Delineating redox cooperativity in water-soluble and membrane multiheme cytochromes through protein design

Benjamin J. Hardy[1], Paulina Dubiel[1], Ethan L. Bungay[1], May Rudin[1], Christopher Williams[2], Christopher J. Arthur[2], Matthew J. Guberman-Pfeffer[3], A. Sofia Oliveira[2], Paul Curnow[1‡], J. L. Ross Anderson[1‡]

[1]School of Biochemistry, University of Bristol, United Kingdom
[2]School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom
[3] Department of Chemistry and Biochemistry, Baylor University, Texas 76706, United States
[‡]*Corresponding authors*

Abstract

Nature has evolved diverse electron transport proteins and multiprotein assemblies essential to the generation and transduction of biological energy. However, substantially modifying or adapting these proteins for user-defined applications or to gain fundamental mechanistic insight can be hindered by their inherent complexity. *De novo* protein design offers an attractive route to stripping away this confounding complexity, enabling us to probe the fundamental workings of these bioenergetic proteins and systems, while providing robust, modular platforms for constructing completely artificial electron-conducting circuitry. Here, we use a set of *de novo* designed, mono- and diheme soluble and membrane proteins to unpick the contributions of electrostatic micro-environments and dielectric constants of the surrounding protein medium on inter-cofactor redox cooperativity that we have previously reported. Experimentally, we find that the two heme sites in both the water-soluble and membrane constructs have broadly equivalent redox potentials in isolation, in agreement with continuum electrostatics calculations. BioDC, a Python program for the estimation of electron transfer energetics and kinetics within multiheme cytochromes, also predicts equivalent heme sites, and reports that burial within the low dielectric environment of the membrane strengthens heme-heme electrostatic coupling. We conclude that redox cooperativity in our diheme cytochromes is largely driven by heme electrostatic coupling and confirm that this effect is greatly strengthened by burial in the membrane. These results demonstrate that while our *de novo* proteins present minimalist, new-to-nature constructs, they enable the dissection and microscopic examination of processes fundamental to the function of vital, yet complex, bioenergetic assemblies.

Introduction

Multiheme cytochromes are an abundant group of water-soluble and transmembrane proteins that contain multiple redox-active heme cofactors, which are generally held in sufficiently close proximity to facilitate inter-cofactor electron transfer. They are essential to aerobic respiration and play many organism-dependent roles in anaerobic respiration and photosynthesis[1–4]. The biophysical properties and reactivity of these heme cofactors are finely controlled by their conformation[5,6], axial ligation, local protein environment[7,8], solvent accessibility and the polarity of the surrounding medium, parameters which have been manipulated and tailored through evolution[9]. As the heme cofactors are bound in close proximity, there is an interdependency between their redox properties[10], resulting from changes in heme-heme electrostatic interactions when an electron is added or removed from one of the pair. This phenomenon, known as redox cooperativity, is thought to promote efficient and rapid directional electron transfer, as observed in the transmembrane cytochrome bc_1[11] and b_6f complexes[12], terminal oxidases[13,14], bacterial photosynthetic reaction centres[15], water-soluble cytochromes *c*3[16] and the soluble and transmembrane domains of long extracellular heme wires[17].

Much effort has been made to unpick the specific factors that control the strength of electrostatic coupling and the resulting splitting of heme potentials in such proteins[12,18–21], though the ease of re-engineering these natural cytochromes is severely limited by their inherent complexity. Whilst there have been notable successes in altering the redox potentials of natural membrane cytochromes through heme ligand substitutions (including the bc_1 complex[22,23] and the decaheme MtrC[24]), most reported mutations confer only minimal effects and often disrupt heme binding sites to the point where the cofactor is lost entirely[25]. In parallel to these experimental efforts, computational studies have been instrumental in elucidating the contributions of electrostatics, electronic coupling and insulation from solvent[1,21,26,27], providing valuable electrochemical insight into otherwise spectroscopically equivalent hemes that can prove resistant to experimental interrogation[28].

An alternative, and experimentally tractable, way of studying the fundamental properties of hemes within proteins is through the use of simple 'maquette' proteins, which are typically water-soluble or amphipathic four-helix bundles designed to bind redox-active porphyrins[29–32]. Maquettes provide a minimal platform where the protein environment is defined and highly mutable, the sequence is devoid of evolutionary history or complexity, and hemes and non-natural porphyrins can be easily added or removed. When multiple hemes are bound within close proximity, redox cooperativity reminiscent of natural cytochromes is commonly observed[30,33,34]. Whilst the ease of binding porphyrins and other cofactors within maquettes has enabled the design of a multitude of highly functional proteins[35–41], these have typically lacked a well-defined, singular structure, hindering subsequent engineering with atomistic precision. To address this issue, we designed a suite of structured, robust and modular coiled-coiled *de novo* heme proteins which are readily expressed in *E. coli*[42].

This family of redox proteins is based upon the diheme protein called 4D2, which binds two bis-histidine coordinated *b*-type hemes, for which a crystal structure was solved to 1.9 Å resolution[33,42]. The two co-planar hemes sit in near-identical binding sites and exhibit noticeable redox cooperativity, resulting in two midpoint potentials split by 63 mV. From 4D2 we generated a rigid monoheme protein, m4D2, by removing one heme binding site and repacking the vacant heme binding site through computational protein design, with 2D NMR spectroscopy revealing a well-structured protein when loaded either with heme *b* or a symmetrical heme analogue. By duplicating the sequence of each helix, we also generated an extended tetraheme cytochrome, e4D2, which could be directly observed by electron microscopy (EM) despite its small size of 24 kDa. To study the properties of the diheme scaffold within the membrane, we used hydrophobic surface design[43] to create a transmembrane version of 4D2, termed CytbX, which was efficiently routed to bacterial membranes where it innately

bound two *b*-type hemes with high affinity. Interestingly, when bound within the transmembrane CytbX, the heme pair exhibits significantly stronger redox cooperativity than 4D2, with its two redox potentials split by 113 mV despite the interior heme binding sites being nearly identical to those in 4D2.

With the successful 4D2 and CytbX designs in hand, we are in the unique position to systematically study the interactions between a pair of *b*-type hemes within the same soluble protein scaffold, and directly compare them with the membrane equivalent. While it has not been possible to obtain a structure for CytbX, all experimental observations are consistent with 4D2 and CytbX having the same overall fold, inter-heme distance, relative heme orientations and internal heme-binding residues. These similarities rule out many confounding factors that can modulate the redox potential splitting of the heme pair. Our approach, therefore, provides a valuable reductionist system in which to directly study the effect of the membrane on heme redox potentials and redox cooperativity.

To this end, we report here the design, cellular production, and characterisation of a full set of single-heme 4D2 and CytbX constructs. We exploit the modular nature of our coiled-coil system to swap the heme-binding and core packing modules of m4D2, yielding the complementary soluble monoheme construct. We also apply the same computational design pipeline to CytbX to yield two transmembrane monoheme cytochromes. Experimentally, we observe that the local environments of each heme binding site contribute minimally to the redox split observed in the diheme constructs and, using two independent computational methods for redox potential prediction, we conclude that redox potential splitting in our system is mainly driven by heme-heme electrostatic interactions. Furthermore, redox potential calculations support the observation that the strengthened redox cooperativity in CytbX is indeed a function of the surrounding low dielectric medium and enable us to assign the low and high potential hemes of 4D2 and CytbX. Our findings demonstrate that both *de novo* soluble and membrane cytochromes can be engineered with atomic precision and demonstrate the utility of our platform for studying the fundamental properties and interactions of hemes within proteins.

Results

Protein Design

We previously reported the design of the monoheme protein m4D2 (Fig. 1a) from the diheme 4D2 (Fig. 1b) by maintaining heme binding at the site nearer the two structured TSN loops visible in the 4D2 and 4D2 T19D crystal structures (PDB IDs: 7AH0, 8CCR) (heme 1) whilst repacking the other heme site (heme 2) to create a compact hydrophobic core[42]. By maintaining heme 1, we sought to preserve the hydrogen bonding interactions observed in the 4D2 crystal structure between a heme propionate group and the asparagine residues of these loops. Here, we wished to design the complementary or inverse protein, with heme 2 preserved and heme site 1 re-packed to form a compact core. To distinguish between these constructs, we term the monoheme designs retaining heme in sites 1 and 2 as m1-4D2 (previously m4D2) and m2-4D2 respectively (Fig. 1c). To design the sequence of m2-4D2, we simply swapped the order of the heme-binding and hydrophobic packing portions of each helix of m1-4D2 (see Fig. 1a and c, and supplementary information) without any additional modifications.

Using Rosetta[44], we generated a theoretical model of m2-4D2 by threading the new sequence onto the crystal structure of 4D2 (containing a reconstructed loop), and subjected this model to structural relaxation (Cα RMSD vs 4D2 = 1 Å). In this model, the swapped modules form a heme-binding pocket and compact core as desired (Fig. 1c). ESMfold[45] predicts a high confidence structure (mean pLDDT = 87.0) that mostly matches the theoretical model (Cα RMSD = 2.63 Å, TM = 0.73; see Table S2, Fig. S5),

containing a heme-shaped cavity in heme site 2 with histidine side chains poised for heme coordination and a compact core containing the m1-4D2 packing module mutations in heme site 1 (Supplementary figure 1a, b). AlphaFold2 (AF2) predicts a high confidence structure (mean pLDDT = 82.3) containing a widened heme-shaped cavity in heme site 2 and compact core in heme site 1 (Supplementary figure 1c, d), though it predicts the helices of m2-4D2 to pack in an unexpected, mirrored topology (Supplementary figure 2).

Interestingly, AF2 also predicts this same mirrored helical topology for many of our *de novo* heme proteins, including m1-4D2 and CytbX, and most notably for 4D2, where two crystal structures are solved in the expected topology (Supplementary figure 3d). Frequently, for related proteins such as m1-4D2, AF2 predicts no heme-shaped cavity at all (Supplementary figure 4). In stark contrast, the ESMfold structure prediction of 4D2 aligns very closely to the crystal structure (Cα RMSD = 1.24 Å, TM = 0.92), with correct helical packing topology, heme-shaped cavities, and with heme-coordinating histidine-threonine H-bonding pairs superimposable (Supplementary figure 3a-c). We, therefore, suggest a general suitability of ESMfold over AF2 for the structure prediction of *de novo* heme proteins, in agreement with other published observations that ESMfold outperforms AF2 on amino acid sequences with no homology to natural proteins[46]. ESMfold structure predictions of the complete suite of proteins discussed in this report are shown in Figure S5, and prediction metrics from ESMfold and AF2 are compared in supplementary tables 1 and 2.

To generate transmembrane single-heme constructs, we first tested whether core residues of the m1-4D2 packing module could be directly transplanted into CytbX, given that it adopts the same overall backbone structure (within 1 Å) as 4D2. As an initial test, we copied the m1-4D2 core residues into the corresponding heme site in CytbX (site 2). Given that the positions of these m1-4D2 mutations were originally chosen as to not alter any knobs-into-holes (KIH) packing residues, they should not impair vital transmembrane helix-helix interactions in CytbX[42,43].

ESMfold predicts a high-confidence structure (mean pLDDT = 86.3) with minimal deviation to CytbX (Cα RMSD = 2.22 Å) and a heme-shaped cavity in the binding site. However, the m1-4D2 packing residues structurally perturb the bundle, slightly levering the helices apart (Supplementary figure 6a, b). When docked with heme and relaxed in Rosetta, the structure adopts a more compact conformation (1.64 Å Cα RMSD vs. CytbX, 2.05 Å Cα RMSD vs. m1-4D2). When the nine m1-4D2 mutations are introduced into the CytbX model through flexible-backbone Rosetta design, generated decoys have 0.7 Å mean Cα RMSD vs. CytbX and compact cores. Although this sequence could adopt suitable structures after Rosetta relaxation, the ESMfold prediction suggested that the sequence could be optimised further.

To generate improved single-heme CytbX variants, we used RosettaMP[47] with the *franklin2019*[48] membrane protein score function to computationally sample mutations for repacking either heme site. Here, we refer to CytbX with heme 1 present as m1-CytbX (Fig. 1d) and heme 2 present as m2-CytbX (Fig. 1f), these being the transmembrane counterparts of m1-4D2 and m2-4D2 respectively (Fig. 1a, c). Using a flexible-backbone design protocol, we allowed substitutions (to any hydrophobic amino acid) at residues in equivalent positions to those of the m1-4D2 packing module in either heme site of CytbX and generated 14,000 decoys per design. Decoys were ranked by their Rosetta score and the top 10 sequences for each design were further analysed. Multiple sequence alignment with ClustalOmega[49] revealed redundancy in the top sequences (Supplementary figure 7). For both constructs, we selected unique sequences that had either the lowest Rosetta score or the most compact core (as reported by the Packstat[50] metric) for bacterial expression. While all four designs (two designs for each of two constructs) were expressed and purified from *E. coli* membranes,

exhibiting identical biophysical properties, only the best expressing sequences from each pair (Supplementary figure 11a) for m1-CytbX and m2-CytbX are discussed further for clarity.

The Rosetta-generated models of m1-CytbX and m2-CytbX contained compact hydrophobic cores (Fig. 1d, f) and minimal structural deviations from the CytbX backbone (M1-CytbX Cα RMSD = 1.02 Å, m2-CytbX Cα RMSD = 0.96 Å). Both designs were predicted to contain four TM helices with N_{in}-C_{in} topology by the DeepTMHMM server (Supplementary figure 10)[51]. ESMfold predicted very high confidence structures closely matching the Rosetta designs (Supplementary figure 5) containing heme-shaped cavities and compact protein cores at the opposite ends of the bundles (Supplementary figure 8). Again, AF2 predicted lower-confidence, incorrectly packed structures lacking heme-binding cavities (Supplementary figure 9). Together, these findings suggest that the Rosetta-designed hydrophobic cores are more suitable for the CytbX scaffold than the unaltered, transplanted m1-4D2 packing module.

Expression and characterisation of mono-heme proteins

We subsequently ordered codon-optimised synthetic genes encoding the designed amino acid sequences for m2-4D2, m1-CytbX and m2-CytbX in standard pET expression vectors (pET-151, pET-29 or pET-21) with TEV (Tobacco etch virus N1A) protease-cleavable *N*-terminal hexahistidine or thrombin-cleavable Strep tags for the soluble and membrane proteins respectively. Following expression in *E. coli* T7-express cells, we purified the soluble m2-4D2 using affinity chromatography procedures as previously described[42], cleaved the purification tag with TEV protease, loaded the incompletely heme-incorporated m2-4D2 with exogenous heme, and purified the holoprotein by size-exclusion chromatography. (Supplementary figure 11). In contrast, we expressed the m1-CytbX and m2-CytbX membrane proteins in *E. coli* C43 (DE3) and used Strep-Tactin affinity chromatography to purify the proteins from isolated membranes, solubilised using the mild non-ionic detergent CYMAL-5. Following tag removal with Thrombin protease, we used SEC with no further addition of heme; both m1-CytbX and m2-CytbX eluted as single monodisperse species, whilst m2-4D2 eluted as two sharp peaks (Supplementary figure 11.)

All purified proteins exhibited an intense red colour with UV-visible absorbance spectra containing prominent ferric Soret peaks at 416-417 nm consistent with bis-histidine coordination of heme *b* [42,43,52] (Fig. 2 a-c). Unexpectedly, the absorbance spectrum of m2-CytbX contained an additional low intensity peak in the Q-band region at 590 nm characteristic of zinc protoporphyrin IX (ZnPPIX)[53] (Fig. 2c), and a broadened Soret peak at 417 nm. When excited at 430 nm, the fluorescence emission spectrum contained two prominent peaks at 593 nm and 648 nm confirming bound ZnPPIX[53] (Supplementary figure 12 ZnPPIX). Using calculated extinction coefficients of bound ferrous heme at 533 nm of 12,400 $M^{-1}cm^{-1}$ and bound ZnPPIX at 592 nm of 23,685 $M^{-1}cm^{-1}$, we determined that fewer than 10% of sites were loaded with ZnPPIX**.** The binding stoichiometry of all three monoheme designs was confirmed as 1:1 protein:heme by native mass spectrometry (MS) (Supplementary figure 13, Supplementary table 3). The small portion of ZnPPIX-bound m2-CytbX could not be identified in the MS spectrum.

We then used far-UV circular dichroism (CD) spectroscopy to confirm that the new holo-proteins were predominantly helical, with all exhibiting remarkable thermostability (Tm > 95 °C) (Fig. 2 d-f). The CD spectra exhibited a small, linear decrease in MRE at 222 nm up to 95 °C with no observable melt transition (Supplementary figure 14). The CD spectrum of m2-4D2 reveals a higher $\varepsilon_{222nm}/\varepsilon_{208nm}$ than both m1-CytbX and m2-CytbX (Fig. 2d), suggesting a more supercoiled structure[54]. Whilst crystal structures for these three proteins could not be obtained, the H^1-N^{15} TROSY spectrum of m2-4D2 shows similarly good peak dispersion to that of m1-4D2, indicating that it is well structured in the heme-bound state (Supplementary figure 15). This confirms that combining the heme-binding and

packing modules of m1-4D2 in either orientation, and in such a simple, modular fashion, yields well-structured holo-proteins.

Redox properties of water-soluble and transmembrane monoheme and diheme proteins

To measure the reduction potentials of the *b*-type hemes bound within these proteins, we employed optically transparent thin layer electrochemistry (OTTLE)[55]. Potentiometric titrations revealed a single redox potential for m2-4D2 at -125 ± 2 mV (vs NHE) (Fig. 3a), very close to the previously-measured potential for m1-4D2 at -117 mV[42], suggesting that the hemes of m1-4D2 and m2-4D2 reside within similar chemical environments despite being at opposite ends of the helical bundle. These potentials are closer to the high-potential heme of 4D2 (-104 mV) than the low-potential heme (-167 mV), and the difference between these potentials (8 mV) represents around 13% of the redox splitting measured for the 4D2 hemes (63 mV) (Supplementary table 4). This strongly suggests that the redox splitting in 4D2 is mainly a result of heme-heme interactions rather than differences in heme environments.

Redox titrations revealed single midpoint potentials for m1-CytbX and m2-CytbX of -66 ± 2 mV and -82 ± 1 mV respectively, with both residing between the two split potentials of diheme CytbX (-10 mV and -121 mV) (Fig. 3b). The small portion of ZnPPIX bound to m2-CytbX does not contribute to the potentiometric data as ZnPPIX is not redox active under these conditions. The difference in redox potentials of the single-heme membrane proteins (16 mV) is double that of the soluble counterparts (8 mV) and represents about 5% of the total redox difference of CytbX (111 mV). As for 4D2, these findings suggest that redox cooperativity, rather than differences in heme environment, dominate the overall redox potential splitting in CytbX. The increased magnitude of the cooperativity in CytbX versus 4D2 suggests a predominantly electrostatic origin for the split in heme *b* redox potentials, as the lower dielectric environment of the membrane protein interior would less efficiently screen the electrostatic field felt by each heme.

To further probe the electrostatic contribution to the redox cooperativity, we calculated electrostatic maps and isosurfaces for each of the soluble and membrane proteins. These revealed that m1-4D2 and m2-4D2 have different distributions of charge densities (Supplementary figure 16 Electrostatics), likely accounting for the small differences in their redox potentials. Of particular note is a smaller density of negative charge near the heme of m2-4D2 (Fig. S17 c), and a greater density of positive charge near the heme of m1-4D2 (Fig. S17 b), relative to 4D2. This highlights that the recombination of individual modules in these new, swapped constructs alters electrostatic properties, and that such changes should be assessed during design. The computed electrostatic isosurfaces also highlight the strong enrichment of positive charge on the cytoplasmic side of CytbX, a direct result of implementing the positive-inside rule to enforce N_{in}-C_{in} transmembrane topology during design (Fig. S17 d-f)[43,56,57], likely influencing the potential of heme 2.

Given these experimental results, we wished to further investigate the influence of electrostatics on the heme potentials of the full suite of 4D2 proteins using two computational methods: (i) an established Poisson-Boltzmann Monte-Carlo (PB-MC) continuum-electrostatic workflow[58–60] and (ii) BioDC, a new Python program developed to model redox cooperativity and conductivity in multi-heme cytochromes.[1]

As inputs to both computational methods, we used the crystal structure of 4D2 (PDB ID: 7AH0) with reconstructed loops and Rosetta-relaxed models of the five other proteins. The PB-MC calculations predicted a relative shift in E_m of -5 mV for m2-4D2 relative to m1-4D2 (Fig. 3c), in excellent agreement with the measured shift of -6 mV (Fig. 3a). For the membrane proteins, the PB-MC calculations predicted a relative shift of +2 mV for m2-CytbX vs m1-CytbX (Fig. 3d), whilst we experimentally

measured a shift of -16 mV (Fig. 3b). The positive shift predicted by PB-MC (although minimal) is consistent with the expected effect of enriching positive charge near heme 2 in m2-CytbX. Whilst the reason for the discrepancy observed between the predicted and measured values is not fully understood, it is likely to be related to the fact that protein dynamics is not explicitly factored into our PB-MC calculations. Adding conformational diversity of the protein (e.g. through the use of molecular dynamics-derived snapshots) to the current PB-MC method should improve the agreement with experimental values. This discrepancy may also suggest that m2-CytbX is adopting a subtly different structure than predicted. Overall, PB-MC predicts the heme site 1 and 2 environments in the water-soluble and membrane proteins to be mostly equivalent, in agreement with experimental measurements.

BioDC predicts the heme redox potential of m2-4D2 to be essentially the same as for m1-4D2, but the potential of m2-CytbX to be shifted by -37 mV relative to m1-CytbX. Both predictions are in good agreement with the experimental observations of -5 and -16 mV shifts (Figure 4).

In the parent diheme constructs, BioDC predicts that heme 2 of 4D2 in an aqueous environment (dielectric constant = 78.2) is shifted by -51 mV relative to heme 1, in good agreement with the 63 mV splitting in potential found experimentally. Heme 1 of CytbX in a membranous environment (dielectric constant = 6.4) is predicted to be shifted by -126 mV relative to heme 2. In good agreement with the 111-mV splitting found experimentally. Interestingly, the heme in site 2 of CytbX has a more positive potential than the heme in site 1, whereas the mono-heme variant that binds a heme in site 2 has a more negative potential than the variant that binds a heme in site 1. The flipping in redox ordering suggests that installing a packing module in place of an adjacent heme has changed the electrostatic environment of the remaining heme, and the remaining heme site is not representative of the equivalent site in the parent diheme protein.

The splitting of redox potentials in the diheme 4D2 and CytbX proteins reflects two contributions: (1) The different electrostatic interactions exerted by the environment on each heme while each is in the reduced state, and (2) the electrostatic destabilisation of the oxidised state of one heme by the oxidation of the adjacent heme. Approximately 16 (32%) and 46 (37%) mV of the redox splitting between the hemes of 4D2 and CytbX, respectively, originate from the environmental effect. The change in heme-heme interaction energy upon oxidation of one of the hemes contributes the remaining 35 (69%) and 80 mV (63%) mV to the redox potential splitting. All of these results from BioDC are insensitive to the choice of two different schemes for assigning atomic partial charges (RESP and CM5) to the heme in the reduced and oxidised states (Figure 4).

The enlarged potential splitting in CytbX versus 4D2 is attributable to the lower dielectric screening from the environment. If CytbX was hypothetically in the same aqueous environment as 4D2, the environment-induced and adjacent-heme-oxidation contributions to the redox potential splitting decrease from 46 to 33 mV, and 80 to 42 mV, respectively.

Discussion

In this work, we have demonstrated the utility of our robust and highly engineerable redox protein platform for the study of redox cooperativity in cytochromes. We demonstrate that single-heme cytochromes can be readily designed from our water-soluble and transmembrane diheme constructs, and that these retain the rigidity, tight heme-binding affinity, and biocompatibility of their parent proteins. In both 4D2 and CytbX, we confirm that redox cooperativity in the diheme 4D2 and CytbX is overwhelmingly (>60%) responsible for the splitting in redox potentials, whereas differences in the electrostatic micro-environments of the binding sites account for the remaining separation in

potentials. Importantly, we find that the stronger redox splitting observed in CytbX is a function of the surrounding low dielectric medium of the membrane environment. Whilst CytbX is highly simplistic, we believe the low dielectric membrane likely strengthens heme-heme electrostatic interactions in natural membrane cytochromes also. Furthermore, through the design of m2-4D2, we have demonstrated that the modularity of our system extends beyond simple duplication of the protein, but that heme-binding and packing modules can be interchanged in a facile manner.

Our experimental measurements and computational redox calculations suggest that individual heme binding sites at opposite ends of the 4D2/CytbX bundle fold are essentially equivalent. As expected, the similarity between sites is greater for the sequence-symmetric 4D2 than the sequence-asymmetric CytbX. It should be noted that the local environments of the hemes in the binding sites of m1-4D2/m2-4D2 and m1-CytbX/m2-CytbX are not entirely representative of the corresponding binding sites in the diheme construct. In the latter case, the proximity of the second heme strongly influences the local environment of the first heme. The sites within the monoheme constructs therefore represent the contribution of the protein only to the local environment of each heme within 4D2 or CytbX. It is also important to note that whilst the repacked cores of m1-4D2 and m2-4D2 contain identical sequences, the repacked cores of m1-CytbX and m2-CytbX were independently designed. Nevertheless, both designed cores contain similar hydrophobic amino acid content and are assumed to have minimal impact of the heme potentials of the CytbX Mono proteins.

Continuing from this work, it would be interesting to investigate the design of *de novo* cytochromes that bind heme pairs with dictated relative orientations, replicating the T-shaped, co-planar and stacked configurations observed in bacterial multiheme cytochromes[26]. Whilst this is non-trivial within the four-helix bundle scaffold, such designs could be developed with the use of ligand-aware generative protein models[61,62]. With full control over heme-heme distance and relative orientation, protein designers could realise more sophisticated electron transport pathways within artificial cytochromes.

In our recent report describing the design of 4D2[42], we commented that, whilst we observed split heme potentials, we could not assign which were the high and low potential hemes. Now, based upon the BioDC calculations, we suggest that heme 2 and heme 1 are the low and high potential hemes (b_L and b_H) in 4D2, whereas heme 1 and heme 2 are b_L and b_H in CytbX. It would be interesting to perform these calculations on artificial cytochromes containing greater numbers of coupled hemes such as the tetraheme e4D2, and this will be an avenue of future research. In earlier work describing the design of heme-binding maquettes, heme sites were also iteratively removed to interrogate the observed redox cooperativity[30]. Robertson *et. al.* observed that, when one heme was removed, the remaining heme exhibited a reduction potential matching the low potential of the coupled hemes in their diheme system. In contrast, we observe that our single-heme constructs have redox potentials in between the two split potentials of the diheme systems.

Assignment of the low and high potential hemes of CytbX has implications for the direction of electron transport across the membrane. When expressed in the cell, the N_{in}-C_{in} topology of CytbX (and monoheme variants) enforced during design means that electron transport would be most favourable from the periplasm to the cytoplasm. When assembling CytbX into proteoliposome systems, the orientation of the protein in the membrane should therefore be considered for most efficient transmembrane electron transport.

The construction of the CytbX-Mono proteins here now demonstrates that transmembrane cytochromes that bind only a single heme are accessible by design and provides two such proteins as blank slates for further construction and engineering. It is interesting to observe that the hyper-

thermostability of CytbX remains in the CytbX-Mono proteins, despite losing the contributions of two histidine-heme ligations in the removed heme site. This strongly suggests that the Rosetta-designed cores of these proteins are well packed and highly stable and suggests that further re-design in the non-heme binding portion of these proteins, for example to bind alternative redox-active cofactors, should be feasible without loss of stability. Immediate candidates to bind within these sites would be flavins, as observed in the STEAP family of proteins[63–65], and quinones.

Methods

Membrane protein design:

A Rosetta-generated relaxed model of CytbX was used as a starting point for the design of the CytbX-Mono proteins. To generate mono-heme models in each topology, either of the two hemes were deleted from the CytbX model in PyMol. Residues facing this cavity were then mutated to produce a tightly packed protein core. These residues were taken as the same mutable residues specified in the design of m1-4D2 from 4D2. Initially, the same m1-4D2 mutations were manually substituted into the equivalent positions in the CytbX sequence, given that CytbX and 4D2 share the same backbone and sequence register. These mutations were H9L, A13I, G48W, M51L, H68F, A71L, F103I, G106W and M109L. However, structure prediction with ESMfold revealed that backbone movements would likely be imposed by this combination of residues. When heme was docked into the ESMfold structure and the model subject to relaxation in Rosetta, this widening of helices was reduced, but the final model still deviated from the CytbX backbone.

To obtain optimised repacked cores for both monoheme CytbX structures, we turned to Rosetta design. Mutable residues as outlined above were allowed to mutate to any hydrophobic amino acid (FAMILYVW), defined in a resfile. Models of CytbX with either heme deleted were used as starting-structures, and histidine coordination of the remaining heme was enforced using a constraints file during all steps. Design was performed using a custom RosettaScripts XML file, implementing a FastDesign step followed by a relax step (for details see supplementary files). 14,000 decoys were produced per heme site. The spanfile defined all helical residues as membrane embedded, and starting structures were oriented in the membrane using the PPM server. Decoys were scored using the *franklin2019* score function with polar pore estimation disabled. The quality of repacked cores was assessed using the Packstat metric. The membrane topology of designed sequences was predicted using DeepTMHMM[51].

Soluble protein design:

The sequence of inverse m1-4D2 was created by introducing the 4D2-to-m1-4D2 mutations into the opposite ends of each helix of 4D2, effectively swapping the heme-binding and core-packing halves around. No further design was performed. The mutations from 4D2 to inverse-m1-4D2 were G20W, M23L, H37L, A41I, Y75I, G78W, M81L, H95F and A99L. A Rosetta model of m2-4D2 was produced by threading the mutated sequence onto the crystal structure of 4D2 (with reconstructed loops) using the SimpleThreadingMover followed by FastRelax implemented in RosettaScripts (see Supplementary). A predicted structure was also produced with ESMfold[45].

Protein structure prediction:

Sequences were subject to structure prediction with ESMfold via the 'ESMfold_advanced' Google Colab environment, with num_recycles set to 3 and the language models contacts option enabled: (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/ESMFold_advanced.ipynb). Structure prediction with AlphaFold2 was performed via the ColabFold environment

(https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=kOblAo-xetgx) in SingleSequence mode, with relaxation disabled and all other settings as default. Alignments of predicted structures and experimental or Rosetta models of proteins was performed using the TM-score web server (https://zhanggroup.org/TM-score/)[66].

Mass spectrometry:

The molecular weight of m1-CytbX and m2-CytbX holoproteins was analysed by native mass spectrometry as previously described for CytbX[43]. The molecular weight of heme-bound inverse-m1-4D2 was measured as described previously for 4D2 and m1-4D2[42].

Molecular biology:

Synthetic genes were ordered in expression vectors from Twist Bioscience. The m2-4D2 gene was purchased cloned between the EcoRI and NotI sites of a pET-21+ vector, with a N-terminal 6xHis tag followed by a tobacco-etch virus (TEV) cleavage site. Synthetic genes encoding 10xHis-tagged CytbX-Mono proteins, sfGFP fusions and Strep3-tagged m1-CytbX-0016 and m2-CytbX-0320 were ordered from Twist Biosciences cloned into the NdeI/XhoI sites of pET-29(b)+.

Protein expression and purification:

Expression and purification methods of membrane proteins (m1-CytbX, m2-CytbX) and soluble proteins (m2-4D2) was as described previously[42,43]. Membrane proteins were expressed in C43 (DE3) *E. coli* and soluble proteins were expressed in T7 express *E. coli* (NEB). Affinity purification of Strep3-tagged proteins was performed using a Strep-Tactin® Superflow® high capacity FPLC column (IBA Lifesciences).

Protein characterisation:

UV-visible (UV-Vis) absorbance spectra of purified proteins were measured from 250-750 nm in a quartz cuvette using a Cary UV-Vis spectrophotometer. Fluorescence spectra of proteins containing ZnPPIX were measured from 550-750 nm in a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) with an excitation wavelength of 430 nm. Optically-transparent thin-layer electrochemistry (OTTLE) measurements of membrane and soluble proteins was performed as described previously[42,43]. OTTLE experiments were performed at pH 7.4 for membrane proteins and at pH 8.6 for water soluble proteins. Midpoint potentials for proteins designed in this study are reported as the mean potential derived from single-electron Nernst fits to three separate experiments.

Continuum-Electrostatics calculations:

The redox potential shifts of the heme groups in 4D2, m1-4D2, m2-4D2, m1-CytbX and m2-CytbX were determined using a combination of Poisson–Boltzmann (PB) calculations and Metropolis Monte Carlo (MC) simulations as described in detail previously[67]. This method, which involves the simulation of the joint binding equilibrium of proton and electrons, uses PB calculations with the MEAD software[68] and MC calculations with the software PETIT[69]. The (individual and pairwise) terms needed for the free energies associated with protonation/reduction changes are computed using the PB equation. Such energies are then used in the MC calculations.

The atomic charges and radii for all the atoms in the protein and heme group were taken from the GROMOS 54A7 force field[70] and from our previous work[42], respectively. The simulations used a temperature of 298 K and a molecular surface defined with a solvent probe radius of 1.4 Å. Dielectric constants of 80 and 20 were used for the solvent and protein[67], respectively. The membrane was modelled as a low-dielectric slab parallel to the x-y plane using a dielectric constant value of 20. Each MC simulation comprises 10^5 MC steps, and the acceptance/rejection of each step followed a Metropolis criterion using the previously determined PB free energies.

Predicted reduction potentials of hemes were obtained by fitting the simulated titration curves for the hemes of m1-4D2, m2-4D2, m1-CytbX and m2-CytbX to a Nernst equation describing a single electron reduction event. This approach was previously shown to perform well in predicting redox potential shifts for m1-4D2 and its mutants[42,71].

BioDC Methods:

The structures of 4D2, m1-4D2, m2-4D2, CytbX, m1-CytbX, and m2-CytbX were provided in PDB format to the Structure Preparation & Relaxation module of BioDC. This module interactively assists the user in preparing Atomic Model Building with Energy Refinement (AMBER) topology and coordinate files, The user can (1) select to mutate none, one, or multiple residues; (2) designate Asp, Glu, His, Tyr, and Lys residues as titratable for constant pH molecular dynamics simulations; (3) select pairs of Cys residues that should have disulphide linkages; (4) assign the reduced or oxidised state to each heme individually; and (5) decide to immerse or not the protein in an aqueous rectangular or octahedral water box with a specified thickness for the water layer, and specified numbers of Na^+ and Cl^- ions. Importantly, the module automatically detects whether each heme is of the *b*- or *c*-type variety and whether each heme has His-His or His-Met axial ligands to correctly assign the appropriate force field parameters for the user-selected redox state. Only these types of hemes are currently supported by BioDC.

The parameters for His-His ligated *b*-type hemes were in part taken from Yang *et al.*[72], and developed with the Metal Center Parameter Builder (MCPB) program[73] of the AmberTools suite[74] based on quantum chemical calculations using the B3LYP approximate density functional and a mixed basis set (LANL2TZ(f) for Fe and 6-31G(d) for all second-row elements). Gaussian 16 Rev. A.03[75] was used for the quantum chemical calculations.

The prepared structures were passed to the Energetic Estimation module of BioDC. This module interfaces with the Poisson-Boltzmann Surface Area (PBSA) program of the AmberTools suite to calculate (1) the oxidation energy of each heme while all other hemes are in the reduced state, and (2) the change in the oxidation energy of each heme due to the oxidation of another heme, while all other hemes (if there are more than two hemes in total) are in the reduced state. These quantities reflect, respectively, the influence of the electrostatic environment and heme-heme interactions on the heme redox potentials.

The interior protein static dielectric constant for the PBSA calculations was estimated from the solvent accessible surface area of the hemes[76]. Static dielectric constants of 6.776 and 6.445 were assigned to the protein interiors of 4D2 and CytbX, respectively. These values were used for the associated mono-heme variants to facilitate comparability of mono- and di-heme constructs.

4D2, m1-4D2, and m2-4D2 were modelled in an aqueous environment with a medium dielectric of 78.2. CytbX, m1-CytbX, and m2-CytbX were modelled in a membranous environment with a dielectric equal to that of the protein interior (6.445), as done previously[76]. The thickness of the implicit membrane was 30.6 (CytbX), 28.6 (m1-CytbX), and 30.4 (m2-CytbX) Å, as predicted with the PPM 3.0 server provided by the Orientation of Proteins in Membranes[77]. Beyond the membranous region of ~30 Å centred on each CytbX-related construct, the medium dielectric was set to 78.2 for an aqueous environment.

The Energetic Estimation module of BioDC can also compute the heme-to-heme electron transfer reorganisation energy, assign a generic electronic coupling value based on the mutual orientation of the cofactors, and calculate the non-adiabatic Marcus theory electron transfer rate. None of these

features were needed for the present study. Likewise, BioDC has a Redox Current Prediction module that was not needed for the present study. This module uses the estimated electron transfer rates to compute the associated redox current in the diffusive and protein-limited steady-state regimes. Results using all the features of BioDC are described in a manuscript under preparation.

NMR spectroscopy

Isotopic labelling of m2-4D2 was performed as previously described[78]. Briefly, the cultures were grown in LB at 37 °C, and upon reaching the OD600 of 0.8 the cells were pelleted and resuspended in cell wash solution (22 mM KH_2PO_4, 48 mM Na_2HPO_4, 8.6 mM NaCl). The cultures were then pelleted again, and the cells from 3 L of bacterial cultures were gently resuspended in 750 mL of M9 minimal media (22 mM KH_2PO_4, 48 mM Na_2HPO_4, 8.6 mM NaCl, 18 mM ^{15}N-NH_4Cl, 4 g/L glucose, 1 x Basal Medium Eagle Vitamins (100x stock; VWR/Lonza #733-1801), 2 mM $MgSO_4$, 0.1 mM $CaCl_2$). These were incubated at 37 °C for 1 h to aid cell recovery and subsequently induced with 1 mM IPTG. The proteins were expressed for 4 h at 37°C and purified as normal. SEC was performed at pH 6.4 (50 mM potassium phosphate, 20 mM potassium chloride).

The concentration of the NMR samples was 350 uM in 50 mM potassium phosphate, 20 mM potassium chloride (pH 6.4) with 10% D_2O. ^1H-^{15}N HSQC-TROSY spectra were acquired using standard Bruker pulse programs at 25°C on a 700 MHz Bruker AVANCE *HD III* NMR spectrometer equipped with a 1.7 mm triple-resonance micro-cryoprobe. All the NMR data were processed using Topspin 3.6 (Bruker, Coventry, UK) and the spectra were visualized and using CcpNmr Analysis version 2.4.2 [79] (Wim et al, 2005).

Acknowledgements

This work was supported at the University of Bristol by the Biological and Biotechnological Sciences Research Council (BBW003449/1 and BB/X009831/1). We wish to thank the University of Bristol for funding a studentship for E.L.B. M J. G-P. wishes to acknowledge the financial support of an Office of Vice Provost Research fellowship at Baylor University and the mentorship of Bryan Shaw.

References

1. Guberman-Pfeffer MJ (2023) From Hot Water to Dry Dirt: Microbes Use Cytochrome 'Nanowires' of Similar Conductivity but Different Structure. bioRxiv [Internet]:2023.06.12.544705. Available from: https://www.biorxiv.org/content/10.1101/2023.06.12.544705v1

2. Baquero DP, Cvirkaite-Krupovic V, Hu SS, Fields JL, Liu X, Rensing C, Egelman EH, Krupovic M, Wang F (2023) Extracellular cytochrome nanowires appear to be ubiquitous in prokaryotes. Cell [Internet] 186:2853-2864.e8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0092867423005317

3. Kim HJ, Khalimonchuk O, Smith PM, Winge DR (2012) Structure, function, and assembly of heme centers in mitochondrial respiratory complexes. Biochim. Biophys. Acta - Mol. Cell Res. [Internet] 1823:1604–1616. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167488912000997

4. Swainsbury DJK, Qian P, Hitchcock A, Hunter CN (2023) The structure and assembly of reaction centre-light-harvesting 1 complexes in photosynthetic bacteria. Biosci. Rep. [Internet] 43. Available from: /bioscirep/article/43/5/BSR20220089/232981/The-structure-and-assembly-of-reaction-centre

5. Senge MO, MacGowan SA, O'Brien JM (2015) Conformational control of cofactors in nature – the influence of protein-induced macrocycle distortion on the biological function of tetrapyrroles. Chem.

Commun. [Internet] 51:17031–17063. Available from: https://pubs.rsc.org/en/content/articlehtml/2015/cc/c5cc06254c

6. Jentzen W, Ma JG, Shelnutt JA (1998) Conservation of the conformation of the porphyrin macrocycle in hemoproteins. Biophys. J. [Internet] 74:753–763. Available from: https://pubmed.ncbi.nlm.nih.gov/9533688/

7. Mao J, Hauser K, Gunner MR (2003) How cytochromes with different folds control heme redox potentials. Biochemistry [Internet] 42:9829–9840. Available from: https://pubs.acs.org/doi/full/10.1021/bi027288k

8. Zheng Z, Gunner MR (2009) Analysis of the electrochemistry of hemes with Ems spanning 800 mV. Proteins Struct. Funct. Bioinforma. 75:719–734.

9. Hosseinzadeh P, Lu Y (2016) Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics. Biochim. Biophys. Acta - Bioenerg. [Internet] 1857:557–581. Available from: https://www.sciencedirect.com/science/article/pii/S000527281500170X

10. Fonseca BM, Paquete CM, Salgueiro CA, Louro RO (2012) The role of intramolecular interactions in the functional control of multiheme cytochromes c. FEBS Lett. [Internet] 586:504–509. Available from: https://onlinelibrary.wiley.com/doi/full/10.1016/j.febslet.2011.08.019

11. Bhaduri S, Stadnytskyi V, Zakharov SD, Hasan SS, Bujnowicz, Sarewicz M, Savikhin S, Osyczka A, Cramer WA (2017) Pathways of Transmembrane Electron Transfer in Cytochrome bc Complexes: Dielectric Heterogeneity and Interheme Coulombic Interactions. J. Phys. Chem. B 121:975–983.

12. Szwalec M, Bujnowicz Ł, Sarewicz M, Osyczka A (2022) Unexpected Heme Redox Potential Values Implicate an Uphill Step in Cytochrome b6f. J. Phys. Chem. B [Internet] 126:9771–9780. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.2c05729

13. Nicholls P, Petersen LC (1974) Haem-haem interactions in cytochrome aa3 during the anaerobic-aerobic transition. BBA - Bioenerg. 357:462–467.

14. Wikström MKF, Harmon HJ, Ingledew WJ, Chance B (1976) A re-evaluation of the spectral, potentiometric and energy-linked properties of cytochrome c oxidase in mitochondria. FEBS Lett. 65:259–277.

15. Pottosin II, Chamorovsky CS, Chamorovsky SK (2007) Cooperative interaction of high-potential hemes in the cytochrome subunit of the photosynthetic reaction center of bacterium Ectothiorhodospira shaposhnikovii. Biochem. [Internet] 72:1254–1260. Available from: https://link.springer.com/article/10.1134/S0006297907110120

16. GAYDA J -P, BENOSMAN H, BERTRAND P, MORE C, ASSO M (1988) EPR determination of interaction redox potentials in a multiheme cytochrome: cytochrome c3 from Desulfovibrio desulfuricans Norway. Eur. J. Biochem. [Internet] 177:199–206. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1432-1033.1988.tb14362.x-i2

17. Guberman-Pfeffer MJ (2022) Assessing Thermal Response of Redox Conduction for Anti-Arrhenius Kinetics in a Microbial Cytochrome Nanowire. J. Phys. Chem. B [Internet] 126:10083–10097. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.2c06822

18. Palmer G, Esposti MD (1994) Application of Exciton Coupling Theory to the Structure of Mitochondrial Cytochrome b. Biochemistry [Internet] 33:176–185. Available from: https://pubs.acs.org/doi/abs/10.1021/bi00167a023

19. Bhaduri S, Stadnytskyi V, Zakharov SD, Hasan SS, Bujnowicz, Sarewicz M, Savikhin S, Osyczka A, Cramer WA (2017) Pathways of Transmembrane Electron Transfer in Cytochrome bc Complexes:

Dielectric Heterogeneity and Interheme Coulombic Interactions. J. Phys. Chem. B [Internet] 121:975–983. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.6b11709

20. Hasan SS, Zakharov SD, Chauvet A, Stadnytskyi V, Savikhin S, Cramer WA (2014) A map of dielectric heterogeneity in a membrane protein: The hetero-oligomeric cytochrome b6 f complex. J. Phys. Chem. B [Internet] 118:6614–6625. Available from: https://pubs.acs.org/doi/full/10.1021/jp501165k

21. Guberman-Pfeffer MJ (2023) Structural Determinants of Redox Conduction Favor Robustness over Tunability in Microbial Cytochrome Nanowires. J. Phys. Chem. B [Internet] 127:7148–7161. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.3c02912

22. Pintscher S, Kuleta P, Cieluch E, Borek A, Sarewicz M, Osyczka A (2016) Tuning of Hemes b Equilibrium Redox Potential Is Not Required for Cross-Membrane Electron Transfer. J. Biol. Chem. [Internet] 291:6872–6881. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925820429436

23. Świerczek M, Cieluch E, Sarewicz M, Borek A, Moser CC, Dutton PL, Osyczka A (2010) An electronic bus bar lies in the core of cytochrome bc1. Science (80-.). [Internet] 329:451–454. Available from: https://www.science.org/doi/10.1126/science.1190899

24. van Wonderen JH, Adamczyk K, Wu X, Jiang X, Piper SEH, Hall CR, Edwards MJ, Clarke TA, Zhang H, Jeuken LJC, et al. (2021) Nanosecond heme-to-heme electron transfer rates in a multiheme cytochrome nanowire reported by a spectrally unique His/Met-ligated heme. Proc. Natl. Acad. Sci. U. S. A. [Internet] 118:e2107939118. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2107939118

25. Osyczka A, Moser CC, Daldal F, Dutton PL (2004) Reversible redox energy coupling in electron transfer chains. Nat. 2004 4276975 [Internet] 427:607–612. Available from: https://www.nature.com/articles/nature02242

26. Breuer M, Rosso KM, Blumberger J (2014) Electron flow in multiheme bacterial cytochromes is a balancing act between heme electronic interaction and redox potentials. Proc. Natl. Acad. Sci. U. S. A. [Internet] 111:611–616. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1316156111

27. Teixeira VH, Soares CM, Baptista AM (2002) Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic detail. J. Biol. Inorg. Chem. [Internet] 7:200–216. Available from: https://link.springer.com/article/10.1007/s007750100287

28. Breuer M, Zarzycki P, Blumberger J, Rosso KM (2012) Thermodynamics of electron flow in the bacterial deca-heme cytochrome MtrF. J. Am. Chem. Soc. [Internet] 134:9868–9871. Available from: https://pubs.acs.org/doi/abs/10.1021/ja3027696

29. Farid TA, Kodali G, Solomon LA, Lichtenstein BR, Sheehan MM, Fry BA, Bialas C, Ennist NM, Siedlecki JA, Zhao Z, et al. (2013) Elementary tetrahelical protein design for diverse oxidoreductase functions. Nat. Chem. Biol. [Internet] 9:826–833. Available from: www.nature.com/naturechemicalbiology

30. Robertson DE, Farid RS, Moser CC, Urbauer JL, Mulholland SE, Pidikiti R, Lear JD, Wand AJ, DeGrado WF, Dutton PL (1994) Design and synthesis of multi-haem proteins. Nat. 1994 3686470 [Internet] 368:425–432. Available from: https://www.nature.com/articles/368425a0

31. North B, Summa CM, Ghirlanda G, DeGrado WF (2001) Dn-symmetrical tertiary templates for the design of tubular proteins. J. Mol. Biol. [Internet] 311:1081–1090. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022283601949009

32. Discher BM, Noy D, Strzalka J, Ye S, Moser CC, Lear JD, Blasie JK, Dutton PL (2005) Design of Amphiphilic Protein Maquettes: Controlling Assembly, Membrane Insertion, and Cofactor Interactions. Biochemistry [Internet] 44:12329–12343. Available from: https://pubs.acs.org/doi/10.1021/bi050695m

33. Ghirlanda G, Osyczka A, Liu W, Antolovich M, Smith KM, Dutton PL, Wand AJ, DeGrado WF (2004) De novo design of a D2-symmetrical protein that reproduces the diheme four-helix bundle in cytochrome bc1. J. Am. Chem. Soc. 126:8141–8147.

34. Goparaju G, Fry BA, Chobot SE, Wiedman G, Moser CC, Leslie Dutton P, Discher BM (2016) First principles design of a core bioenergetic transmembrane electron-transfer protein. Biochim. Biophys. Acta - Bioenerg. [Internet] 1857:503–512. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0005272815002479

35. Fry BA, Solomon LA, Leslie Dutton P, Moser CC (2016) Design and engineering of a man-made diffusive electron-transport protein. Biochim. Biophys. Acta - Bioenerg. 1857:513–521.

36. Watkins DW, Jenkins JMX, Grayson KJ, Wood N, Steventon JW, Le Vay KK, Goodwin MI, Mullen AS, Bailey HJ, Crump MP, et al. (2017) Construction and in vivo assembly of a catalytically proficient and hyperthermostable de novo enzyme. Nat. Commun. [Internet] 8:358. Available from: http://www.nature.com/articles/s41467-017-00541-4

37. Stenner R, Anderson JLR (2020) Chemoselective N-H insertion catalyzed by a de novo carbene transferase. Biotechnol. Appl. Biochem. [Internet] 67:527–535. Available from: https://onlinelibrary.wiley.com/doi/10.1002/bab.1924

38. Richard S, W. SJ, Annela S, Ross AJL (2018) A de novo peroxidase is also a promiscuous yet stereoselective carbene transferase. bioRxiv [Internet]:328484. Available from: https://www.biorxiv.org/content/10.1101/328484v2

39. Ennist NM, Zhao Z, Stayrook SE, Discher BM, Dutton PL, Moser CC (2022) De novo protein design of photochemical reaction centers. Nat. Commun. 2022 131 [Internet] 13:1–10. Available from: https://www.nature.com/articles/s41467-022-32710-5

40. Anderson JLR, Armstrong CT, Kodali G, Lichtenstein BR, Watkins DW, Mancini JA, Boyle AL, Farid TA, Crump MP, Moser CC, et al. (2014) Constructing a man-made c-type cytochrome maquette in vivo: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. Chem. Sci. [Internet] 5:507–514. Available from: http://xlink.rsc.org/?DOI=C3SC52019F

41. Lichtenstein BR, Farid TA, Kodali G, Solomon LA, Anderson JLR, Sheehan MM, Ennist NM, Fry BA, Chobot SE, Bialas C, et al. (2012) Engineering oxidoreductases: maquette proteins designed from scratch. Biochem. Soc. Trans. [Internet] 40:561–566. Available from: https://portlandpress.com/biochemsoctrans/article/40/3/561/64332/Engineering-oxidoreductases-maquette-proteins

42. Hutchins GH, Noble CEM, Bunzel HA, Williams C, Dubiel P, Yadav SKN, Molinaro PM, Barringer R, Blackburn H, Hardy BJ, et al. (2023) An expandable, modular de novo protein platform for precision redox engineering. Proc. Natl. Acad. Sci. [Internet] 120:e2306046120. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2306046120

43. Hardy BJ, Martin Hermosilla A, Chinthapalli DK, Robinson C V, Anderson JLR, Curnow P (2023) Cellular production of a de novo membrane cytochrome. Proc. Natl. Acad. Sci. [Internet] 120:2022.12.06.519282. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2300137120

44. Leman JK, Weitzner BD, Lewis SM, Adolf-Bryfogle J, Alam N, Alford RF, Aprahamian M, Baker D, Barlow KA, Barth P, et al. (2020) Macromolecular modeling and design in Rosetta: recent methods

and frameworks. Nat. Methods 2020 177 [Internet] 17:665–680. Available from: https://www.nature.com/articles/s41592-020-0848-2

45. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, et al. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. Science (80-.). [Internet] 379:1123–1130. Available from: https://www.science.org/doi/10.1126/science.ade2574

46. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK (2023) Before and after AlphaFold2: An overview of protein structure prediction. Front. Bioinforma. [Internet] 3. Available from: /pmc/articles/PMC10011655/

47. Koehler Leman J, Mueller BK, Gray JJ (2016) Expanding the toolkit for membrane protein modeling in Rosetta. Bioinformatics [Internet] 33:btw716. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btw716

48. Alford RF, Fleming PJ, Fleming KG, Gray JJ (2020) Protein Structure Prediction and Design in a Biologically Realistic Implicit Membrane. Biophys. J. [Internet] 118:2042–2055. Available from: https://linkinghub.elsevier.com/retrieve/pii/S000634952030237X

49. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. [Internet] 50:W276–W279. Available from: https://europepmc.org/articles/PMC9252731

50. Sheffler W, Baker D (2009) RosettaHoles: Rapid assessment of protein core packing for structure prediction, refinement, design, and validation. Protein Sci. [Internet] 18:229–239. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pro.8

51. Hallgren J, Tsirigos KD, Damgaard Pedersen M, Juan J, Armenteros A, Marcatili P, Nielsen H, Krogh A, Winther O (2022) DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv [Internet]:2022.04.08.487609. Available from: https://www.biorxiv.org/content/10.1101/2022.04.08.487609v1

52. Lundgren CAK, Sjöstrand D, Biner O, Bennett M, Rudling A, Johansson A-L, Brzezinski P, Carlsson J, von Ballmoos C, Högbom M (2018) Scavenging of superoxide by a membrane-bound superoxide oxidase. Nat. Chem. Biol. [Internet] 14:788–793. Available from: http://www.nature.com/articles/s41589-018-0072-x

53. Tangar A, Derrien V, Lei R, Santiago Estevez MJ, Sebban P, Bernad S, Miksovska J (2019) Utility of fluorescent heme analogue ZnPPIX to monitor conformational heterogeneity in vertebrate hexa-coordinated globins. Metallomics [Internet] 11:906–913. Available from: https://academic.oup.com/metallomics/article/11/5/906/5962186

54. Lombardi A, Pirro F, Maglio O, Chino M, DeGrado WF (2019) De novo design of four-helix bundle metalloproteins: One scaffold, diverse reactivities. Acc. Chem. Res. [Internet] 52:1148–1159. Available from: https://pubs.acs.org/doi/abs/10.1021/acs.accounts.8b00674

55. Leslie Dutton P (1978) Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems. Methods Enzymol. 54:411–435.

56. Baker JA, Wong W-C, Eisenhaber B, Warwicker J, Eisenhaber F (2017) Charged residues next to transmembrane regions revisited: “Positive-inside rule” is complemented by the “negative inside depletion/outside enrichment rule.” BMC Biol. [Internet] 15:66. Available from: http://bmcbiol.biomedcentral.com/articles/10.1186/s12915-017-0404-4

57. HEIJNE G, GAVEL Y (1988) Topogenic signals in integral membrane proteins. Eur. J. Biochem. [Internet] 174:671–678. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1988.tb14150.x

58. Teixeira VH, Soares CM, Baptista AM (2002) Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic detail. JBIC J. Biol. Inorg. Chem. [Internet] 7:200–216. Available from: http://link.springer.com/10.1007/s007750100287

59. Gunner MR, Baker NA Continuum Electrostatics Approaches to Calculating pKas and Ems in Proteins. In: Methods in Enzymology. Vol. 578. NIH Public Access; 2016. pp. 1–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27497160

60. Zheng Z, Gunner MR (2009) Analysis of the electrochemistry of hemes with Ems spanning 800 mV. Proteins Struct. Funct. Bioinforma. [Internet] 75:719–734. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/prot.22282

61. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, Ahern W, Borst AJ, Ragotte RJ, Milles LF, et al. (2023) De novo design of protein structure and function with RFdiffusion. Nat. 2023 6207976 [Internet] 620:1089–1100. Available from: https://www.nature.com/articles/s41586-023-06415-8

62. Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, Lee GR, Morey-Burrows FS, Anishchenko I, Humphreys IR, et al. (2023) Generalized Biomolecular Modeling and Design with RoseTTAFold All-Atom. bioRxiv [Internet]:2023.10.09.561603. Available from: https://www.biorxiv.org/content/10.1101/2023.10.09.561603v1

63. Kleven MD, Dlakić M, Lawrence CM (2015) Characterization of a Single b-type Heme, FAD, and Metal Binding Sites in the Transmembrane Domain of Six-transmembrane Epithelial Antigen of the Prostate (STEAP) Family Proteins. J. Biol. Chem. [Internet] 290:22558–22569. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925820448136

64. Oosterheert W, Reis J, Gros P, Mattevi A (2020) An elegant four-helical fold in NOX and STEAP enzymes facilitates electron transport across biomembranes - Similar vehicle, different destination. Acc. Chem. Res. [Internet] 53:1969–1980. Available from: https://dx.doi.org/10.1021/acs.accounts.0c00400

65. Oosterheert W, Gros P (2020) Cryo-electron microscopy structure and potential enzymatic function of human six-transmembrane epithelial antigen of the prostate 1 (STEAP1). J. Biol. Chem. [Internet] 295:9502–9512. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925817489691

66. Zhang Y, Skolnick J (2004) Scoring function for automated assessment of protein structure template quality. Proteins Struct. Funct. Bioinforma. [Internet] 57:702–710. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/prot.20264

67. Teixeira VH, Cunha CA, Machuqueiro M, Oliveira ASF, Victor BL, Soares CM, Baptista AM (2005) On the use of different dielectric constants for computing individual and pairwise terms in poisson-boltzmann studies of protein ionization equilibrium. J. Phys. Chem. B [Internet] 109:14691–14706. Available from: https://pubs.acs.org/doi/full/10.1021/jp052259f

68. Bashford D, Karplus M (1990) pKa's of Ionizable Groups in Proteins: Atomic Detail from a Continuum Electrostatic Model. Biochemistry 29:10219–10225.

69. Baptista AM, Soares CM (2001) Some Theoretical and Computational Aspects of the Inclusion of Proton Isomerism in the Protonation Equilibrium of Proteins. Available from: https://pubs.acs.org/sharingguidelines

70. Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur. Biophys. J. [Internet] 40:843–856. Available from: http://link.springer.com/10.1007/s00249-011-0700-9

71. Oliveira ASF, Rubio J, Noble CEM, Anderson JLR, Anders J, Mulholland AJ (2023) Fluctuation Relations to Calculate Protein Redox Potentials from Molecular Dynamics Simulations. J. Chem. Theory Comput. [Internet]. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jctc.3c00785

72. Yang L, Skjevik ÅA, Han Du W-G, Noodleman L, Walker RC, Götz AW (2016) Data for molecular dynamics simulations of B-type cytochrome c oxidase with the Amber force field. Data Br. [Internet] 8:1209–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27547799

73. Li P, Merz KM (2016) MCPB.py: A Python Based Metal Center Parameter Builder. J. Chem. Inf. Model. [Internet] 56:599–604. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jcim.5b00674

74. Case DA, Aktulga HM, Belfon K, Cerutti DS, Cisneros GA, Cruzeiro VWD, Forouzesh N, Giese TJ, Götz AW, Gohlke H, et al. (2023) AmberTools. J. Chem. Inf. Model. [Internet] 63:6183–6191. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jcim.3c01153

75. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. and DJF (2016) Guassian 16, Revision A.03.

76. Jiang X, Van Wonderen JH, Butt JN, Edwards MJ, Clarke TA, Blumberger J (2020) Which Multi-Heme Protein Complex Transfers Electrons More Efficiently? Comparing MtrCAB from Shewanella with OmcS from Geobacter. J. Phys. Chem. Lett. [Internet] 11:9421–9425. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpclett.0c02842

77. Lomize AL, Todd SC, Pogozheva ID (2022) Spatial arrangement of proteins in planar and curved membranes by PPM 3.0. Protein Sci. [Internet] 31:209–220. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pro.4219

78. Marley J, Lu M, Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. J. Biomol. NMR [Internet] 20:71–75. Available from: https://pubmed.ncbi.nlm.nih.gov/11430757/

79. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct. Funct. Genet. [Internet] 59:687–696. Available from: https://pubmed.ncbi.nlm.nih.gov/15815974/

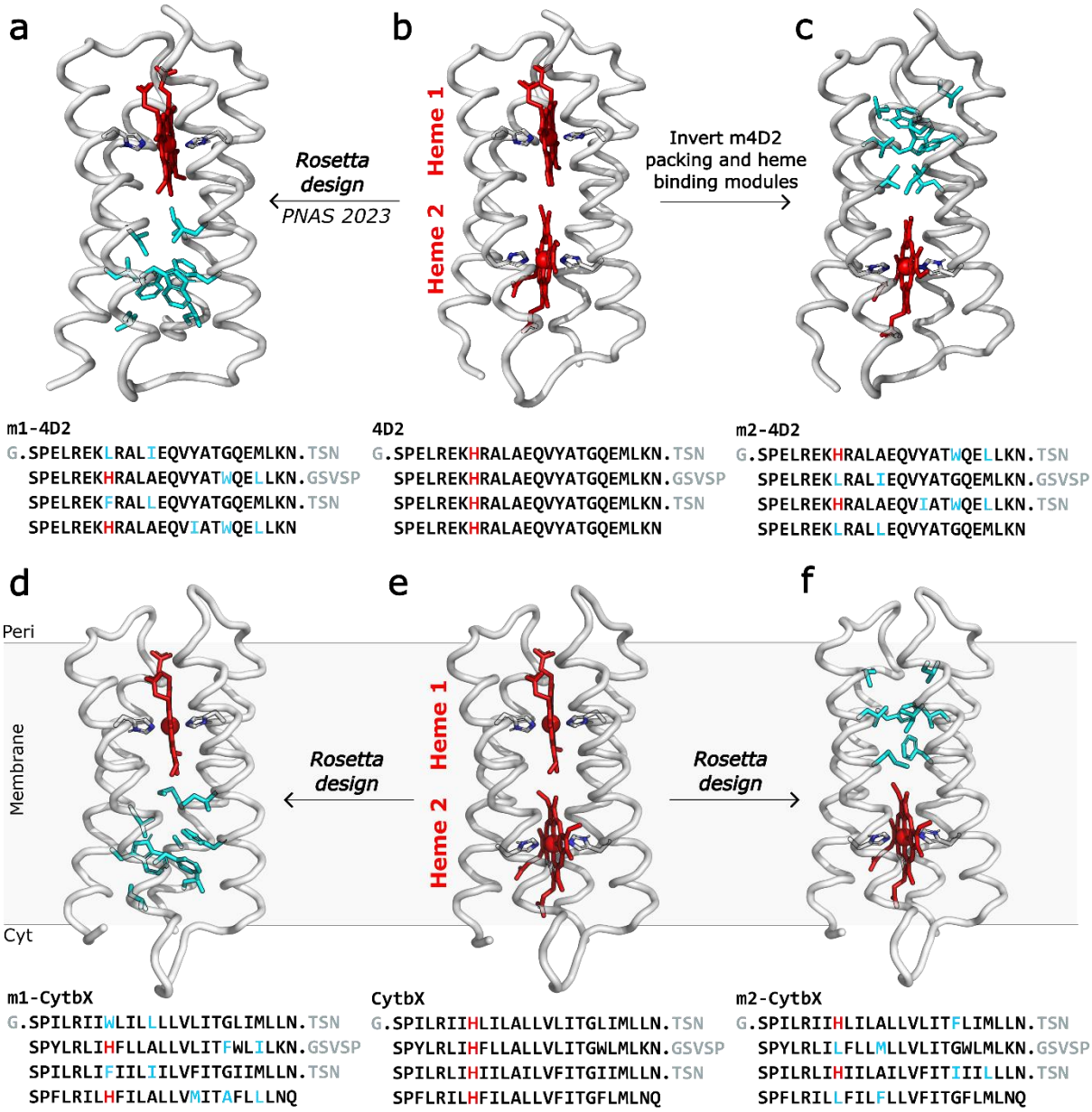

Figure 1. The suite of monoheme and diheme water-soluble and transmembrane modular redox proteins based on the coiled-coil diheme protein 4D2. (**a**) m1-4D2 contains heme 1 but heme site 2 is re-designed. (**b**) 4D2 is the basis of the protein suite and contains two identical bis-histidine heme binding sites. (**c**) m2-m4D2 contains heme 2 but heme site 1 is re-designed, containing the same core packing residues as in site 2 if m1-4D2. (**d**) m1-CytbX contains heme 1 of CytbX but heme site 2 was re-designed with Rosetta. (**e**) CytbX is the transmembrane version of 4D2. (**f**) m2-CytbX contains heme in site 2 but heme site 1 was re-designed with Rosetta. All structures are Rosetta-generated models, apart from 4D2 which is the crystal structure (PDB ID: 7AH0) with reconstructed loops. Protein backbones are shown as grey ribbons, hemes (red), mutated core residues (cyan) and heme-coordinating residues (grey) are shown as grey sticks. The sequence of each protein is shown, with mutated residues highlighted in cyan, histidines highlighted in red, and loop residues in grey.

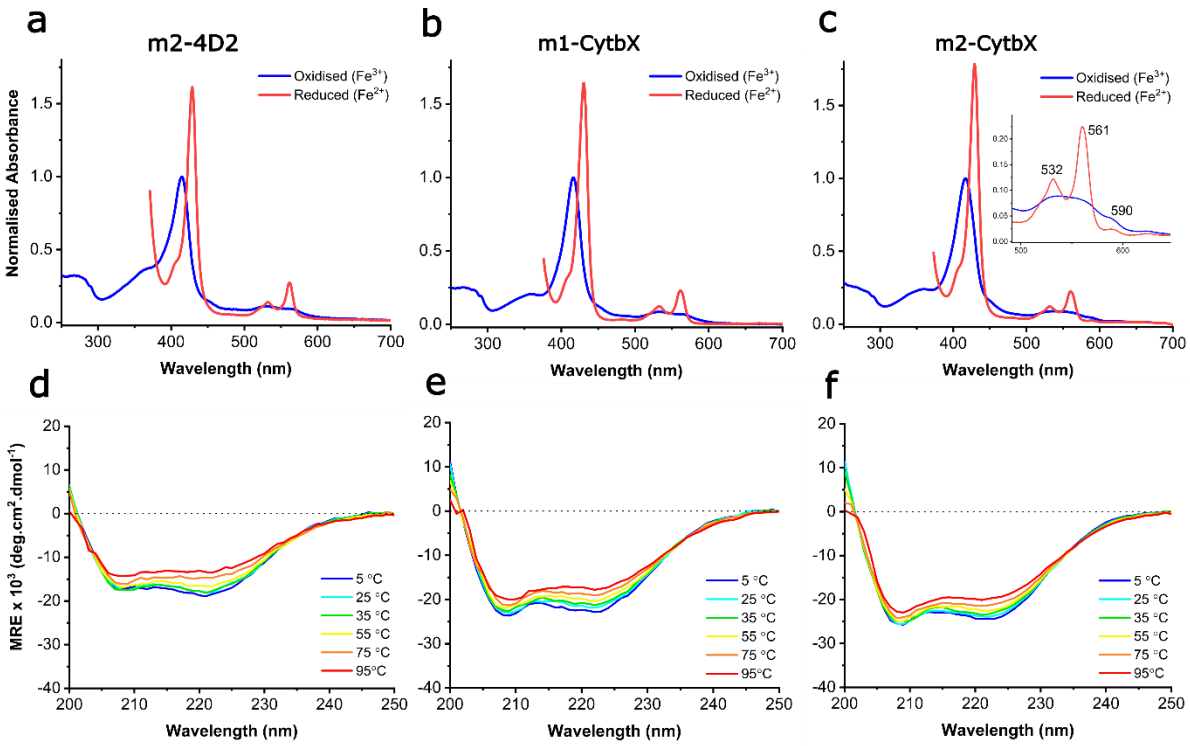

Figure 2. Biophysical characterisation of monoheme designs. UV-Vis absorbance spectra of purified (**a**) m2-4D2, (**b**) m1-CytbX and (**c**) m2-CytbX. Reduced spectra are cropped at the point where the detector is saturated by dithionite absorbance. Absorbance is normalised to the oxidised Soret peak maximum. Circular dichroism spectra during thermal melts of purified (**d**) m2-4D2, (**e**) m1-CytbX and (**f**) m2-CytbX from 5 °C to 95 °C. All designs bind heme, are helical and highly thermostable as designed.

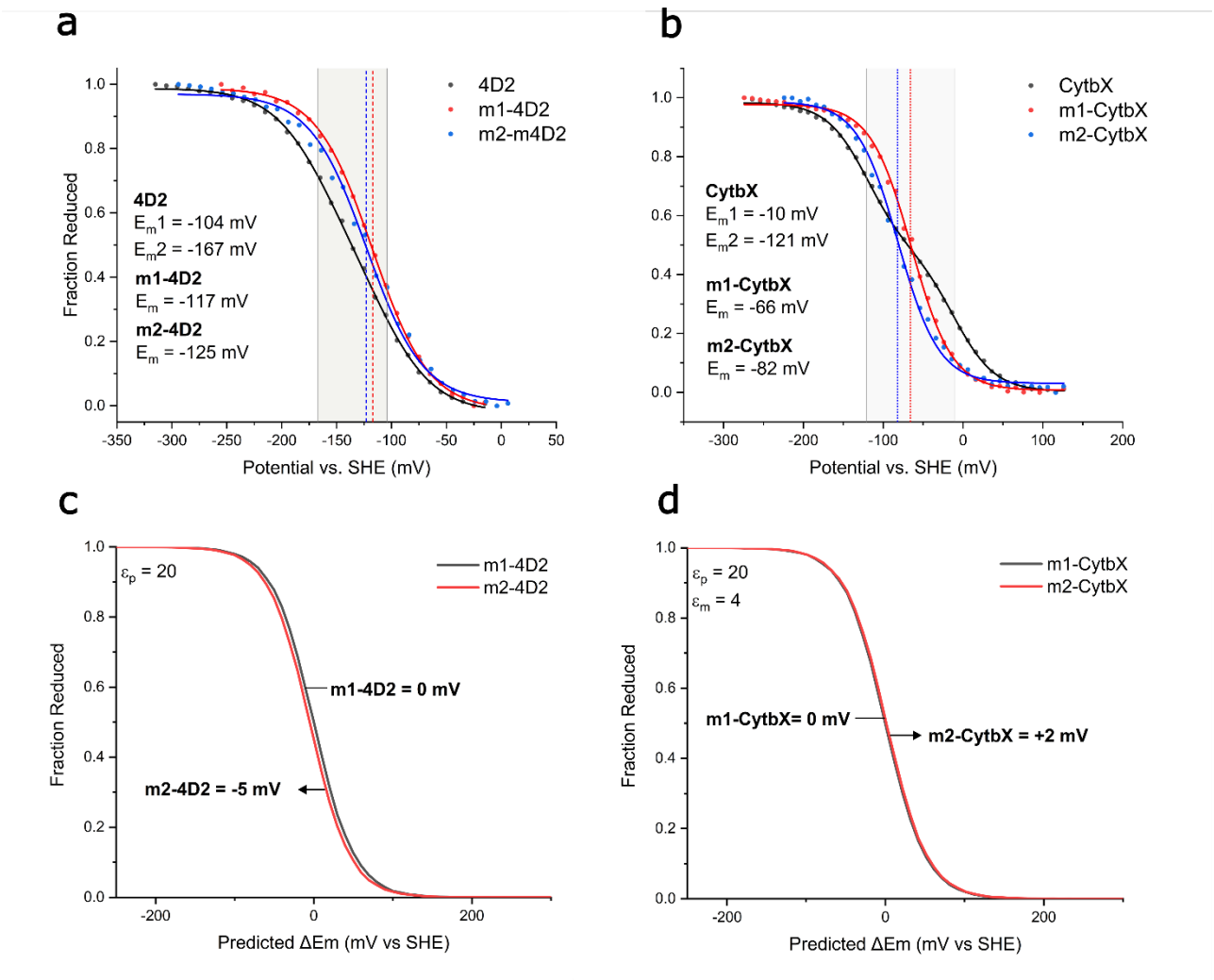

Figure 3. Redox potentiometry of the monoheme and diheme (a) membrane proteins and (b) water soluble proteins. Measurements were performed with purification tags removed from all proteins. Membrane proteins: pH = 7.4, soluble proteins: pH = 8.6. Buffer for membrane proteins: 50 mM NaP (pH 7.4), 150 mM NaCl, 5% glycerol, 0.08% CYMAL-5. Poisson-Boltzmann Monte-Carlo electrostatics calculations of redox titrations for (c) m1-4D2 vs. m2-4D2 and (d) m1-CytbX vs. m2-CytbX. Water-soluble proteins were modelled in a solvent with a dielectric constant (ε_p) of 20 at pH 8.6, and membrane proteins were modelled in an implicit membrane with ε_m=4 at pH 7.4. Predicted E_m shifts are reported using m1-4D2 and m1-CytbX as references for m2-4D2 and m2-CytbX respectively. Redox potentials were derived by fitting to a Nernst equation describing a single electron transfer.

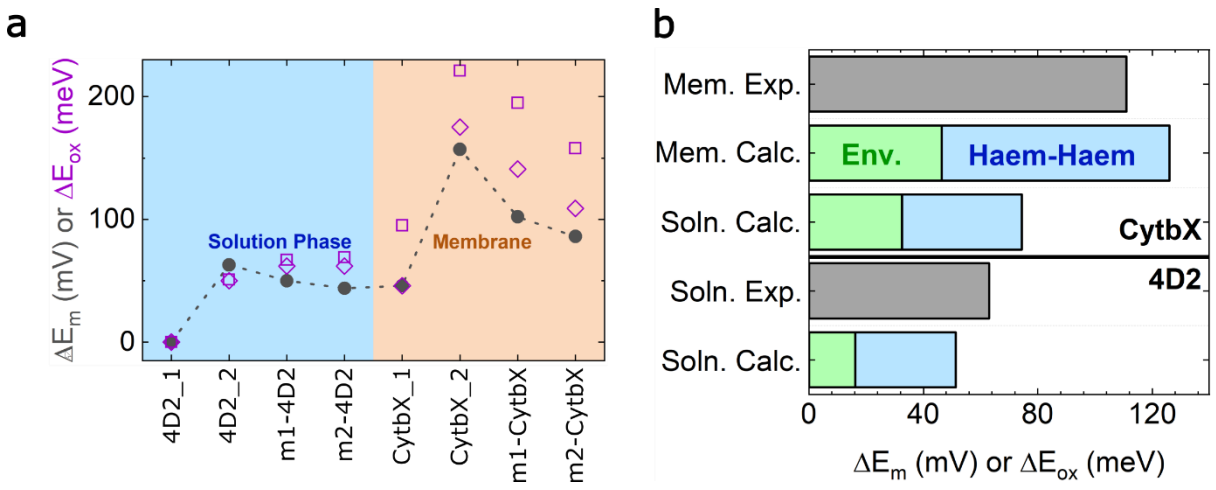

Figure 4. (*a*) BioDC reproduces the trend in redox potential shifts across the mono- and diheme constructs in aqueous and membranous environments, and (*b*) delineates the contributions to the redox potential splitting in the dihemes in terms of environmental electrostatics and heme-heme interactions. Hemes in site 1 and 2 of the diheme constructs are denoted by suffixes (e.g. 4D2_1 denotes 4D2 heme 1). Note that the relative redox potentials for the dihemes from experiment were assigned to heme site 1 and 2 based on the BioDC results. The purple diamonds and squares correspond to results obtained using atomic partial charges for the heme derived according to the Restrained Electro-Static Potential (RESP) and Charge Method 5 (CM5) schemes, respectively.