the UV-vis and EPR spectra of the mini-heme proteins for 6 ≤ pH ≤ 11 were very similar when using 1:2 (Fig. 3, red trace) and 1:3 (Fig. S6, red trace) heme to monomeric peptide ratio. Yet, the circular dichroism characterization showed that having extra peptide, *i.e.* using a 1:3 ratio, guarantees that the mini-heme protein remains well folded at very basic pHs, in particular for pH as high as 11 (Fig. S3B).

The high ionic strength of the TRIS-maleate buffer is crucial to maintain the pH at 7 prior to mixing with hemin, in view of the acidity of the stock solution peptide (in water) and thus preventing hemin from aggregation in acidic aqueous solution, a well-documented phenomenon for hemin. It is of note that once hemin is complexed to the peptide, such aggregation phenomenon is not observed anymore. The choice of the TRIS-maleate buffer is related to its wide pH range, making it suitable for EPR measurements in frozen solution, i.e. preventing large and buffer-induced differences in pH upon freezing the sample. We routinely used this buffer for EPR studies in heme enzymes, having pH effects that are relevant to the catalytic intermediates. [10]

4. Circular Dichroism (CD) Spectroscopy of mini-heme proteins

CD spectra were collected on a Jasco J-1500 CD spectropolarimeter at 25 °C by using 10 mm pathlength quartz cuvettes. Heme was complexed with the peptide as described above (see Section 3), except for using 10 mM sodium phosphate/borate buffer at pH 7.7. Heme incorporation and expected coordination was confirmed by UV-Vis spectroscopy (Fig. 3A) before and after the CD measurements. CD spectra were recorded for the apo peptides and the resulting mini-heme proteins upon heme complexation with the apo peptides.

The spectrum of the GRW-L16CL30H apo peptide (Fig. S2, black trace) showed a double minimum at 208 nm and 222 nm and a stronger maximum at ca. 193 nm, characteristic for a pure α-helical secondary structure. A very similar spectrum was observed for the holo peptide (Fig. S2, red trace), showing that the α-helical secondary structure is conserved in the GRW-L16CL30H mini-heme protein.

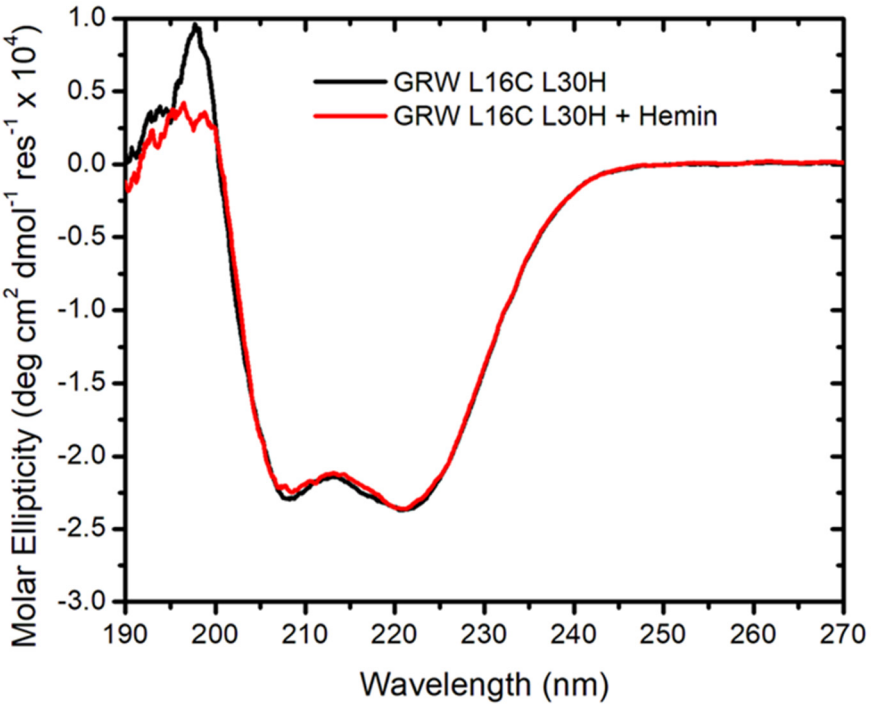


Figure S2: CD spectra of GRW-L16CL30H peptide before (black trace) and after (red trace) complexation with hemin. Samples contained 10 mM sodium phosphate/borate buffer pH 7.7, 15 μM peptide for the apo peptide, and 1:3 heme to peptide ratio when hemin was added.

The CD spectrum of the GRW-L16CL30H mini-heme protein was also recorded as a function of pH (Fig. S3A) to assess possible changes in the α-helical structure induced by pH. Each spectrum was recorded after increments of *ca.* one pH unit, showing that there were no changes in ellipticity upon heme complexation up to pH 10 (Fig. S3A, black, red and green traces), with a relatively small loss of ellipticity (15-20 %) detected at pH 11 (Fig. S3A, blue trace). All in all, the peptide remained well folded up to pH 11. Accordingly, we can rule out that the pH-dependent spectral changes observed by UV-Vis absorption spectroscopy between pH 7.5 and 11 (Fig. 3A) are induced by unfolding of the mini-heme protein and that any changes in secondary structure are minor.

Thermal denaturation of apo and heme-bound proteins were analyzed for His-bound heme (pH 6.7) and Cys/OH-bound heme (pH 10.5), and monitoring the ellipticity (θ) at 222 nm as a function of temperature, with the aim of assessing the stability of the α-helical structure of the GRW-L16CL30H mini-heme protein (Fig. S3B). The thermal denaturation behavior of the His-coordinated GRW-L16CL30H mini-heme protein at pH 6.7 (solid olive-green squares) was quite similar to the apo protein (open black squares) for the whole temperature range, even if the apo protein folds as trimers (3SCC) and converts to 4SCCs (dimers of antiparallel 2SCCs, with one heme per 2SCC) upon

heme complexation. For the hexacoordinated Cys/OH mini-heme protein at pH 10.5, a clear improvement in the thermal stability was induced by the heme complexation (Fig. S3B, solid red circles). The midpoints of unfolding (T_M), determined by fitting the experimental data of the apo (black trace) and mini-heme protein (red trace) to a two-state unfolding model (CDpal program), are (54.9 ± 1.0) °C and (64.2 ± 0.5) °C, respectively.

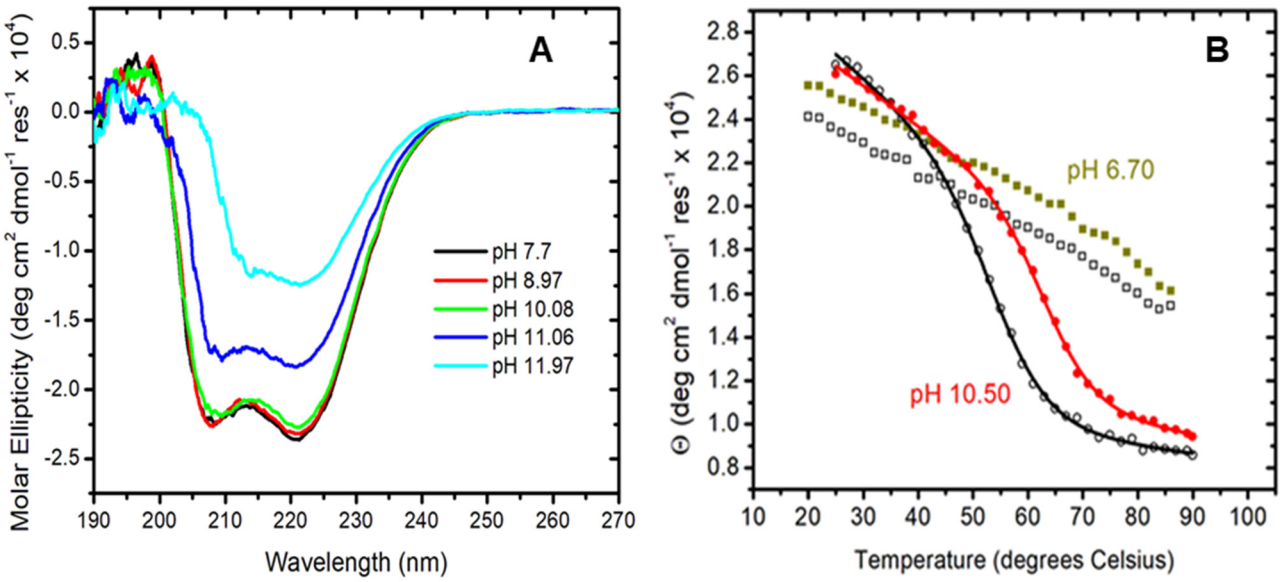


Figure S3: (**Panel A**) The CD spectrum of the GRW-L16CL30H mini-heme protein as a function of pH. Experimental conditions are the same as in Fig. S2. (**Panel B**) Temperature dependence of the CD signal monitored at 222 nm during thermal denaturation for GRW-L16CL30H in complex with heme, in the His coordinated (pH 6.70, solid olive-green squares) and Cys/OH$^-$ coordinated (pH 10.50, solid red circles) configurations. The apo peptide at pH 6.7 (open black squares) and pH 10.5 (open black circles) are shown for comparisons. The fits for the pH 10.5 experimental data are also shown.

5. Analytical Ultracentrifugation (AUC) studies of apo and holo peptides

In order to assess the oligomeric assembly states of the apo and holo (i.e. complexed with heme) synthetic proteins, we performed sedimentation velocity and sedimentation equilibrium experiments. Analytical ultracentrifugation experiments were performed at the Center for Analytical Ultracentrifugation for Macromolecular Assemblies at the University of Texas, Health Science Center (San Antonio, Texas, US). Samples were analyzed as previously described. [11] Apo GRW-L16CL30H peptide was measured for (monomeric peptide) concentrations of 21 µM and 63 µM to determine if there was a mass action effect, and at two pH values, 5.6 and 10.0 (as controls for holo peptide samples) in 50 mM potassium phosphate/borate buffer. Apo GRW-L16C peptide (in 50 mM potassium phosphate) was also measured at 14 µM and 42 µM (monomeric peptide) concentrations and pH values of 5.3 and 7.8. The holo GRW-L16CL30H peptide samples, i.e. peptides in complex

with heme, were measured at pH 6.2 and pH 10.0 in 250 mM TRIS-maleate buffer. These samples had 375 μM monomeric peptide and 125 μM hemin (final) concentrations, thus a heme to monomeric peptide ratio of 1:3 (see Section 3). Heme incorporation and pH-dependent expected coordination was confirmed by UV-Vis spectroscopy (Fig. 3A) before and after the AUC measurements.

Data were collected in intensity mode at wavelengths appropriate for the dynamic range of the instrument, given the protein concentration and choosing the appropriate absorption band for each sample. Apo peptides were tracked following their absorption band at 220 nm. The holo peptides were measured at 550 nm (Q-band spectral region of the heme, see Fig. 3A). All samples were measured at 60,000 rpm, 20 ºC temperature, and using 2-channel upon centerpieces (Beckman Coulter, Indianapolis). Data analysis was performed with UltraScan-III (ver. 4, release 5819). [12] Pseudo-absorbance data were fitted with the two-dimensional spectrum analysis, [13] subtracting time- and radially-invariant noise contributions, and fitting the meniscus and bottom position, simultaneously correcting for density and viscosity of the buffer. Data were further refined with the parametrically constrained spectrum analysis. [14] D20,w and s20,w values were transformed to molar mass, assuming a partial specific volume of 0.7599 ml/g, estimated from sequence with UltraScan.

Sedimentation profiles did not change for the different concentrations in the apo and holo peptides (data not shown), suggesting the absence of mass action effects. This is consistent with the absence of spectral differences as a function of concentrations tested in the UV-Vis absorption, EPR and cyclic voltammetry experiments. Our results show that changes in sedimentation of apo peptides were largely insensitive to pH changes in the range 5.6 to 10.0, and that all samples sedimented with a major species displaying a molar mass that is consistent with a trimer (Fig. S4A). Upon complexation with heme, both pH 10.0 (Fig. S4B) and pH 6.2 (Fig. S4C) samples showed a clear shift towards higher oligomeric forms. The exact oligomeric forms were difficult to ascertain for the GRW-L16CL30H mini-heme proteins. At pH 10.0, the AUC results suggest the presence of tetramers and hexamers (Fig. S4B), while at pH 6.2 larger s-values were observed, reflecting hexamers and dimers of tetramers (Fig. S4C). A small amount of even larger species, as well as monomeric species was detectable at both pHs.

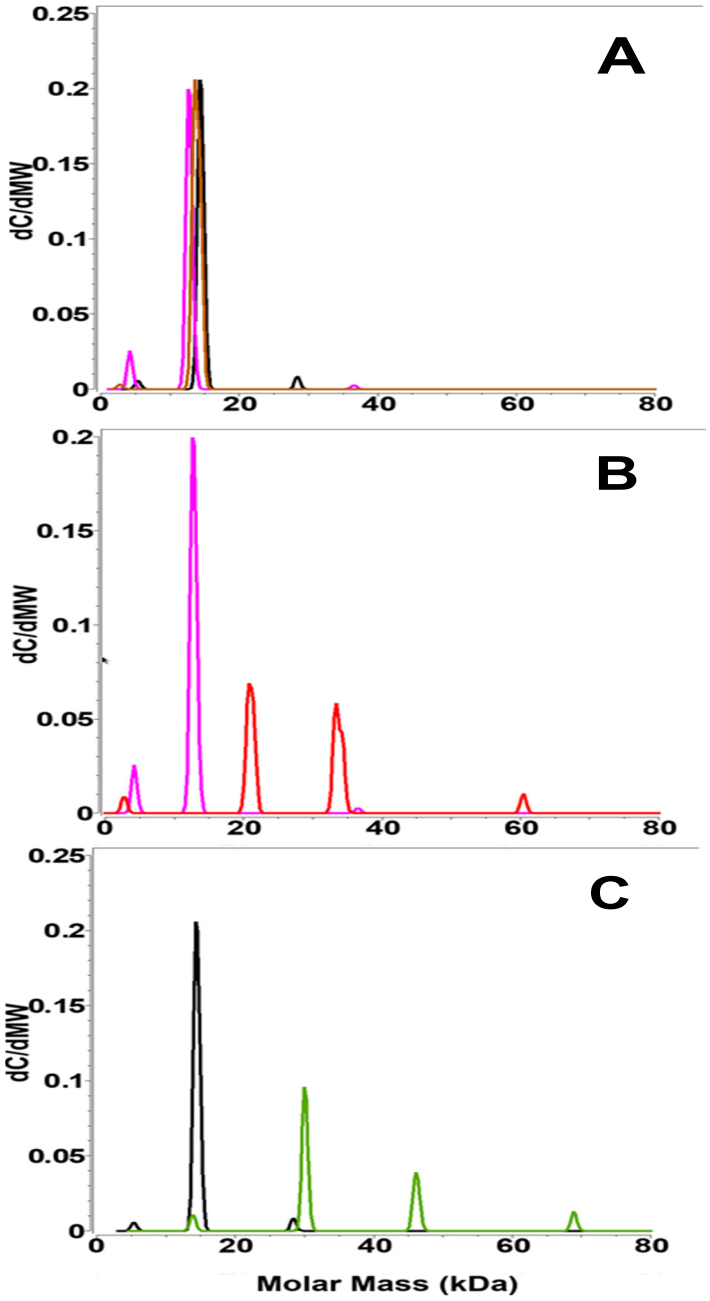


Figure S4.

(**Panel A**) Molar mass distributions obtained from the analytical ultracentrifugation experiments for the apo GRAND peptides as a function of pH. The apo GRW-L16CL30H peptide at pH 5.6 (black trace) and pH 10.0 (magenta trace), and the apo GRW-L16C peptide at pH 7.8 (brown trace) are shown. A clear trimeric arrangement is observed.

(**Panel B**): Molar mass distributions of the GRW-L16CL30H mini-heme protein at pH 10.0 (red trace). The peptides assemble into higher order structures, consistent with tetramers and hexamers. The apo peptide at the same pH is shown for comparisons (magenta trace).

(**Panel C**): Molar mass distributions of the GRW-L16CL30H mini-heme protein at pH 6.2 (olive green trace). The peptides assemble into higher order structures, consistent with hexamers and dimers of tetramers. The apo peptide at the same pH is shown for comparisons (black trace).

6. Representation of the proposed structural model for the GRW-L16CL30H mini-heme protein

To illustrate our proposed model for the GRW-L16CL30H mini-heme protein, that is a dimer of antiparallel 2SCC in complex with heme (one heme per 2SCC), a structural model (shown in Fig. 1) was created in PyMol by using the crystal structure of a *de novo* designed antiparallel 4SCC previously reported (PDB code: 2b1f). The color version is shown below (Fig. S5). In order to reflect the three distinct heme binding modes, i.e., His (Panel **A**), Cys (panel **B**) and Cys/hydroxide (Panel **C**) concluded from the spectroscopic characterization of the GRW-L16CL30H mini-heme protein

(Fig. 3), we recreated these binding sites replacing Asn16 by Cys on α-helix B (shown in blue) of the GCN4 Leucine Zipper, a four heptades repeat peptide, [15] plus Asn20 and Val23 by Glu and His, respectively on the zipper α-helix C (shown in gray).

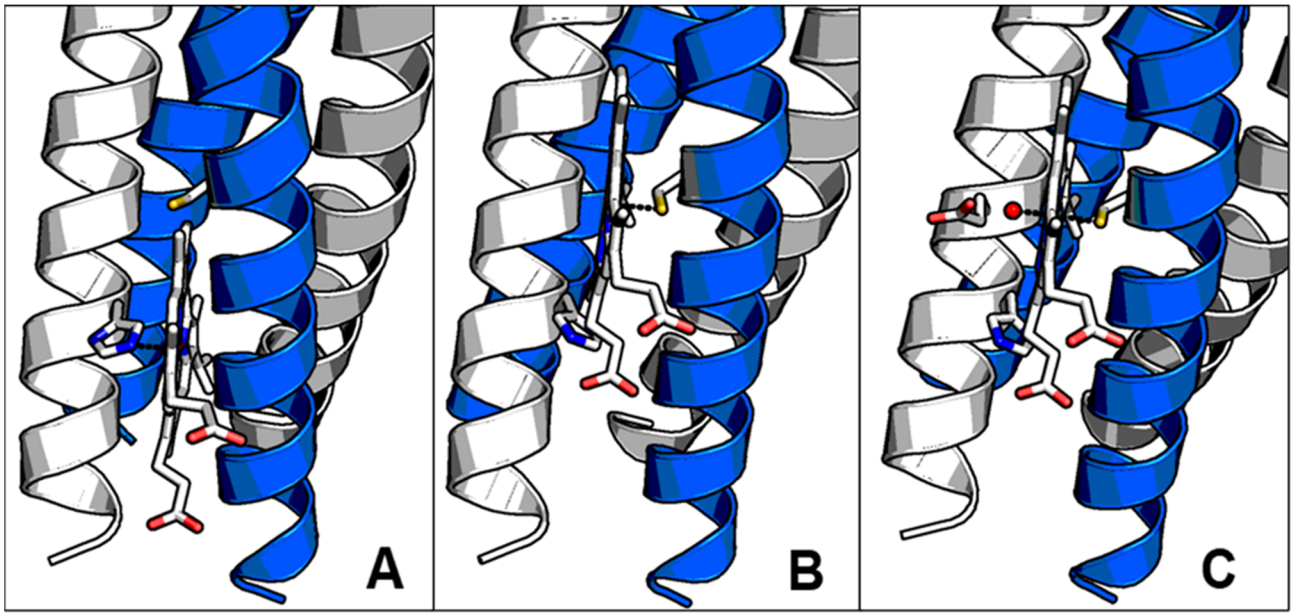


Figure S5. PyMol structural model representing the dimer of antiparallel 2SCC (with one heme per 2SCC) proposed for the folding of GRW-L16CL30H peptides upon complexation with heme. The selected pH values correspond to the three cofactor configurations, *i.e.* His-pentacoordinated heme (pH 7.0, Panel **A**), Cys-pentacoordinated heme (pH 9.0, Panel **B**) and Cys/hydroxide hexacoordinated heme (pH 10.5, Panel **C**), that were concluded from the spectroscopic characterization. Only the heme in complex with the antiparallel 2SCC being closer to the viewer is shown for clarity.

The replacements in chain C of the GCN4 Leucine Zipper do not match the actual positions of His (His30) and Glu (Glu27) in the GRW-L16CL30H construct because of the 5 heptads vs. 4 heptads difference in the monomeric peptides. Yet, by repositioning the replaced Glu and His one heptad towards the N-terminus of the GCN4 Leucine Zipper, we were able to recreate their relative positioning in the GRW-L16CL30H construct, between Chains B and C and also their relative position, in the GRW-L16CL30H construct, from the C-terminus of Chain C. The Fe-protoporphyrin IX molecule [16] was loaded into PyMol and manually positioned within the 4SCC to illustrate possible configurations to obtain the expected Cys or His coordinated heme in the GRW-L16CL30H mini-heme protein at different pH values, and according to the spectroscopic results. The resulting Fe^{III}–N_{His}, Fe^{III}–S_{Cys} and Fe^{III}–O bond lengths (2.1 Å, 2.4 Å and 1.8 Å, respectively) for the coordinated heme were fixed to those observed in the crystal structures of natural heme peroxidases and P450 monooxygenases [17]. It is important to note that this structural representation is only used to visualize the proposed folding pattern of the mini-heme proteins.

7. Low temperature 9-GHz EPR spectroscopy of the mini-heme proteins

The 9-GHz EPR spectra were recorded at 4K and 15K-25K on a Bruker EleXsys E500 spectrometer equipped with a standard Bruker ER4102 X-band resonator and a liquid helium cryostat (Oxford Instruments, ESR 900). EPR quartz tubes of 4 mm external diameter were used. All samples were measured under non-saturating experimental conditions (temperature and microwave power). Because of the different temperature saturation behavior of the high-spin and low-spin ferric heme samples, 15K (or higher) was used to record spectra in which low-spin or both low- and high-spin signals were contributing to the EPR spectrum.

Figure S6 shows the comparisons of the EPR spectra of the GRW-L16CL30H mini-heme protein in 250 mM TRIS-maleate buffer at pH 6.00 (olive-green trace) and pH 9.14 (blue trace) with those of the hemin in TRIS-maleate buffer solution at pH 6.0 (dash-dotted gray trace) and hemin in cysteine buffer solution at pH 12.00 (dashed gray trace).

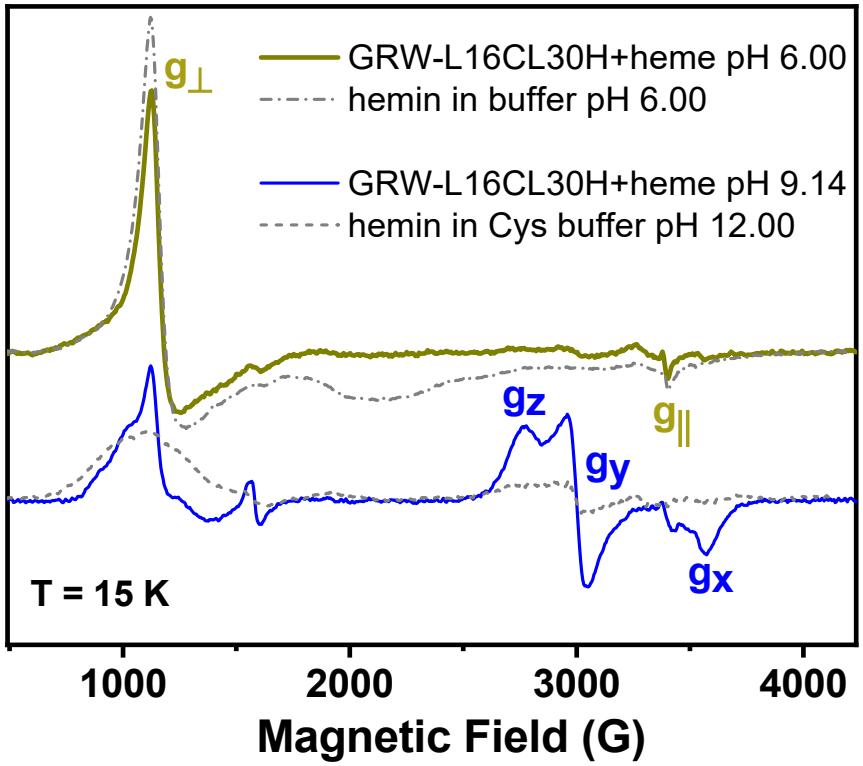


Figure S6: Comparisons of the 9-GHz EPR spectra of the GRW-L16CL30H mini-heme protein at pH 6.00 (olive-green trace) and pH 9.14 (blue trace) with hemin in buffer solution at pH 6.00 (dash-dotted gray trace) and hemin in Cys buffer pH 12.00 (dashed gray trace). Experimental conditions: 9.48 GHz frequency, 100 kHz modulation frequency. The high-spin heme EPR spectra (top) were recorded at 4K, 10 G modulation amplitude and 1.2 mW microwave power (4 scans). The EPR heme spectra (bottom) with contributions of both low-spin and high-spin signals, were recorded at 15K, 25 G modulation amplitude and 0.6 mW microwave power (8 scans).

Hemin in buffer solution at pH 6.00 shows a characteristic axial EPR signal of a ferric high-spin species with a broad $g_{\perp}$ (peak-to-trough of 166 G) and a negative absorption resonance being

part of the broad signal (dash-dotted gray trace). Such an EPR signal results from the axially symmetric environment of water molecules. When adding a high molar excess of imidazole (100 fold), the heme becomes penta-coordinated with a nitrogenous axial ligand (from imidazole) resulting in an EPR spectrum with a narrower $g_{\perp}$ component (Fig. 3B, dash-dotted trace), that is virtually identical to the mini-heme proteins spectrum at pH 6.00 (olive-green trace). In the case of hemin in cysteine buffer (300 μM hemin in 200-fold molar excess of cysteine, pH adjusted to 12) both EPR signals of the high-spin and low-spin Cys/thiolate coordinated heme contribute to the spectrum (dashed gray trace). The high-spin hemin thiolate signal has also a broad $g_{\perp}$ component, resulting from the symmetric environment. In the case of the mini-heme protein at pH 9.14 (blue trace), a rhombic distortion was observed, reflecting the influence of the heme distal-side environment. The EPR signal of the low-spin hemin thiolate shows the expected unique rhombic signal with low g-anisotropy as in the case of the mini-heme protein (blue trace).

8. UV-Vis electronic absorption spectroscopy of the mini-heme proteins

Measurements were performed at room temperature with a Cary 5E UV/Vis spectrophotometer and using Quartz cuvettes with septum, 10 mm pathlength. Mini-heme protein samples in the ferric oxidation state were degassed in an anaerobic glove box (less than 8 ppm O_2) for 20 min, in order to remove dissolved O_2 in the buffer solution. The complete removal of O_2 was tested by cyclic voltammetry (see Section 8). UV-Vis measurements, on the fully degassed samples, were performed under a constant flow of Argon gas. A molar excess of degassed sodium dithionite was used to fully reduce the samples to the ferrous state, and subsequent addition of degassed ammonium persulfate was used to re-oxidized the samples to the ferric state. The oxyferrous state at pH 6.0 was obtained by reduction of a ferric sample degassed with an Argon flow only (no glove box).

A crucial aspect for the synthetic mini-heme proteins is to conserve the heme axial ligand upon changes in oxidation states, as happens in natural heme-containing enzymes. Hence, spectral changes due to the ferric to ferrous oxidation states were followed by UV-Vis absorption spectroscopy (Fig. S7). The GRW-L16CL30H mini-heme protein (48 μM monomeric peptide and 16 μM heme final concentrations) at pH 10.6 (Fig. S7 top, red trace) showed the characteristic split Soret band (360 nm and 420 nm) and the β and α bands (540 nm and 570 nm, respectively) of a low-spin thiolate heme in the ferric oxidation state (Fig. 3A). Upon anaerobic reduction of the sample with a molar excess of sodium dithionite, a single Soret band at 425 nm and two asymmetric Q-bands (visible spectral region) at 526 nm and 556 nm were observed, indicative of a low-spin 6-coordinated

ferrous heme (Fig. S7 top, dotted trace). This spectrum is virtually identical to the previously reported ferrous DGCR8 (DiGeorge critical region 8) heme protein, for which it was proposed that the heme is coordinated by neutral Cys axial ligand(s), [18] and also to the case of a mutant myoglobin H93G with *bis*-thioether ligands. [19] Upon anaerobic re-oxidation of the sample, using ammonium persulfate (Fig. S7 top, gray trace), the ferric low-spin thiolate heme spectrum was recovered. The broadening (and shift of the maximum to 410 nm) of the Soret band is due to the relatively higher contribution of the high-spin Cys-coordinated species (see Fig. 3A, blue trace) revealed by the ligand-to-metal charge transfer (LMCT) band at 625 nm (Fig. S7 top, *Inset*). The change in relative equilibrium of high- and low-spin ferric heme forms of the sample, upon the reduction/re-oxidation cycle, can be easily explained by a repositioning of the water/hydroxy molecule relative to the heme iron upon re oxidation. All in all, the absorption experiments confirm that the Cys16/thiolate ligation is conserved in the ferric/ferrous redox cycle.

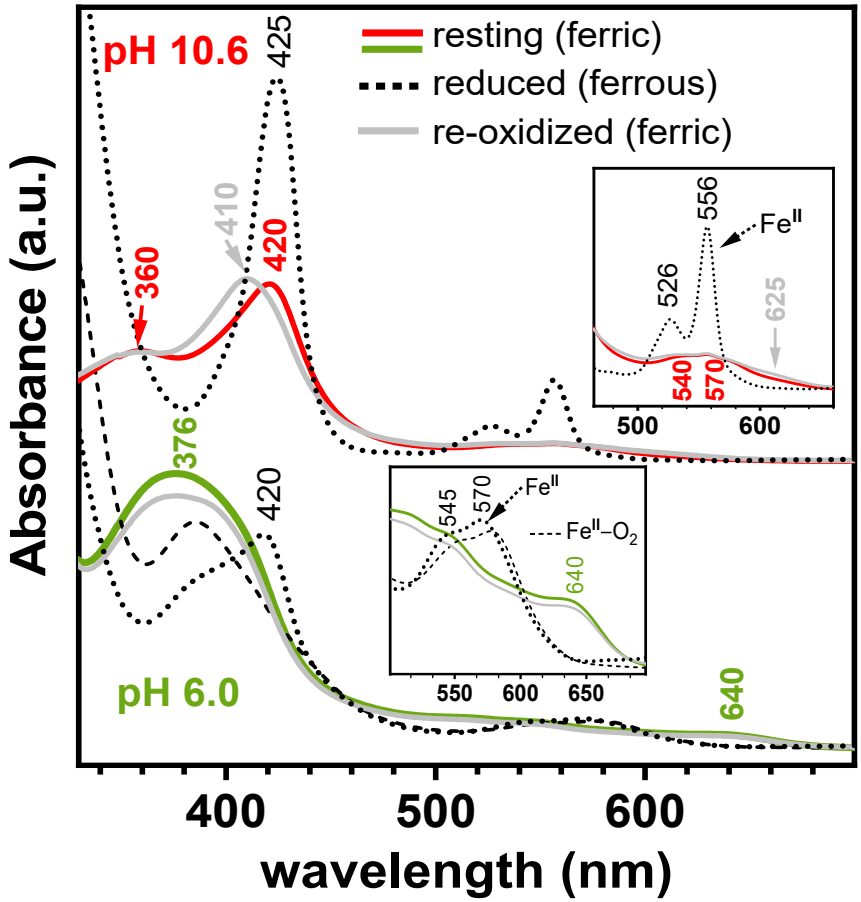


Figure S7: UV-Vis absorption spectra of the GRW-L16CL30H mini-heme protein in the ferric oxidation state at pH 10.6 (Cys-coordinated heme, red trace) and pH 6.0 (His-coordinated heme, olive-green trace) recorded at room temperature and in strict anaerobic conditions (see text). Samples were reduced with dithionite (dotted traces) and subsequently re-oxidized with ammonium persulfate (gray traces). The spectrum of the oxyferrous intermediate of the His-coordinated heme at pH 6.0 (dashed trace) is also shown.

The spectra of the GRW-L16CL30H mini-heme protein (96 μM monomeric peptide and 32 μM heme final concentrations) in the ferric oxidation state at pH 6.0 (Fig. S7 bottom, red trace) is characteristic of a 5-coordinated heme, with an LMCT band at 640 nm typically observed in His-coordinated peroxidases. It is of note that, in peroxidases and KatGs, the Soret band is centered at ca. 400-406 nm while in the mini-heme protein, the broad Soret band shows a maximum at 376 nm. Such a blueshift of the Soret band of the ferric spectrum has been previously reported for His-coordinated peroxidases at acidic pHs, and has been correlated to a longer heme Fe^{III} to N_{His} bond length using Resonance Raman spectroscopy. [20] The ferrous and oxyferrous spectra of the pH 6.0 sample (Fig. S7 bottom, dotted and dashed black traces, respectively) show bands in the Q-band spectral region reported for His-coordinated peroxidases. Not surprisingly, and consistent with the observed shift of in the ferric oxidation state, the Soret band of the ferrous state (Fig. S7 bottom, dotted black trace) is blue-shifted as compared to ferrous peroxidases. Upon re-oxidation, the absorption spectrum (Fig. S7 bottom, solid gray trace) was identical to the initial ferric spectrum, confirming that the heme ligation to His30/imidazole is also conserved through the ferric/ferrous redox cycle.

9. Cyclic voltammetry of the mini-heme proteins

Cyclic voltammetry (CV) was performed using a PAR potentiostat controlled by Echem software. All experiments were performed in an anaerobic chamber (less than 8 ppm O_2). The electrochemical cell was equipped with three electrodes, a pyrolytic graphite (PG) working electrode, a platinum wire as an auxiliary electrode and an Ag/AgCl electrode (saturated NaCl solution) as the reference electrode. The conversion factor to the normal hydrogen electrode (NHE) for measured potentials is +210 mV. The PG electrode surface (S = 0.07 cm^2) was renewed by polishing with fine sand paper (P1200), and then briefly sonicated to remove free carbon particles. The membrane electrode configuration was used to entrap 2 μl of mini-heme protein samples in a thin layer between the electrode and a dialysis membrane of suitable cutoff. [21]

Strict anaerobic conditions, required to avoid fast self-reoxidation of the mini-heme proteins by the dissolved O_2 in the buffer solution, were obtained by using the anaerobic glove box and purging the electrochemical cell with argon during CV measurements. The effect of dioxygen on the voltammograms of the mini-heme proteins was studied by adding increasing volumes of O_2 gas in the electrolyte with a gas tight syringe (Fig. 4 and Fig. S8). The GRW-L16CL30H mini-heme protein at pH 10.1 (Fig. 4) showed a pH-dependent cathodic wave, with no anodic counterpart, at a mid-wave potential of *ca.* 200 mV more positive than the midpoint potential of the anaerobic Fe^{III}/Fe^{II} redox

couple (-460 mV vs Ag/AgCl). Its amplitude increased with increasing O_2 concentrations (see arrow in Fig. 4). The occurrence and redox potential of this wave agrees well with the catalytic reduction of dioxygen, upon binding of O_2 to the ferrous mini-heme protein. [22] The current is limited by O_2 depletion in the diffusion layer. Hence, the anaerobic Fe^{III}/Fe^{II} couple was recovered with similar magnitude, both anodic and cathodic, when the presence of O_2 diminished.

Figure S8 (Panel A) shows the voltammograms of the GRW-L16C mini-heme protein (90 µM heme concentration) at pH 9.0 in anaerobic conditions (solid trace) and aerated buffer (dash dotted trace). The Fe^{III}/Fe^{II} redox couple is characterized by a well-defined redox wave at -390 mV vs Ag/AgCl (solid trace). The presence of O_2 in the buffer solution induced a cathodic wave observed at a more positive potential (*ca.* -190 mV vs Ag/AgCl), reflecting the catalytic reduction of O_2 driven by the mini-heme protein. The reduction of dioxygen by the PG electrode (no protein present) is shown for comparisons (Panel B). Direct reduction of O_2 at the PG electrode proceeds through a slow irreversible electrochemical process with a mid-wave potential at *ca.* -300 mV vs Ag/AgCl. (Panel A), clearly showing the difference with the catalytic reduction driven by the mini-heme protein (Panel B).

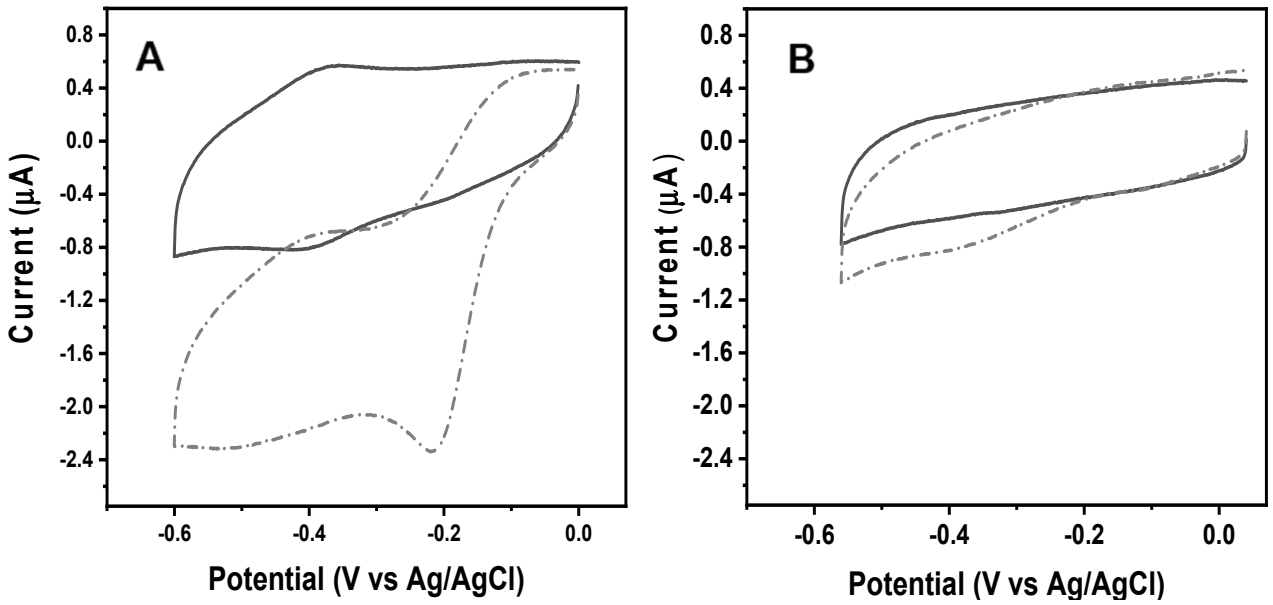


Figure S8: (Panel **A**) Cyclic voltammograms of the GRW-L16C mini-heme protein (270 µM monomeric peptide and 90 µM heme final concentrations) in 250 mM TRIS-maleate buffer at pH 9.0 measured in anaerobic conditions (solid trace) and after O_2 addition (dash dotted trace). 20 mV/s scan rate. (Panel **B**) Cyclic voltammograms of the PG electrode in the membrane configuration, 250 mM TRIS-maleate buffer at pH 9.0, measured in the absence (solid trace) and presence of O_2 (dash dotted trace). 20 mV/s scan rate.

The stability of the mini-heme proteins, related to the heme coordination for the whole pH range, was confirmed by the absence of changes in redox potentials during several

reduction/oxidation cycles on the PG electrode (data not shown). The heterogeneous electron-transfer rate constant was determined to be 9 s^{-1} for the GRW-L16C mini-heme protein at pH 10.0, based on the variation of the redox peaks as a function of scan rates and according to Laviron's formalism. [23] A very similar value was reported for cyt P450 BM3 monooxygenase at pH 7 in SDS (10 s^{-1}), [24] both being lower than the electron-transfer rate constant reported for cyt P450 BM3 at pH 7.4 in DDAB films (221 s^{-1}). [22]

10. Spectroscopic characterization of the GRW-L16C and GRW-L30H mini-heme proteins

Fig. S9 shows the pH-dependent UV-Vis absorption and EPR spectra of the GRW-L16C mini-heme protein, that was obtained using the GRAND scaffold with a single mutation (Leu16 to Cys) as putative heme binding site. For pH ≥ 9.0, both the electronic absorption (Panel A) and the EPR (Panel B) spectra are very similar to those of the GRW-L16CL30 mini-heme protein (see Fig. 3 and corresponding description). Accordingly, they confirm our assignments of the heme being pentacoordinated to Cys16 at pH 9.0 (blue traces) and converting to a low-spin hexacoordinated form (with an OH trans to Cys16) at pH ≥ 10.5. It is of note that the HS to LS equilibrium of the Cys16 coordinated species is not the same as in the mini-heme protein with only the Cys16 binding site, and the full conversion to the LS conformation was observed at a higher pH value. This difference between the two mini-heme proteins possibly reflects changes in positioning of the heme and/or packing of the helices in the dimer, given the two (Cys16 and His30) vs one (Cys16) mutations. At neutral pH, and in the absence of an amino acid residue that could act as a putative axial ligand to the heme, no specific coordination of the heme to the mini-protein was expected. Yet, both UV-Vis absorption and EPR spectra (Fig. S9A and S9B, brown traces) differ from those of hemin in buffer solution (dash-dotted gray traces), and are consistent with heme in interaction with the peptide dimer (*i.e.* accommodated), with no nitrogenous axial ligand. It is of note that differences between accommodated and His-coordinated heme are small but measurable, in particular for the absorption spectrum (compare Fig. S9A brown trace and Fig. 3A, olive-green trace), in which the maximum of the Soret and LMCT bands are clearly not the same. Also, the axial EPR signal of the GRW-L16C at pH 7.0 showed a peak-to-trough width of the $g_{\perp}$ resonance clearly different to that of hemin in buffer solution (Fig. S8B) and the heme coordinated to GRW-L16CL30H dimer (Fig. 3B) at the same pHs.

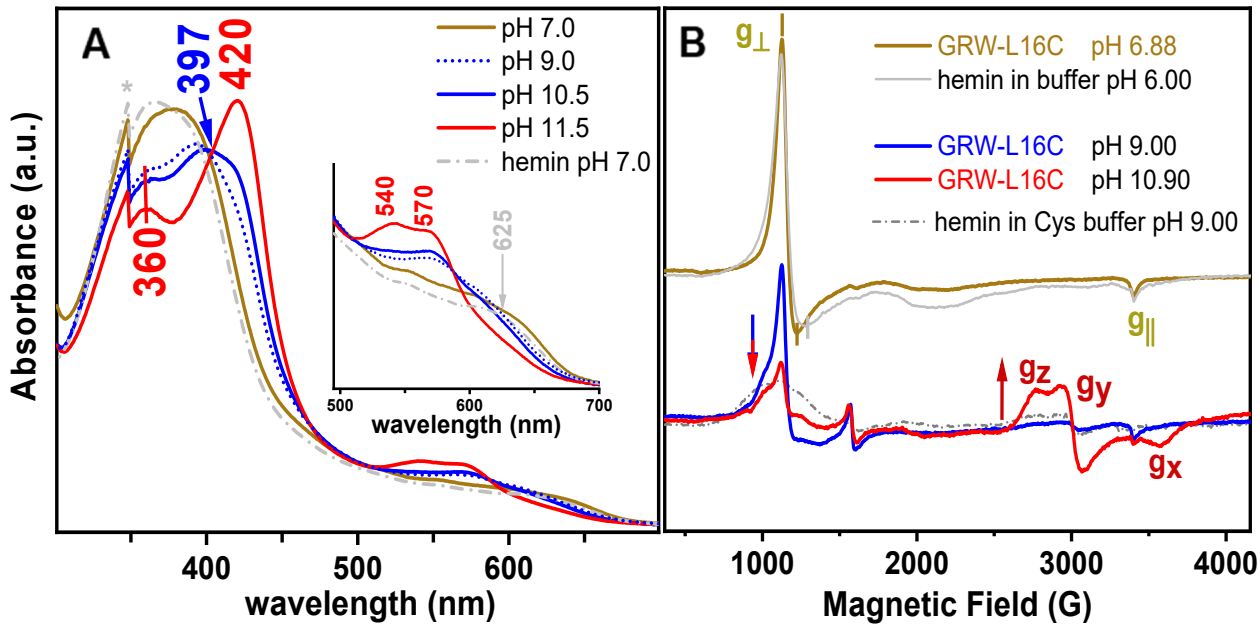


Figure S9: UV-Vis electronic absorption spectra (Panel A) and 9-GHz EPR spectra (Panel B) of the GRW-L16C ferric mini-heme protein (300 µM and 30 µM heme concentration, respectively) as a function of pH. Experimental conditions for EPR measurements are the same as in Fig. S6.

The characterization of the GRAND scaffolds with the Leu30 to His single mutation, the GRW-L30H mini-heme protein, as a function of pH and using EPR spectroscopy (Fig. S10) showed partially unanticipated results. The EPR spectra for the samples within the pH range 6.50 ≤ pH ≤ 8.50 (brown and grass green traces) are characteristic of heme iron in the ferric (Fe^{III}) high-spin state, and virtually identical to that of penta-coordinated hemin with imidazole as axial ligand (black dash-dotted trace). Hence, considering the peptide sequence, we can confidently conclude that a penta-coordinated heme with His30 as axial ligand exists in this pH range, as shown in the case of the GRW-L16CL30H mini-heme protein for the same pH range (Fig. 3B). Also, within the pH range 6.50 ≤ pH ≤ 8.50, the ferric/ferrous redox cycling showed that the heme ligation is conserved (see below, Fig. S11). The unexpected behavior was related to the fact that at pHs lower than 6.5 and higher than 8.5, the EPR spectra (black and dark green traces, respectively) were those of hemin in buffer solution (dashed gray traces), that is in a symmetric (water) environment. Such a behavior can be reconciled with the occurrence of a less constrained environment of the heme in the artificial mini-protein (no "heme cavity" as in natural heme-containing enzymes), with the absence of an H-bond to the δ-nitrogen of the His axial ligand or differences in the pK values of the His induced by the microenvironment. [25]

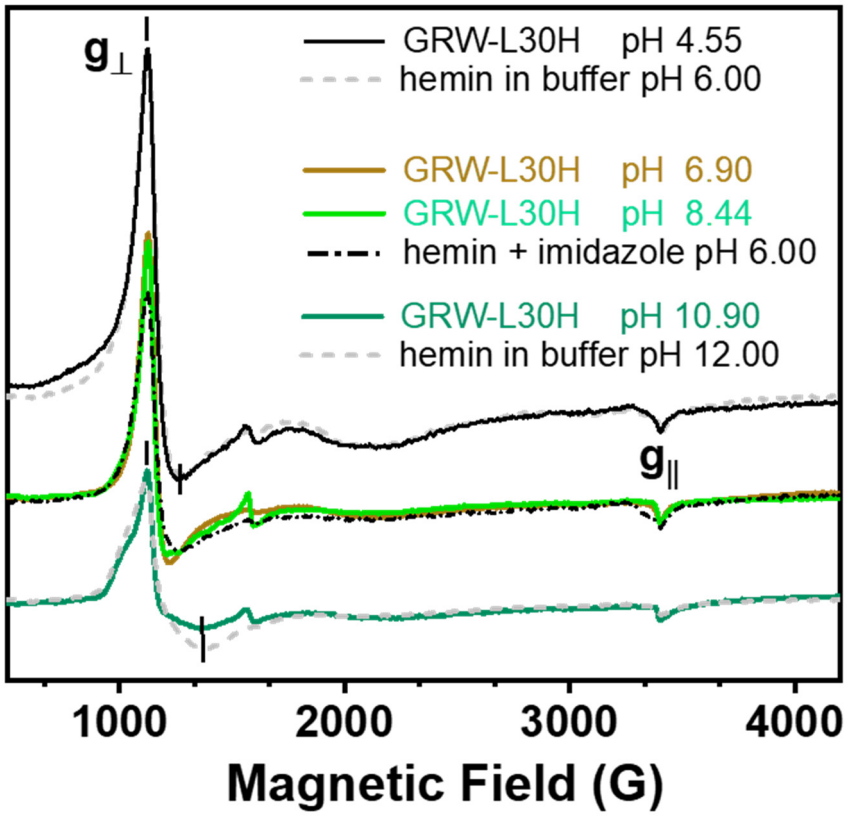


Figure S10: 4K 9-GHz EPR spectra of the ferric GRW-L30H mini-heme protein (300 μM heme concentration) for selected pHs within the range 6.00 ≤ pH ≤ 11, that are representative of the heme coordination pattern. The spectra of hemin in buffer at pH 6.0 and pH 12.0 (gray dashed traces), and hemin in imidazole buffer pH 6.00 (dash-dotted black trace) are shown for comparisons.

Taken together, our results on the GRW-L16C and GRW-L30H mini-heme proteins confirm that the pH-induced heme preference for coordination to His or to Cys uncovered in the GRW-L16CL30H mini-heme protein is irrespective of the presence of a competing binding site.

Spectral changes due to the redox cycling of the GRW-L16C and GRW-L30H mini-heme proteins were followed by UV-Vis absorption spectroscopy (Fig. S11). Experimental conditions were the same as described for the GRW-L16CL30H scaffold (see above, Section 8), except for the higher concentration of the samples (270 μM monomeric peptide and 90 μM heme final concentrations), thus we used a 2 mm pathlength Quartz cuvette with septum.

The electronic absorption spectra of the ferric and ferrous oxidation states of the GRW-L16C mini-heme protein at pH 10.8 were virtually identical to those of the GRW-L16CL30H mini-heme protein at pH 10.6 (Fig. S7). The subsequent re-oxidation of the ferrous sample also showed the contribution of the high-spin Cys-coordinated species, although to a lesser extent than in the case of

the GRW-L16CL30H mini-heme protein at pH 10.6, and reflected by the less pronounced shift of the Soret band, with maximum at 421 nm in this case (Fig. S11 top, gray trace).

The spectra of the GRW-L30H mini-heme protein in the ferric oxidation state at pH 6.8 (Fig. S11 bottom, green trace) is characteristic of a 5-coordinated heme, with the Soret band maximum at 395 nm and the LMCT band at 640nm. The differences in the Soret band of the GRW-L30H (maximum at 395 nm, Fig. S11) and the GRW-L16CL30H (maximum at 379 nm, Fig. S7) mini-heme proteins at neutral pHs, confirms our proposal of the blue shift in the latter originating from a longer length of the heme Fe^{III} to N_{His} bond. It may also reflect a small, yet significant for the coordination, difference in packing of the 4SCC scaffold with peptides having a single His30 mutation, as compared to the double mutant (His30 and Cys16) GRW-L16CL30H mini-heme protein. The ferrous GRW-L30H mini-heme protein at pH 6.8 showed a spectrum of a predominant oxyferrous heme (Fig. S11 bottom, dashed trace), possibly due to the presence of adventitious dissolved O_2 in solution. Upon re-oxidation, the resulting absorption spectrum (Fig. S11 bottom, gray trace) was almost identical to the initial ferric spectrum, confirming that the heme ligation to His30/imidazole is robust and conserved through the ferric/ferrous redox cycle.

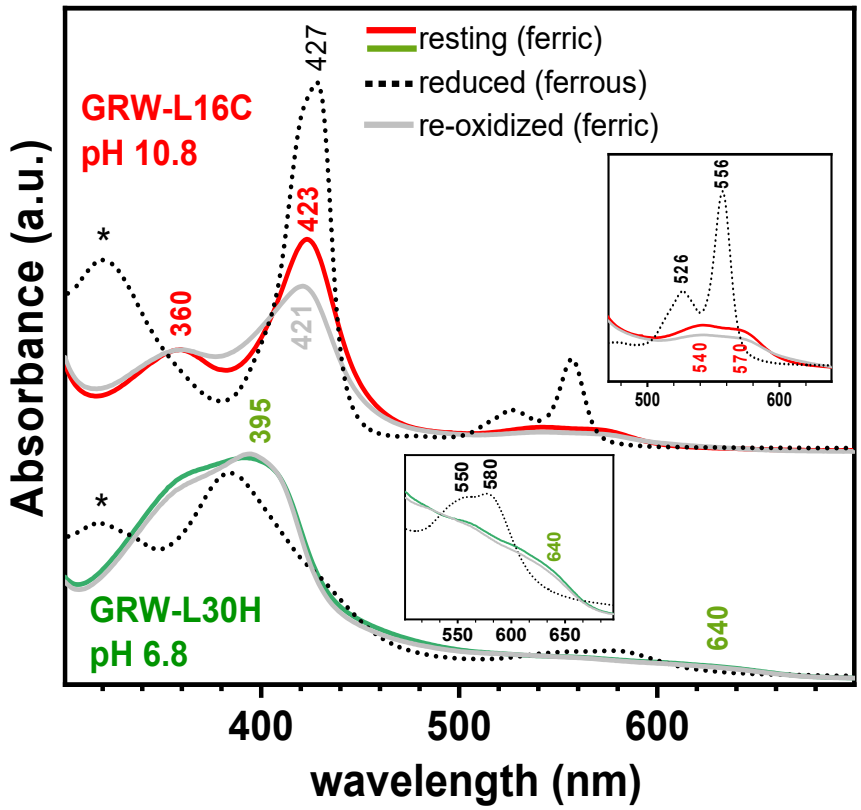


Figure S11: UV-Vis absorption spectra of the GRW-L16C and GRW-L30H mini-heme proteins in the ferric oxidation states at pH 10.8 (Cys-coordinated heme, red trace) and pH 6.8 (His-coordinated heme, green trace), respectively recorded at room temperature and in anaerobic conditions. Samples were reduced with dithionite (dotted traces) and subsequently re-oxidized with ammonium persulfate (gray traces). The asterisks show the contribution of the dithionite at 320nm.

11. Catalytic reactivity of the high-valent intermediate formed in the GRW-L16C and GRW-L30H mini-heme proteins.

The high-valent catalytic intermediate, known as Compound I, that is formed upon oxidation of the ferric heme active site in peroxidases [26], KatGs [10] and cyt P450 enzymes (Fig. 1A), results from the $2e^-$ oxidation reaction of the ferric heme with H_2O_2 (oxidant) to form the [Fe^{IV}=O $Por^{\bullet+}$] species. Such intermediate typically reacts with substrates by proton abstraction and $1e^-$ oxidation, as schematically shown in Fig. S1A. In KatGs [27], cyt P450s [5] and most peroxidases (reviewed in [26]), the [Fe^{IV}=O $Por^{\bullet+}$] species is a short-lived intermediate, yet it can be trapped within seconds in some cases when performing the reaction of the enzyme with hydrogen peroxide or peroxyacetic acid (PAA) at low temperatures (typically on ice) and then unequivocally identified by 9-GHz EPR spectroscopy.[28, 29]

We have applied this approach to the GRW-L30H and GRW-L16C mini-heme proteins (Fig. S11). As in the case of natural His-coordinated heme peroxidases, the His-coordinated mini-heme protein (at pH 7) forms a short-lived [Fe^{IV}=O $Por^{\bullet+}$] intermediate that could be trapped when using PAA as oxidant and mixing in ice for 2s. The very broad (ca. 2000 G) EPR spectrum of the flash-freeze sample after 2s mixing (Fig. S11A, light green trace) is characteristic of a $Por^{\bullet+}$ in magnetic interaction (exchange coupling) with the heme iron (Fe^{IV}=O), as has been shown in the case of turnip and lignin peroxidases [28, 29] for example. We could also show that when the reaction was performed in the presence of ABTS (2,2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) as substrate, the [Fe^{IV}=O $Por^{\bullet+}$] intermediate readily oxidizes the substrate resulting in the formation of the ABTS cation radical product, as shown by the characteristic narrow radical signal at $g \sim 2$ [30] (Fig. S11A, dark cyan trace) and the 'EPR-silent' (integer spin) Fe^{IV}=O intermediate. The choice of ABTS as substrate was based on the fact that the $1e^-$ oxidation of ABTS results in a very stable radical product, with a distinct EPR signal.

In the case of the Cys-coordinated mini-heme protein at pH 10.5, and as in the case of natural cyt P450s [5], the [Fe^{IV}=O $Por^{\bullet+}$] intermediate is much shorter lived (milliseconds time scale, even at low temperatures) hence we could not trap enough to detect the EPR signal. Yet, we could indirectly show that Compound I was formed since it reacted with ABTS substrate, resulting in the formation of the ABTS cation radical (Fig. S11B, dark cyan trace). Such an oxidation reaction was monitored by the disappearance of the ferric EPR signal of the thiolate heme (Fig. S11B, red trace) and concomitant formation of the ABTS cation radical (Fig. S11B, dark cyan trace), through the peroxide shunt pathway (Fig. S1A).

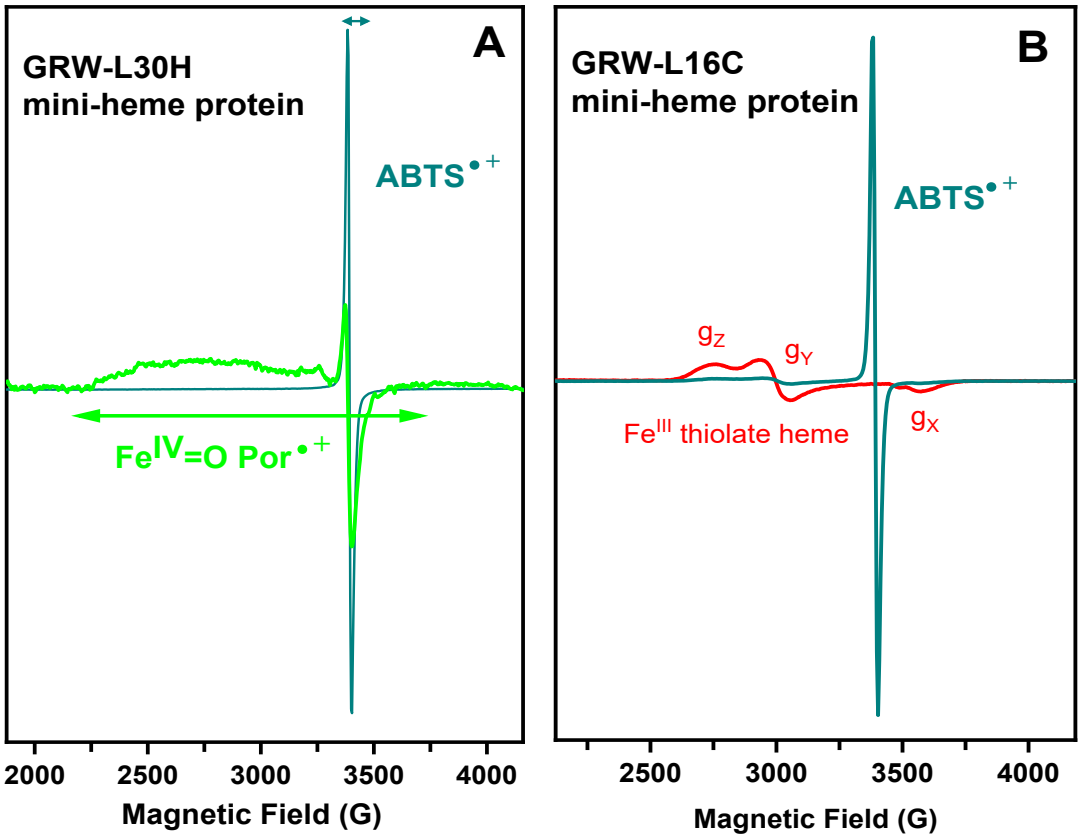


Figure S12: (Panel A) The 9-GHz EPR spectrum of the short-lived [Fe^{IV}=O $Por^{\bullet+}$] intermediate (light green trace) obtained upon reaction of the GRW-L30H ferric mini-heme protein at pH 7.0 with hydrogen peroxide as oxidant. The EPR spectrum of the $ABTS^{\bullet+}$ product (dark cyan trace) was detected when the reaction with peroxyacetic acid H_2O_2 was performed in the presence of ABTS as substrate. (Panel B) The reaction of the GRW-L16C ferric mini-heme protein at pH 10.5 (red trace) with hydrogen peroxide as oxidant and in the presence of ABTS as substrate, showing the formation of the $ABTS^{\bullet+}$ product (dark cyan trace). EPR experimental conditions: 9.48 GHz frequency, 100 kHz modulation frequency, 8 G modulation amplitude, 1.0 mW microwave power, T = 4K. A 5fold-molar excess of PAA and a 2 fold-molar excess of ABTS were used for the reactions with the mini-heme proteins (0.3 mM heme concentration, 3:1 peptide to protein ration).

10. References

[1] M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, *Chem. Rev.* **1996**, *96* (7), 2841-2887.

[2] I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chem. Rev.* **2005,** *105* (6), 2253-2277

[3] P. R. O. de Montellano, *Chem. Rev.* **2010**, *110* (2), 932-948.

[4] J. T. Groves, *Nat. Chem*. **2014**, *6* (2), 89-91.

[5] J. Rittle, M. T. Green, *Science* **2010**, *330* (6006), 933-937.

[6] (a) T. L. Poulos, B. C. Finzelx, A. J. Howard, *J. Mol. Biol.* **1987**, *195*, 687-700. (b) V. C. Murarka, D. Batabyal, J. A. Amaya, I. F. Sevrioukova, T. L. Poulos, *Biochemistry* **2020**, *59* (29), 2743-2750.

[7] C. S. Mocny, V. L. Pecoraro, *Acc. Chem. Res.* **2015**, *48* (8), 2388-2396.

[8] K J. Koebke, V. L. Pecoraro, *ACS Catal.* **2018**, *8* (9), 8046-8057.

[9] A. E. Tolbert, C. S. Ervin, L. Ruckthong, T. J. Paul, V. M. Jayasinghe-Arachchige, K. P. Neupane, J. A. Stuckey, R. Prabhakar, V. L. Pecoraro, *Nat. Chem.* **2020**, *12*, 405-411.

[10] A. Ivancich, L. J. Donald, J. Villanueva, B. Wiseman, I. Fita, P. C. Loewen, *Biochemistry* **2013**, 52, 7271-7282.

[11] B. Demeler, *Cur. Protoc. Prot. Sci.* **2010**, *60* (1), 7.13.1-7.13.24.

[12] B. Demeler, G. Gorbet in *Analytical Ultracentrifugation: Instrumentation, Software, and Applications*. (Eds: S. Uchiyama, W. F. Stafford, T. Laue), Springer, **2016**, pp. 119-143.

[13] E. Brookes, W. Cao, B. Demeler, *Eur Biophys J.* **2010**, *39* (3), 405-414.

[14] G. Gorbet, T. Devlin, B. Hernandez Uribe, A. K. Demeler, Z. Lindsey, S. Ganji, S. Breton, L. Weise-Cross, E. M. Lafer, E. H. Brookes, B. Demeler, *Biophys. J.* **2014**, *106* (8), 1741-1750.

[15] Y. Deng, J. Liu, Q. Zheng, D. Eliezer, N. R. Kallenbach, M. Lu, *Structure* **2006**, *14* (2), 247-255.

[16] The heme molecule was taken from the crystal structure of lignin peroxidase, PDB code: 1B82.

[17] T. L. Poulos, *Chem. Rev.* **2014**, *114* (7), 3919-3962.

[18] H. M. Girvan, J. M. Bradley, M. R. Cheesman, J. R. Kincaid, Y. Liu, K. Czarnecki, K. Fisher, D. Leys, S. E. J. Rigby, A. W. Munro, *Biochemistry* **2016**, *55*, 5073-5083.

[19] R. Perera, M. Sono, J. A. Sigman, T. D. Pfister, Y. Lu, J. H. Dawson, *Proc. Natl. Acad. Sci. USA* **2003**, *100* (7), 3641-3646.

[20] G. Smulevich, M. Paoli, G. De Sanctis, A. R. Mantini, F. Ascoli, M. Coletta, *Biochemistry* **1997**, *36,* 640-649.

[21] M. M. C. dos Santos, P. M. P. de Sousa, M. L. S. Goncalves, L. Krippahl, J. J. G. Moura, E. Lojou, P. Bianco, *J. Electroanal. Chem.* **2003**, *541*, 153-162.

[22] B. D. Fleming, Y. Tian, S. G. Bell, L.-L. Wong, V. Urlache, H. A. O. Hill, *Eur. J. Biochem.* **2003**, *270*, 4082-4088.

[23] E. Laviron, *J. Electroanal. Chem.***1979**, *101*, 19-28.

[24] A. Udit, M. G. Hill, H. B. Gray, *Langmuir* **2006,** *22,* 10854-10857.

[25] M. Machuqueiro, B. Victor, J. Switala, J. Villanueva, C. Rovira, I. Fita, P. C. Loewen, *Biochemistry* **2017**, *56* (17), 2271-2281.

[26] G. Smulevich, A. Feis, B. D. Howes, A. Ivancich, in *Handbook of Porphyrin Science* **2010** (Kadish, K. M., Smith, K. M., Guilard, R. Eds), vol. 6, Chapter 31, pp 367-453, World Scientific, Hackensack, NJ.

[27] C. Jakopitsch, M. Auer, A. Ivancich, F. Ruker, P. G. Furtmuller, C. Obinger, *J. Biol. Chem.* **2003**, *278*, 20185-20191.

[28] A. Ivancich, G. Mazza, A. Desbois, *Biochemistry* **2001**, *40*, 6860-6866.

[29] A. T. Smith, W. A. Doyle, P. Dorlet, A. Ivancich, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16084-16089.

[30] A. J. Fielding, R. Singh, B. Boscolo, P. C. Loewen, E. Ghibaudi, A. Ivancich, *Biochemistry* **2008**, *47*, 9781-9792