

Cite this: *Chem. Sci.*, 2019, **10**, 3031

All publication charges for this article have been paid for by the Royal Society of Chemistry

## One fold, two functions: cytochrome P460 and cytochrome $c'$ - $\beta$ from the methanotroph *Methylococcus capsulatus* (Bath)†

Hannah R. Adams, <sup>a</sup> Callie Krewson, <sup>b</sup> Jenny E. Vardanega, <sup>b</sup> Sotaro Fujii, <sup>c</sup> Tadeo Moreno, <sup>a</sup> Chicano, <sup>a</sup> Yoshihiro Sambongi, <sup>c</sup> Dimitri Svistunenko, Jordi Paps, <sup>a</sup> Colin R. Andrew <sup>\*b</sup> and Michael A. Hough <sup>\*a</sup>

Nature is adept at utilising highly similar protein folds to carry out very different functions, yet the mechanisms by which this functional divergence occurs remain poorly characterised. In certain methanotrophic bacteria, two homologous pentacoordinate  $c$ -type heme proteins have been identified: a cytochrome P460 (cyt P460) and a cytochrome  $c'$ - $\beta$  (cyt cp- $\beta$ ). Cytochromes P460 are able to convert hydroxylamine to nitrous oxide ( $N_2O$ ), a potent greenhouse gas. This reactivity is similar to that of hydroxylamine oxidoreductase (HAO), which is a key enzyme in nitrifying and methanotrophic bacteria. Cyt P460 and HAO both have unusual protein-heme cross-links, formed by a Tyr residue in HAO and a Lys in cyt P460. In contrast, cyts cp- $\beta$  (the only known cytochromes  $c'$  with a  $\beta$ -sheet fold) lack this crosslink and appears to be optimized for binding non-polar molecules (including NO and CO) without enzymatic conversion. Our bioinformatics analysis supports the proposal that cyt cp- $\beta$  may have evolved from cyt P460 via a gene duplication event. Using high-resolution X-ray crystallography, UV-visible absorption, electron paramagnetic resonance (EPR) and resonance Raman spectroscopy, we have characterized the overall protein folding and active site structures of cyt cp- $\beta$  and cyt P460 from the obligate methanotroph, *Methylococcus capsulatus* (Bath). These proteins display a similar  $\beta$ -sheet protein fold, together with a pattern of changes to the heme pocket regions and localised tertiary structure that have converted a hydroxylamine oxidizing enzyme into a gas-binding protein. Structural comparisons provide insights relevant to enzyme redesign for synthetic enzymology and engineering of gas sensor proteins. We also show the widespread occurrence of cyts cp- $\beta$  and characterise their phylogeny.

Received 22nd November 2018  
Accepted 20th January 2019

DOI: 10.1039/c8sc05210g  
[rsc.li/chemical-science](http://rsc.li/chemical-science)

## Introduction

Understanding the means by which enzymes can utilise a common fold and structural motifs to evolve very different functions is essential for synthetic biology and rational enzyme design for biotechnology. A useful guide in nature can be found in proteins within the same organism that diverge to adopt different functional roles, typically following a gene duplication event. Methanotrophic bacteria, whether or not they are also nitrifiers, are of considerable interest for biotechnological applications. Methanotrophs are a subgroup of methylotrophic

bacteria, which utilise methane as their sole source of carbon and energy<sup>1</sup> and possess the enzyme methane monooxygenase (MMO) which catalyses the oxidation of methane to methanol.<sup>2</sup> Some methanotrophs, such as *Methylococcus capsulatus* (Bath), are also an ammonia-oxidising nonlithotrophic bacteria (ANB), organisms that can aerobically oxidise ammonia to nitrite but (unlike nitrifiers) do not use this as their source of energy.<sup>3</sup> Nitrous oxide is produced by ANB from nitrite reduction (part of nitrifier denitrification) and also hydroxylamine oxidation. In *M. capsulatus* (Bath), two related  $\beta$ -sheet proteins (31% sequence identity<sup>4</sup>) have been previously identified, a cytochrome P460 (McP460) and a cytochrome  $c'$  (McCP- $\beta$ )<sup>5,6</sup> encoded by the *CytL* and *CytS* genes respectively. Cytochromes P460, defined by an unusual porphyrin-Lys cross link at the 13' meso carbon, catalyse the oxidation of hydroxylamine ( $NH_2OH$ ) to form  $N_2O$  under anaerobic conditions. The apparent production of nitrite during this reaction under aerobic conditions has more recently been shown to be due to oxidation of an NO intermediate by oxygen in a non-enzymatic manner.<sup>7</sup> The hydroxylamine substrate is a toxic intermediate generated as

<sup>a</sup>School of Biological Sciences, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK. E-mail: [mahough@essex.ac.uk](mailto:mahough@essex.ac.uk)

<sup>b</sup>Department of Chemistry and Biochemistry, Eastern Oregon University, La Grande, Oregon 97850, USA. E-mail: [candrew@eou.edu](mailto:candrew@eou.edu)

<sup>c</sup>Graduate School of Biosphere Science, Hiroshima University, Kagamiyama 1-4-4, Higashi-Hiroshima, Hiroshima, 739-8528, Japan

† Electronic supplementary information (ESI) available. See DOI: [10.1039/c8sc05210g](https://doi.org/10.1039/c8sc05210g)



a result of ammonia monooxygenase or methane monooxygenase activity in ammonia oxidising and methane oxidising bacteria (AOB or MOB) respectively.<sup>8</sup> Hydroxylamine oxidation also occurs in the unrelated hydroxylamine oxidoreductases (HAO) which, in addition to containing multiple heme centers, have a different type of P460 porphyrin cross-link involving a conserved Tyr, as well as very different protein folds.<sup>9–11</sup> The HAO P460 heme pocket also contains Tyr, Asp and His residues implicated in reactivity.<sup>9–11</sup> Structures of cyt P460<sup>12</sup> and HAO<sup>9,10</sup> from the ammonia oxidising bacterium *Nitrosomonas europea*, as well as cyt P460 from *Nitrosomonas* sp. AL212,<sup>13</sup> have previously been determined, but no structural data have been available for their methanotrophic equivalents. *Nitrosomonas europea* cyt P460 (NeP460) and *Nitrosomonas* sp. AL212 (NsALP460) have been the focus of several recent spectroscopic studies.<sup>12–15</sup> A relationship between heme P460 character, the protein crosslink, and distortion of the heme geometry away from planarity has recently been proposed.<sup>13</sup>

Cytochromes *c'* (cyts cp) are defined by a pentacoordinate heme Fe with a CXXCH c-heme-binding motif located close to the C-terminus.<sup>16,17</sup> Whereas all cyt cp crystal structures to date have revealed 4  $\alpha$ -helix bundles, circular dichroism studies of McCP- $\beta$  predicted a predominantly  $\beta$ -sheet fold,<sup>5</sup> suggesting that there are two distinct types of cytochrome *c'* proteins, designated cyt cp- $\alpha$  or cyt cp- $\beta$  according to their secondary structure. Within the cyt cp- $\alpha$  family, reactivity at the distal heme site is governed by a buried and crowded hydrophobic environment that selects for small non-polar ligands (NO and CO) that do not require H-bond stabilization (reviewed in ref. 16)). Moreover, distal steric constraints, coupled with a solvent-exposed proximal heme pocket, result in an unusual proximally-bound five-coordinate heme nitrosyl complex (5cNO or 5c{FeNO})<sup>7</sup> formed *via* distal six-coordinate heme nitrosyl (6cNO or 6c{FeNO})<sup>7</sup> and a transient dinitrosyl species.<sup>18–20</sup> Functional roles for cyt cp- $\alpha$  in defence against nitrosoative stress and NO transport have been proposed, while the mechanism of 6cNO  $\rightarrow$  5cNO conversion is similar to that of the mammalian NO-sensor, soluble guanylate cyclase,<sup>21</sup> for which no structural data of the haem domain are currently available.

By contrast with cyts cp- $\alpha$ , much less is known about the cyt cp- $\beta$  family, the sole characterized example to date being that of *M. capsulatus* (Bath) (McCP- $\beta$ ). Data including the amino acid sequence and circular dichroism spectra have been reported for McCP- $\beta$ ,<sup>4,5</sup> together with some ligand binding and spectropotentiometric studies.<sup>6</sup> These studies suggested that McCP- $\beta$  is a homodimeric heme protein with a somewhat higher molecular weight (16 kDa per monomer) than cyts cp- $\alpha$  ( $\sim$ 12 kDa per monomer), and with a predicted  $\beta$ -sheet structure. Although McCP- $\beta$  exhibits some of the ligand binding and spectroscopic properties of cyts cp- $\alpha$ , its reduction potential ( $-205$  mV) is much lower than that of the cyts cp- $\alpha$  ( $-10$  to  $+202$  mV).<sup>22</sup> A transcriptomics study of *M. capsulatus* (Bath) revealed that exposure to ammonia greatly increased expression (28.5–40-fold) of *CytS* (McCP- $\beta$ ) and of HAO, but not *CytL* (McP460).<sup>23</sup> This strongly suggests a role for McCP- $\beta$  in NO scavenging and protection against nitrosoative stress, for example dealing with NO produced by HAO or nitrite reductases in nitrifier

denitrification. It has also been suggested that HAO and cyt cp- $\beta$  are required for N<sub>2</sub>O formation in *M. capsulatus* (Bath)<sup>24</sup> although upregulation of *CytS* in fact may reflect a protective function against the production of NO by HAO. The lack of *CytL* upregulation in the presence of ammonia may suggest that HAO plays the major role in hydroxylamine oxidation under the conditions used. Interestingly, McCP- $\beta$  has been shown to act as an electron acceptor from cyt P460 that has been reduced by hydroxylamine, but this is unlikely to be a primary function since cyt cp- $\beta$  and cyt P460 are not always present in a single organism.<sup>7</sup> Notably, in NeP460, mutation of the cross-linking Lys to Arg, Ala or Tyr resulted in catalytically inactive proteins with spectroscopic properties similar to those of the cyts cp.<sup>25</sup>

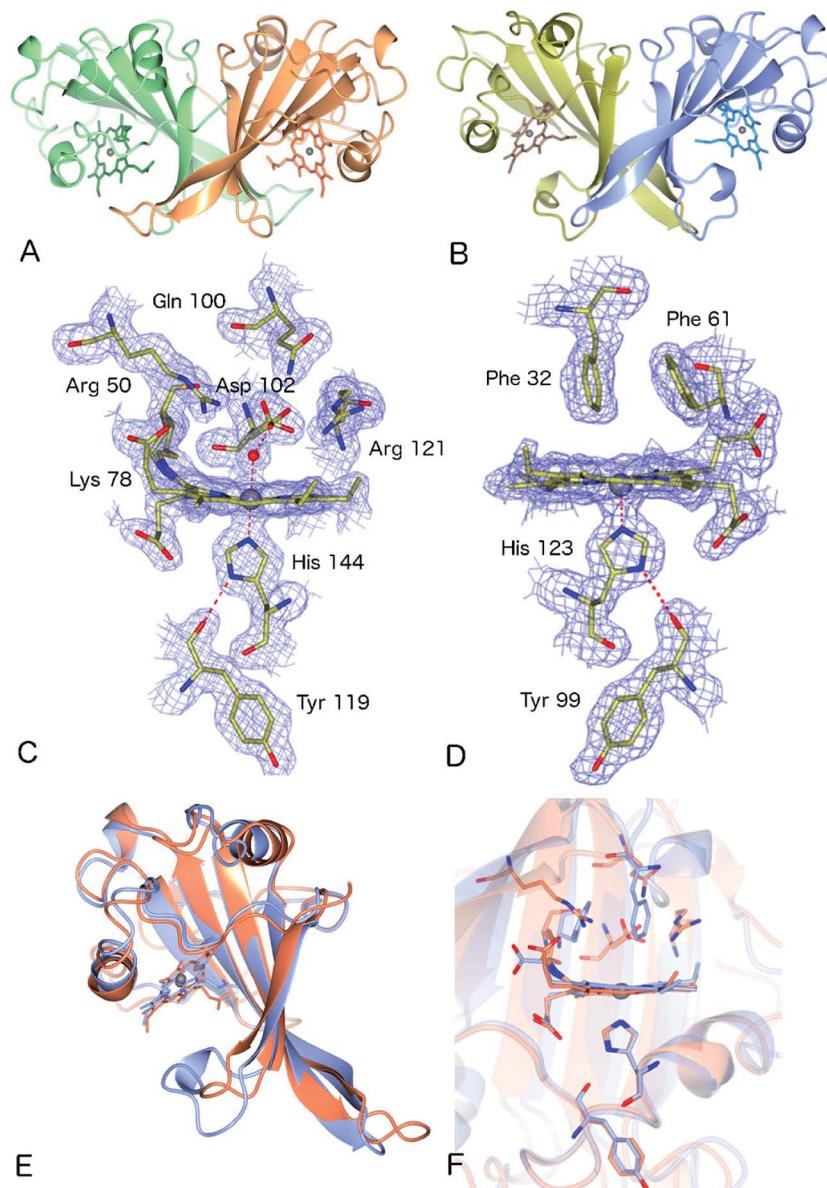
In this study we present high resolution crystal structures of McCP- $\beta$  and McP460 in their as-isolated states, representing the first crystal structure of any cyt cp- $\beta$ , and the first for a methanotrophic cyt P460. Despite possessing similar  $\beta$ -sheet protein folds, the heme environment of McCP- $\beta$  has undergone major structural changes relative to McP460, replacing a highly charged and hydrophilic distal heme pocket with a hydrophobic pocket where two Phe residues act as a 'cap' over the distal NO binding position. Characterization of McCP- $\beta$  and McP460 using UV-visible absorption, EPR and resonance Raman spectroscopies highlights differences in their heme environments, while facilitating comparison with previously characterized proteins from other organisms. Finally, our phylogenetic analysis identifies the occurrence and evolutionary history of these proteins and provides clues to their functional roles. Analysis of these changes in the context of an otherwise highly conserved protein fold provides insights for synthetic enzyme design and synthetic biology.

## Results and discussion

### Crystal structure of McP460 and comparison with other cyts P460

The crystal structure of as-isolated (ferric) McP460 was determined to 1.36 Å resolution (Fig. 1, Table 1) with residues 19 to 161 observed in the electron density maps. McP460 exists as a homodimer, with each monomer consisting of a twisted five-stranded antiparallel  $\beta$ -sheet, two  $\alpha$ -helices and two helical turns, Fig. 1A – a protein fold similar to that of NeP460<sup>12</sup> and NsALP460<sup>13</sup> from AOB. The defining P460 covalent cross-link is evident between Lys 78 and the 13'-*meso* carbon of the McP460 porphyrin ring (Fig. 1C). In agreement with a recent study of NsALP460, structural parameters for the *meso* carbon correspond to sp<sup>2</sup> hybridization, consistent with the retention of porphyrin aromaticity (Table S4†). No hydroxyl modification of the porphyrin (observed at the 5'-*meso* carbon position in NeP460, but not in NsALP460) is evident in our McP460 structure. The proximal pocket of McP460 is also quite similar to that of NeP460 and NsALP460, a minor difference being that the His 144 ligand of McP460 is H-bonded to the Tyr 119 peptide carbonyl, whereas the His ligands of NeP460 and NsALP460 are H-bonded to a Met carbonyl.

The distal side of the McP460 heme is easily accessible to solvent as there is a large opening on the side of the protein



**Fig. 1** The homodimeric structures of (A) McP460 and (B) McCP- $\beta$  showing the predominantly  $\beta$ -sheet fold; (C) the heme environment in McP460, note the distal water ligand and hydrophilic, charged pocket with several Arg and Asp residues in a position to interact with bound substrates; (D) the heme environment of McCP- $\beta$  showing the hydrogen bonding between the proximal His 123 and the backbone carbonyl of Tyr 99. The distal pocket is devoid of water molecules and formed of hydrophobic residues including the 'Phe Cap' of Phe 32 and Phe 61; (E) superposition of McCP- $\beta$  (blue) and McP460 (coral) overall structures; (F) superimposition of the heme environments of McCP- $\beta$  (blue) and McP460 (coral).

(Fig. S1†). Within the distal heme pocket are a large number of polar residues (Arg 43, Gln 46, Asn 48, Arg 50, Gln 100, Asp 102 and Arg 121) (Fig. 1C)), with a well-defined water molecule bound at a distance of 2.43 Å (Table 2) to produce a 6-coordinate Fe species, in a manner similar to that of the HAO P460 cofactor.<sup>10</sup> The distal water ligand in McP460 is 3.1/3.2 Å from residue Asp 102, and we note that a bound hydroxylamine substrate displacing the distal water would likely be in hydrogen-bonding distance to several additional pocket residues such as Arg 43, Arg 50, Asp 102 or Arg 121. Interestingly, the distal heme environment of McP460 exhibits significant

differences relative to cyts P460 characterized from AOB. Notably, NsALP460 has a much more hydrophobic distal pocket environment than McP460 (or NeP460), without residues likely to form a stabilising hydrogen bonds to distal ligands, Fig. S2.† Indeed, in certain respects, the NsALP460 structure resembles McCP- $\beta$  as much as it does NeP460 or McP460 (see below). The NsALP460 heme is 5c with no evidence of a distal water molecule or other ligand,<sup>13</sup> while in the NeP460 crystal structure,<sup>12</sup> a phosphate anion from the crystallisation medium obscures the coordination of the Fe(III) resting state. By contrast, our structure of McP460 clearly shows a 6c Fe(III) heme resting state,

**Table 1** Data collection and processing statistics for the crystal structure of McCP- $\beta$  and McP460

Dataset	McCP- $\beta$	McP460
Resolution (Å)	61.18–1.61	59.0–1.36
Space group	$P2_13$	$P2_12_12_1$
Unit cell, (Å)	$a = b = c = 106.0$	$a = 46.6$ $b = 81.6$ $c = 85.3$
Unique reflections	52 377	70 585
Completeness (%)	99.9 (99.4)	99.9 (98.5)
Redundancy	7.6 (4.0)	5.4 (4.4)
$R_{\text{merge}}$	0.047 (0.93)	0.043 (1.048)
$I/\sigma(I)$	20.9 (1.4)	13.7 (1.2)
$CC_{1/2}$	0.99	1.00 (0.49)
Wilson B-factor (Å <sup>2</sup> )	24.3	16.0
$R_{\text{work}}$	0.177	0.197
$R_{\text{free}}$	0.202	0.224
RMSD bond length (°)	0.015	0.021
RMSD bond angles (Å)	2.15	1.84
Ramachandran favoured (%)	96.1	94.6
PDB accession code	6HIH	6HIU

similar to that of the HAO P460 cofactor. The presence of a water ligand in NeP460 remains an open question and awaits determination of a structure for that enzyme in the absence of phosphate.

### Crystal structure of McCP- $\beta$ : a P460-like fold with a hydrophobic distal pocket

The crystal structure of McCP- $\beta$  was solved by SAD using the anomalous signal of the intrinsic heme Fe, and the structure determined to 1.61 Å resolution. McCP- $\beta$  displays a highly  $\beta$ -sheet fold that is quite unlike any known cytochrome  $c'$  structures (all previous examples being 4  $\alpha$ -helix bundles), yet very similar to McP460 (Fig. 1 and S3†). McCP- $\beta$  exhibits a novel distal pocket structure relative to NeP460, McP460 and all cyt cp- $\alpha$  structures. The pocket is comprised of hydrophobic residues, with two aromatic Phe residues (Phe 32 and Phe 61) centrally located above the Fe in a ‘capping’ position well-suited to sterically influencing the binding of diatomic ligands (we name this structural feature the ‘Phe Cap’). This residue arrangement is different from that of the single Phe residue found to lie near-parallel to and above the heme in certain cyt cp- $\alpha$  structures. In contrast to the structures of cyt P460, no heme-protein crosslink is observed in the McCP- $\beta$  structure and the equivalent residue, Phe 61, now forms part of the Phe cap. The proximal pocket contains the conserved haem ligand His 123 and an aromatic Phe 133 residue (oriented perpendicular to the imidazole plane), analogous to all reported cyt P460

structures this arrangement differs from that of cyt cp- $\alpha$  in which a conserved positive, polar residue (Arg or Lys) lies adjacent to the His ligand. The McCP- $\beta$  structure also shows that the His ligand forms a H-bond to the carbonyl of Tyr 99, analogous to the situation in McP460 (Fig. 1). We note that the His ligands of NeP460 and NsALP460 are also H-bonded to a peptide carbonyl (albeit from a Met residue).

The thermal stability difference between McCP- $\beta$  and McP460 may be considered in relation to their respective structures. McP460 has a significantly more hydrophilic pocket, which may allow water ingress and aid in denaturation. McCP- $\beta$  in contrast is highly hydrophobic, with water molecules excluded from the heme region.

### Spectroscopic characterization and thermostability

As-isolated Fe(III) McP460 (pH 7.0) exhibits a Soret absorption band at 419 nm, with a shift to the characteristic 460 nm band in the Fe(II) state (Fig. S4, Table S1†). The  $\lambda_{\text{max}}$  value for the Fe(III) state is significantly blue-shifted relative to that of NeP460 (434–440 nm) and NsALP460 (440 nm) (Table S1†), a fact that we attribute to the presence of a distal H<sub>2</sub>O ligand in McP460. Room-temperature Resonance Raman (RR) spectra of fresh McP460 were obtained at room temperature using laser excitations at 407 nm (for the ferric state) and 442 nm (for the ferrous state) (Fig. 2a and b). A notable feature of the P460 RR spectra (particularly apparent in the Fe(II) state) is the increased number of observable porphyrin vibrations relative to McCP- $\beta$  (Fig. 2c and d), a consequence of the Lys cross-link which lowers the porphyrin symmetry. Although the porphyrin modes of P460 hemes have yet to be systematically assigned, the most intense porphyrin RR bands of the Fe(III) (1367 cm<sup>−1</sup>) and Fe(II) (1352 cm<sup>−1</sup>) states are characteristic of  $\nu_4$  redox marker modes, with the 1566 cm<sup>−1</sup> band of the Fe(III) state resembling the  $\nu_2$  spin-state marker mode (Fig. 2a, Table S2†). The low frequency RR region also contains numerous RR bands (Fig. 2b). Unlike other heme proteins in which the most intense band in the 200–700 cm<sup>−1</sup> region is the  $\nu_7$  mode (typically observed near 680 cm<sup>−1</sup>), Fe(III) McP460 exhibits its most intense low-frequency RR band at 709 cm<sup>−1</sup>, while that of the Fe(II) state is at 785 cm<sup>−1</sup>. Recent RR measurements of Fe(III) NeP460 (using 458 nm excitation) also showed a complex RR spectrum.<sup>15</sup> However, in contrast to McP460, the RR modes of NeP460 exhibited different intensity patterns, and no porphyrin marker bands could be assigned. Neither could porphyrin marker bands be identified in RR spectra of the Fe(II) P460 fragment of NeHAO.<sup>26</sup> While the exact cause of these differences is unclear, it likely results from differences in porphyrin

**Table 2** Heme site parameters for McCP- $\beta$  and P460 structures

Structure	McCP- $\beta$ (6HIH)	McP460 (6HIU)	NeP460 (2J3E)	NeALP460 (6AMG)
Fe–His N (Å)	2.10/2.09	2.06	2.25	2.15
Fe–water (Å)	—	2.43	—	—
Lys–heme (Å)	—	1.32/1.36	1.56	1.37



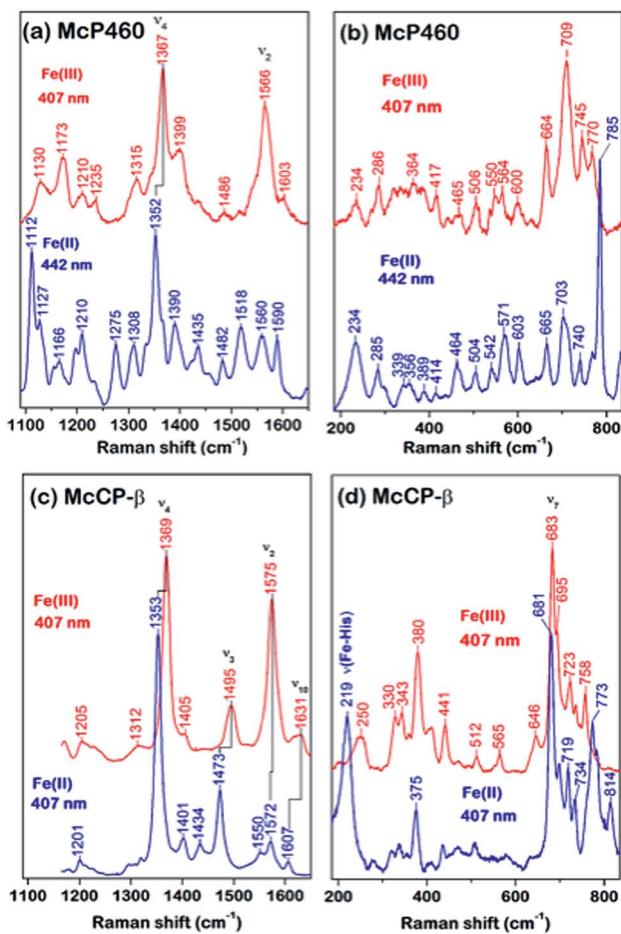


Fig. 2 RR spectra of ferric and ferrous McP460 in the high-frequency (a) and low frequency (b) regions; RR spectra of ferric and ferrous McCP-β in the high-frequency (c) and low-frequency (d) regions. Data obtained at room temperature using either 407 nm or 442 nm excitation as indicated. Solutions contained  $\sim$ 150  $\mu$ M protein (in haem) in pH 7.0 buffer (50 mM MOPS).

conformation/symmetry between the different P460 cofactors (*vide infra*).

A low temperature EPR analysis of McP460 showed a spectrum representative of a single high spin species (Fig. S5†) with  $g_1 = 6.18$ ,  $g_2 = 5.57$  and  $g_3 = 1.99$  resembling that of an isolated enzyme.<sup>27</sup> Table S3.† In our hands McP460 undergoes a spectroscopic shift if not snap-frozen relatively soon following the final stage of purification, with the 'degraded' variant exhibiting absorption maxima at 415 nm in the Fe(III) form and 455 nm in the Fe(II) state, with a colour change from green (fresh protein) to brown (degraded protein) (Fig. S6†). This process also affects the RR spectra (Fig. S7†) and is discussed in more detail in the ESI.† We note here that all data presented are, unless stated otherwise, for the fresh, green form of McP460.

UV-visible absorption bands of as-isolated recombinant McCP-β (Fig. S8, Table S1†) are similar to those previously reported for the native protein,<sup>6</sup> with a Soret maximum at 399 nm (and weak shoulder at  $\sim$ 379 nm), together with an unresolved  $\alpha/\beta$  band near 500 nm and a CT1 band at  $\sim$ 640 nm, the latter characteristic of 5c high-spin (5cHS) ferric heme. The

spectroscopic properties of Fe(III) McCP-β remain essentially unchanged over the pH range 4.0–10.0 (Fig. S9†). Resonance Raman (RR) spectra of Fe(III) McCP-β (obtained with 407 nm excitation) reveal porphyrin marker band frequencies:  $\nu_4$  (1369  $\text{cm}^{-1}$ ),  $\nu_3$  (1495  $\text{cm}^{-1}$ ),  $\nu_2$  (1575  $\text{cm}^{-1}$ ),  $\nu_{10}$  (1631  $\text{cm}^{-1}$ ) (Fig. 2c upper trace, Table S2†). The  $\nu_3$  and  $\nu_{10}$  spin-state frequencies are close to those attributed to the 5cHS Fe(III) states of cys cp- $\alpha$  ( $\nu_3 \sim$ 1490–1494  $\text{cm}^{-1}$  and  $\nu_{10} \sim$ 1625–1629  $\text{cm}^{-1}$ ).<sup>28</sup> The EPR spectrum of Fe(III) McCP-β at pH 9 shows a mixture of two high spin species (Fig. S10, Table S3†) consistent with previous reports.<sup>29</sup>

Reduction of Fe(III) McCP-β with dithionite yields the Fe(II) state which exhibits a Soret maximum at 431 nm and a broad  $\alpha/\beta$  band centered at 552 nm (Fig. S8, Table S1†). RR spectra of Fe(II) McCP-β (obtained with 407 or 442 nm excitation) resemble those of Fe(II) cys cp- $\alpha$ , and have frequencies characteristic of 5cHS Fe(II) hemes:  $\nu_4$  (1353  $\text{cm}^{-1}$ ),  $\nu_3$  (1473  $\text{cm}^{-1}$ ),  $\nu_2$  (1572  $\text{cm}^{-1}$ ),  $\nu_{10}$  (1607  $\text{cm}^{-1}$ ) (Fig. 2c lower trace, Table S2†). In the low frequency region, the  $\nu(\text{Fe-His})$  vibration of Fe(II) McCP-β is evident at 219  $\text{cm}^{-1}$ , which is  $\sim$ 10–15  $\text{cm}^{-1}$  lower than for Fe(II) cyt cp- $\alpha$  proteins,<sup>30</sup> suggesting that McCP-β has a weaker proximal bond. The Fe–His bond strength in McCP-β may have important ramifications for its proposed heme–NO binding role, and in particular its ability to form 5c vs. 6c  $\{\text{FeNO}\}^7$  complexes. It should also be noted that McCP-β does not degrade in the same fashion as McP460.

The circular dichroism (CD) spectra of McP460 and McCP-β showed negative peaks from 205 to 225 nm. From the spectral data, the ratios of  $\alpha$ -helix :  $\beta$ -sheet : others in McP460 and McCP-β were estimated as 9 : 36 : 65 and 6 : 36 : 58 (Fig. 3a), respectively. Previously reported ratios for NeP460 also estimated a high level of  $\beta$ -sheet, although the amount of  $\alpha$ -helix was estimated to be much higher than that of McP460 and McCP-β.<sup>5</sup> This indicates that these proteins have similar content of  $\beta$ -sheet as observed in these crystal structures.

The thermal stability of the two proteins was characterised by CD ellipticity changes at 222 nm (Fig. S11†). The melting temperature ( $T_m$ ) values obtained were  $58.4 \pm 2.8$  °C for McP460 and  $64.4 \pm 3.2$  and  $94.1 \pm 0.9$  °C for McCP-β (Fig. 3b), indicating that the latter has a two-step transition during its thermal denaturation and more stable than the former. The unfolding

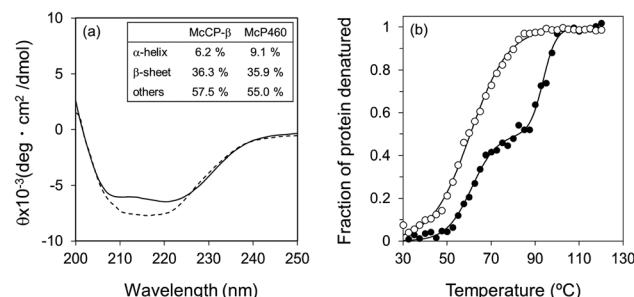


Fig. 3 CD spectra (a) of McCP-β (dashed line) and McP460 (solid line). (b) Thermal denaturation curves for McCP-β (closed circles) and McP460 (open circles) as measured by CD. Representative normalized raw data plots are shown at the temperature interval of 2.5 °C.



transitions are much broader for McP460 and McCP- $\beta$  as is typical for  $\beta$ -sheet proteins in comparison to those of cys cp- $\alpha$ .<sup>31,32</sup> We note for context that the optimum growth temperature of *M. capsulatus* (Bath) is 45 °C.<sup>33</sup>

### Evolution of cyt P460 and cyt cp- $\beta$ proteins

McCP- $\beta$  has 31% sequence identity to McP460 and also possesses 18% identity to cyt P460 from *N. europaea*. In contrast, McCP- $\beta$  has very low sequence identity to the well-characterised cyt cp- $\alpha$  proteins. A maximum likelihood phylogenetic tree displays maximum support for the branch separating a clade containing the cyt cp- $\alpha$  genes from a group comprising the cyt cp- $\beta$  and the cys P460 genes (Fig. 4). The cyt P460 family shows a paraphyletic structure, with a monophyletic cyt cp- $\beta$  (44% bootstrap support) nested within. The group of P460 sequences which includes McP460 is not closely related to the cys cp- $\beta$  clade. Another group of P460 sequences, which contains both NeP460 and NsALP460, are more closely

related to the cys cp- $\beta$  (49%) instead. It can be inferred from this result that the cyt cp- $\beta$  proteins evolved from the cyt P460, as the latter are nested several nodes within the cyt P460, although these nodes have moderate bootstrap supports. Notably this topology is different to the one recovered in a previous study;<sup>5</sup> this may be explained by the use of a much larger dataset in our study, together with more sophisticated phylogenetic methods (e.g., maximum likelihood instead of neighbour joining). There are 14 organisms within the aligned sequences which possess both a cyt P460 and a cyt cp- $\beta$ . These genes are not close to each other on the genome, arguing against a recent gene tandem duplication event. In all but two of these (*Nitrosococcus halophilus* and *Methyloimonas methanica*) the cyt P460 sequences are placed on a clade that is not related to the cys cp- $\beta$  genes.

Alignment of all 144 cyt P460 sequences showed full conservation of only 5 residues (Trp 24, Gly 73, Lys 78, Lys 115 and Trp 17 in McP460) along with the CXXCH motif. The

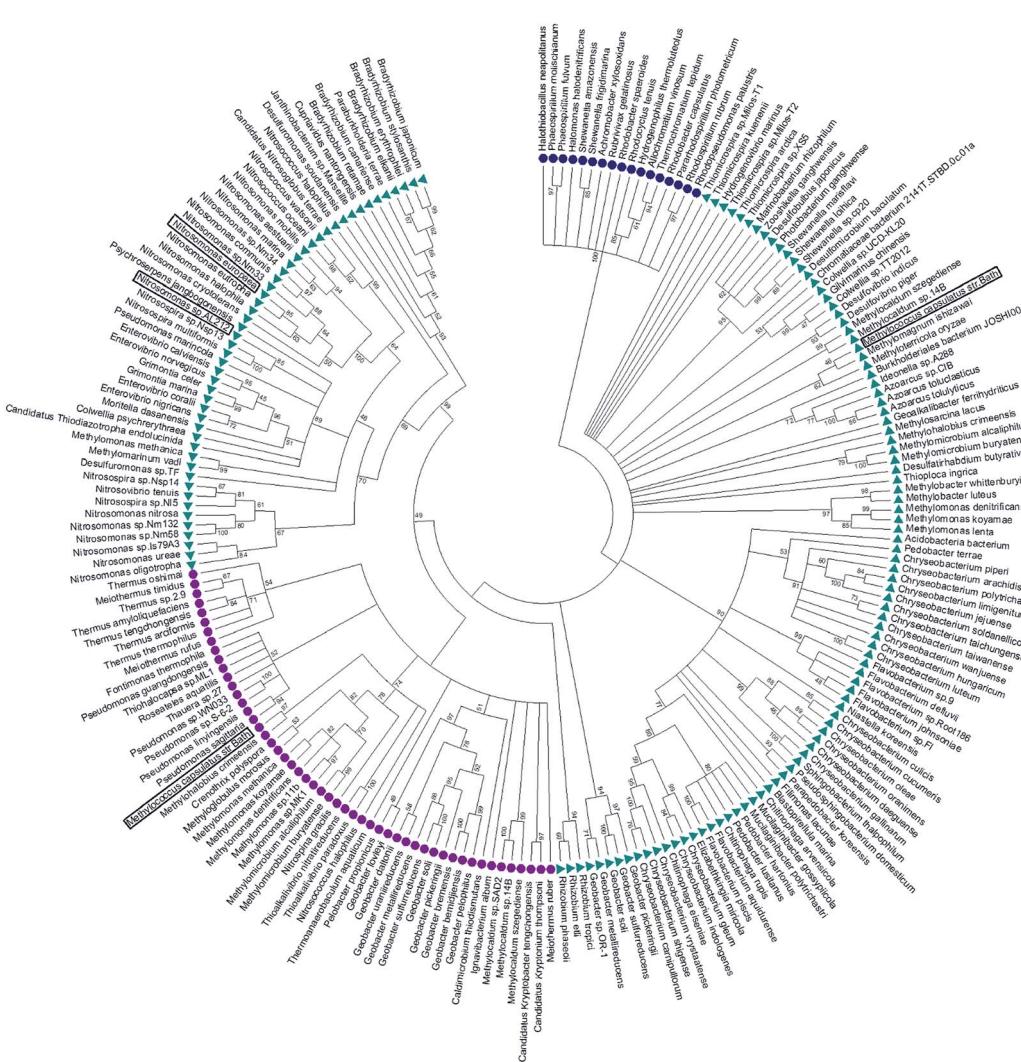


Fig. 4 Maximum likelihood tree of cyt cp- $\alpha$  (blue circles), cyt cp- $\beta$  (purple circles) and cyt P460 (green triangles) sequences. The cyt cp- $\alpha$  genes are on a separate branch to the cyt cp- $\beta$  and the cys P460 genes. The cyt P460 family displays a paraphyletic structure, with a monophyletic cyt cp- $\beta$  nested within suggesting that the cyt cp- $\beta$  evolved from the cyt P460. Image was prepared using Mega7.<sup>54</sup>



conserved lysine cross link in cyts P460 (Lys 78 in McP460) has been shown to be important in relation to the catalytic ability of the protein,<sup>25</sup> whilst most of the other conserved residues are, as in the P460 of *N. europaea*, clustered at the base of the monomer.<sup>12</sup> The group of cyts P460 that appear more closely related to the cyts cp- $\beta$  have a higher level of homology to each other with an additional 20 conserved residues. The residues present in the distal pocket of NsALP460 (Phe 76 and His 80), are highly conserved within this subgroup (being present in 45 and 39 (out of 50) of the sequences respectively) with the alternative residues having similar properties.

The other group of cyts P460 (containing McP460) show five further conserved residues (Arg 50, Asn 55, Ala 58, Pro 71, Met 104 in McP460) when aligned together in comparison to the alignment of all 144 cyt P460 sequences. Two of these residues are positioned near the heme of McP460, Met 104 being to the side of the heme in the proximal pocket and Arg 50 sitting directly above the heme in the distal pocket. The type of residue in the distal pocket for this subgroup of P460s appears reasonably conserved with the majority of them being charged, hydrophilic, polar residues. Arg 50 for example is present in 85 (out of 95) of these sequences. As previously shown,<sup>5</sup> most of the cyt P460 and cyt cp- $\beta$  sequences are from a wide range of proteobacteria.

### Structure-function relationships in cyt P460: implications for synthetic enzyme design

McP460 and NeP460 are well separated in the phylogenetic tree of P460 enzymes yet both are able to oxidise hydroxylamine, as can the unrelated enzyme HAO. In the mechanism proposed for hydroxylamine oxidation in NeP460,<sup>14</sup> an initial  $\text{Fe}(\text{III})-\text{NH}_2\text{OH}$  substrate complex undergoes a three electron oxidation to form a  $\{\text{FeNO}\}^6$  (Enemark-Feltham notation) intermediate *via* a 6c  $\{\text{FeNO}\}^7$  precursor. Nucleophilic attack of the  $\{\text{FeNO}\}^6$  species by a second  $\text{NH}_2\text{OH}$  molecule, together with an additional one-electron oxidation, results in the formation of  $\text{N}_2\text{O}$  product and the  $\text{Fe}(\text{III})$  heme resting state.

The distinguishing structural feature of P460 cofactors is the Lys(Tyr) porphyrin cross-link. Notably, in NeP460, mutation of the cross-linking Lys to Arg, Ala or Tyr resulted in catalytically inactive proteins with spectroscopic properties similar to those of the cyts cp.<sup>25</sup> It was recently proposed that the Lys-heme cross-link in NeP460 provides structural rigidity and prevents formation of an off-pathway 5c  $\{\text{FeNO}\}^7$  species. Out of plane heme deformations are another feature of P460 centers. Normal-coordinate structural decomposition (NSD) analyses of crystallographic data from NeP460 (J2E3) and NsALP460 (6AMG)<sup>34</sup> revealed that these P460 heme cofactors have relatively high degrees of ruffling (B1u) and saddling (B2u) distortions,<sup>13</sup> whereas the P460 center of NeHAO (4FAS) is distorted predominantly by ruffling. It has been proposed that a high degree of ruffling (rather than saddling) is a hallmark of P460 cofactors. Consistent with this proposal, NSD analysis of our McP460 structural data also reveal a high degree of ruffling with relatively little saddling, whereas McCP- $\beta$  exhibits a degree of saddling comparable with NeP460 and NsALP460, but has very little ruffling (Fig. 5, Table S5†). It has been further proposed

that heme properties such as lower reduction potentials<sup>35</sup> and stronger bonding between the Fe and axial ligand<sup>36</sup> are associated with increased heme ruffling, suggesting that these distortions are important to P460 function. In particular, a low reduction potential has been proposed as the reason why HAO and P460 avoid auto-reduction upon NO binding. In this context, we also note that McP460 has a significantly lower  $E^\circ$  value ( $-380$  to  $-300$  mV) than that of McCP- $\beta$  ( $-205$  mV),<sup>6</sup> although structural differences other than ruffling (e.g. axial heme environments) may also influence heme redox properties.

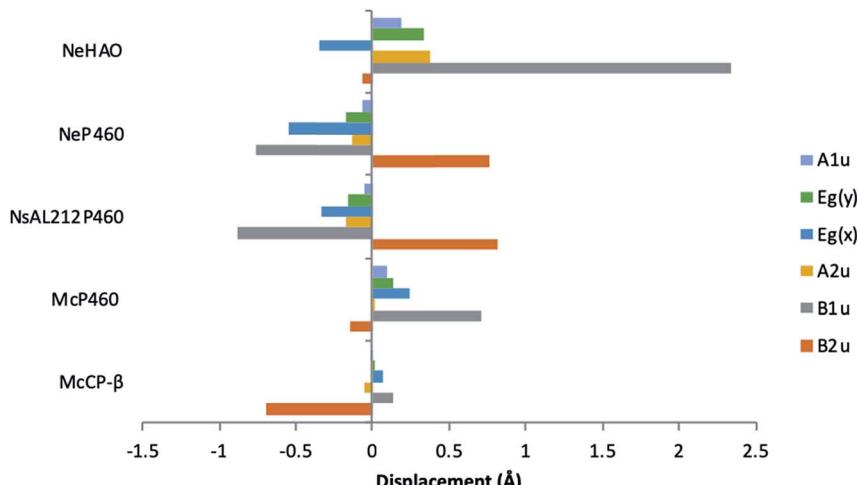
In addition to the presence of a porphyrin-Lys (or Tyr) cross-link, crystal structures of McP460, NeP460, and HAO show polar distal heme pockets with side-chains positioned to form hydrogen bonds with exogenous ligands. For example, in McP460, a 3.1 Å hydrogen bond is formed between Asp 102 and the heme-bound water molecule (Fig. 1). Other distal ligands relevant to the mechanism, such as hydroxylamine, NO or  $\text{N}_2\text{O}$  would likely be in a position to form hydrogen bonds with several Arg or Asp residues within the pocket, Fig. 1. Interestingly, the recently determined structure of NsALP460 (PDB 6AMG) revealed a quite different distal pocket environment. The heme retains its P460 character with the presence of a Lys cross-link, but the distal pocket is far more hydrophobic with residue Phe 76 in a similar position to that of Phe32 in McCP- $\beta$ . It is conceivable that NsALP460 could provide clues as to the evolutionary divergence between McCP- $\beta$  and cyts P460, representing a half-way house situation. Testing of this hypothesis awaits the demonstration of enzymatic activity in NsALP460. We also note that a conserved Tyr residue in the heme pocket of HAO from AOBs has been proposed to play a key role in the catalytic mechanism.<sup>37</sup> In methanotroph HAOs this residue is absent, leading to suggestions that the final stage of the HAO mechanism is arrested. In P460 structures, the equivalently located residue is an Arg (50 McP460 and 44 NeP460) or His (NsAL212-P460). The extent to which enzymatic activity in P460 heme proteins is related to distal pocket structure *versus* heme geometry will be addressed in future studies.

### Functional evolution of cyt cp- $\beta$ : a possible NO binding role?

The availability of high resolution crystal structures of a cyt cp- $\beta$  and a cyt P460 from the same organism allows the architecture of the two related proteins to be considered in a structure-function context independent of species differences. In contrast to McP460, there is no porphyrin-Lys cross-link or potential H-bonding residue to interact with heme-bound water or diatomic ligands in McCP- $\beta$ , and indeed the heme is 5-coordinate in the resting state with a vacant distal heme binding site.

There is low sequence conservation within the cyt cp- $\beta$  family, with 7 other residues in addition to the CXXCH motif being fully conserved across the 52 aligned cyt cp- $\beta$  sequences (Trp14, Gly34, His36, Tyr39, Met84, Thr94, Trp97 in McCP- $\beta$ ). This is similar to the low conservation of sequence in cyts cp- $\alpha$ .<sup>22</sup> Phe 32 is highly conserved within cyt cp- $\beta$ , being present in all but one of the aligned sequences (*Meiothermus ruber*). Phe 61 is slightly less conserved, being present in 39 of the sequences, although all are hydrophobic residues. Within the proximal





**Fig. 5** Out-of-plane displacements (minimal basis) for the hemes of McCP- $\beta$ 1 (6HIH), McP460 (6HIU), NsALP460 (6AMG), *N. europaea* P460 (J2E3) and the P460 heme of *N. europaea* HAO (4FAS). Out of plane distortions are characterized in terms of displacements along the normal coordinates of the  $D_{4h}$ -symmetric porphyrin macrocycle. NsALP460 and NeP460 display much higher overall distortions to both *M. capsulatus* (Bath) cytochromes. McP460 displays a high amount of ruffling (B1u) which has been proposed to be important in the function of the P460 cytochromes.<sup>13</sup> In comparison to the other two cytochromes P460, McP460 does not have an equal amount of ruffling and saddling, a pattern that is more similar to that of the P460 heme of *N. europaea* HAO. McCP- $\beta$  correspondingly does not display a high level of ruffling, although it does demonstrate a large amount of saddling (B2u). See also Table S5.†

pocket, Tyr 131 is conserved in just over two thirds of the cyt cp- $\beta$  sequences (36 out of 52), although all other residues are hydrophobic (either Phe or Leu). In addition, Phe 133 is conserved in 38 of the 52 cyt cp- $\beta$  sequences.

No organisms identified in our study have both a cyt cp- $\beta$  and a cyt cp- $\alpha$  (or indeed a cyt P460 and cyt cp- $\alpha$ ). Between the two cyt cp families (alpha and beta) the only conservation is in the CXXCH motif. Nevertheless, the conserved 'Phe cap' in McCP- $\beta$  is reminiscent of the occluding distal aromatic or aliphatic residue in cyt cp- $\alpha$ . In the case of cyts cp- $\alpha$ , the highly hydrophobic and sterically crowded distal pockets are important in their ability to discriminate between gaseous ligands by allowing small neutral ligands, such as NO and CO to bind to the heme.<sup>16</sup> The cyt cp- $\alpha$  family of proteins occur in denitrifying, photosynthetic, methanotrophic, sulphur-oxidising and nitrogen-fixing bacteria. In each of these groups, the ability to bind NO may have functional importance. In denitrifiers, NO is produced by nitrite reductases (NiRs) and further reduced to N<sub>2</sub>O by nitric oxide reductase (NOR) enzymes. It has been proposed that cyt cp- $\alpha$  could act as a shuttle carrying this potentially dangerous free radical from NiR to the membrane-bound NOR. There is also evidence that cyt cp- $\alpha$  could itself act as an NO-reductase. Recent studies suggesting that HAO produces NO as an end product suggests a role for an NO-sequestering protein such as McCP- $\beta$  to work in concert with HAO deal with excessive NO levels. The introduction of a hydrophobic distal heme environment by evolution in the cyt cp- $\beta$  family is consistent with a functional requirement for the heme to selectively bind NO. We note that McCP- $\beta$  has previously been shown to form complexes with both NO and CO.<sup>6</sup> Notably, in contrast to cyts cp- $\alpha$ , in both McCP- $\beta$  and McP460 the proximal His forms a H-bond to a main chain carbonyl

which could influence any dissociation from Fe upon distal NO binding, a process shown to occur in cyts cp- $\alpha$ .

## Conclusion

We have characterised the closely related cytochrome P460 and cytochrome  $c'$ - $\beta$  proteins from *Methylococcus capsulatus* (Bath). High resolution crystal structures reveal a highly similar tertiary structure between the two proteins but major structural differences around the heme pocket consistent with a switch of function between hydroxylamine oxidation (P460) and NO binding/sequestration (cyt cp- $\beta$ ). Interestingly, our structures reveal that the protein fold of McP460 is rather more similar to McCP- $\beta$  than it is to NeP460, with the exception of the heme binding pocket where a number of localised structural alterations are associated with functional conversion. In particular, the lysine that forms the heme cross-link of McP460 has been replaced in McCP- $\beta$  by a Phe residue that instead provides steric hindrance to distal ligand binding as part of the 'Phe Cap' motif. The heme pocket of McCP- $\beta$  is much less solvent-exposed than in McP460 with a highly hydrophobic distal pocket with no water molecules observed within the McCP- $\beta$  heme pocket. Phylogenetic analysis of the occurrence of these two proteins suggests an evolutionary path such that the cyt cp- $\beta$  (*CytL* gene) proteins evolved from the cyts P460 (*CytS* gene), likely in response to the need to capture excess nitric oxide.

## Materials and methods

### Expression and purification

The genes encoding McCP- $\beta$  and McP460 (Accession numbers AAD39218 and AAU93287 from the NCBI RefSeq database) were



commercially synthesised (Epoch Life Sciences) in a pBSK plasmid with restriction enzyme sites (EcoRI and XbaI respectively) flanking either end of the gene. The gene was cut from this plasmid and ligated into the expression vector pMMB503EH.<sup>38</sup> *E. coli* XL-1 Blue cells (Stratagene USA) were used for general cloning, and were routinely grown at 37 °C in Luria Broth (LB) medium or plates. When required, antibiotics were used at the following final concentrations: streptomycin (50 µg ml<sup>-1</sup>), chloramphenicol (20 µg ml<sup>-1</sup>). *E. coli* BL21(DE3) (Novagen) competent cells containing pEC86,<sup>39</sup> a plasmid carrying chloramphenicol resistance and the ccmA-H genes from *E. coli* necessary for the maturation of c-type cytochromes, were transformed with the McCP-β containing pMMB503EH vector. Cells were grown in LB or Terrific Broth (TB) and shaken at 200 rpm at 37 °C. For protein production, IPTG was added in mid-log phase to a final concentration of 119 µg ml<sup>-1</sup>. Ferri-protoporphyrin IX chloride was added to a final concentration of 30 µg ml<sup>-1</sup> after 5 h and 1 ml of a metal ion mix (2 mM Ni<sup>2+</sup>, 2 mM Co<sup>2+</sup>, 10 mM Zn<sup>2+</sup>, 10 mM Mn<sup>2+</sup> and 50 mM Fe<sup>3+</sup>) was added after 24 h in a procedure adapted from ref. 40.

Cultures were grown for 96 h and cells harvested by centrifugation for 20 min at 4 °C and at 4000 rpm. Harvested cells were resuspended in 20 mM Tris-HCl pH 8, and disrupted by 2 passages through an EmulsiFlex (12 000 psi). The crude McCP-β extract was prepared by centrifugation at 18 000 rpm for 30 min followed by 24 h dialysis against ddH<sub>2</sub>O. The extract was then centrifuged at 18 000 rpm for 20 min and loaded onto a DEAE Sepharose Fast Flow anion-exchange column previously equilibrated with two column volumes of 20 mM Tris-HCl, pH 8. McCP-β was eluted using 10 mM NaCl and was further purified to homogeneity by gel filtration using a Sephadex G75 column. A final yield of 6.4 mg L<sup>-1</sup> McCP-β was obtained. The crude McP460 extract was prepared by centrifugation at 18 000 rpm for 30 min followed by the addition of 70% ammonium sulfate. The extract was then centrifuged for 20 min at 18 000 rpm and loaded onto a phenyl-sepharose hydrophobic interaction column previously equilibrated with two column volumes of 0.5 M ammonium sulfate, 50 mM sodium phosphate pH 8. McP460 was eluted over a decreasing salt gradient and further purified to homogeneity by gel filtration using a Sephadex G75 column. The purity of the proteins was determined by SDS-PAGE analysis and by UV-vis spectroscopy using published molar absorptivity of McP460 and McCP-β heme Soret bands: Fe(III) McP460:  $\epsilon_{419} = 78.5$  mM cm<sup>-1</sup>;<sup>29</sup> Fe(III) McCP-β:  $\epsilon_{400} \sim 70$  mM cm<sup>-1</sup>.<sup>6</sup>

### Protein crystallisation and X-ray crystallography

Crystals of McCP-β were grown using the hanging drop vapour diffusion method at 18 °C. Brown cubic crystals were obtained after 3–5 days using 2 µl of 15 mg ml<sup>-1</sup> protein in 0.1 M HEPES buffer, pH 7.5, with an equivalent volume of reservoir solution containing 0.01 M ZnSO<sub>4</sub>, 35% PEG 550 (v/v) and 0.1 M MES, pH 6.5. Prior to X-ray data collection crystals were cryoprotected by transfer to reservoir solution comprising mother liquor supplemented with 10% (v/v) glycerol and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at beamline I03,

Diamond Light Source, using a Pilatus 3 6 M detector. A dataset measured at a wavelength of 1.74 Å was used to solve the structure using the anomalous signal of the intrinsic heme Fe. All data were processed in Xia2 using XDS<sup>41</sup> and AIMLESS.<sup>42</sup> The initial model was built using Buccaneer.<sup>43</sup> A high-resolution dataset was measured from a second MCCP crystal at a wavelength of 0.97 Å and refined from the starting model from the SAD experiment. Structures were refined by maximum likelihood methods using REFMAC5.<sup>44</sup> Model building between cycles of refinement, including addition of water molecules and ligands was performed in Coot,<sup>45</sup> and the quality of the structures was monitored using the MOLPROBITY<sup>46</sup> and the JCSG Quality Control servers. Data collection and refinement statistics are shown in Table 1. Crystals of McP460 were grown using the hanging drop vapour diffusion method at 18 °C. Green orthorhombic crystals were obtained after 4–5 days using 2 µl of 15 mg ml<sup>-1</sup> protein in 20 mM Tris buffer, pH 8, with an equivalent volume of reservoir solution containing 0.1 M Tris, 2.4 M ammonium sulfate pH 8. Prior to X-ray data collection crystals were cryoprotected by transfer to reservoir solution comprising mother liquor supplemented with 1.7 M sodium malonate and flash-cooled in liquid nitrogen. Data were measured at beamline I04-1, Diamond Light Source at a wavelength of 0.979 Å, processed using DIALS and the structure solved using Auto-Rickshaw,<sup>47</sup> with an initial SAD electron density map being used as the target for molecular replacement with the structure of NeP460 [PDB 2je3]<sup>12</sup> as the search model. Model building was completed in Arp-Warp<sup>48</sup> and subsequently refined and validated as described for McCP-β above. Protein Data Bank accession numbers are given in Table 1.

### Spectroscopic analysis

A Varian Cary 50 spectrophotometer was used to measure UV-visible absorption spectra. Reduction of McCP-β and McP460 to the ferrous state was achieved through addition of excess sodium dithionite under anaerobic conditions.

The EPR spectra were measured on a Bruker EMX EPR spectrometer (X band) at 10 K. A spherical high-quality Bruker resonator ER 4122 SP 9703 and an Oxford Instruments liquid helium system were used to measure the low temperature EPR spectra. Wilmad SQ EPR tubes (Wilmad Glass) were filled with the McCP-β or McP460 solutions and frozen in methanol kept on dry ice. The tubes were then transferred to liquid nitrogen. The spectra were measured at the following instrumental conditions: microwave frequency  $\nu_{MW} = 9.4668$  GHz; microwave power  $P_{MW} = 3.17$  mW; modulation frequency  $\nu_M = 100$  kHz; modulation amplitude  $A_M = 5$  G; scan rate  $\nu = 22.6$  G s<sup>-1</sup>; time constant  $\tau = 81.92$  ms; conversion time, at a 2048 data point scan range,  $t_{conv.} = 81.92$  ms.

Protein solutions for RR measurements (100–150 µM in heme) were prepared in 50 mM MOPS buffer (pH 7.0) containing 0.10 M NaCl. Samples of as-isolated Fe(III) proteins were transferred to glass capillary tubes for RR measurements. Reduced proteins were prepared with O<sub>2</sub>-free solutions inside an anaerobic glove box (using excess dithionite to reduce the Fe(III) proteins to the Fe(II) states) and transferred to septum-



sealed glass capillary tubes for RR measurements. RR spectra (room-temperature, 90°-scattering geometry) were recorded on a custom McPherson 2061/207 spectrograph (100  $\mu\text{m}$  slit width, 0.67 m focal length, 2400 grooves/mm holographic grating) equipped with a Princeton Instruments liquid N<sub>2</sub>-cooled (LN-1100PB) CCD detector. Excitation wavelengths were provided by Kr ion laser (406.7 nm) and a He-Cd laser (441.6 nm). RR spectra were recorded for periods of 1–3 min using laser powers of 10–30 mW (measured at sample). An indene standard was used to calibrate Raman shifts to an accuracy of  $\pm 1 \text{ cm}^{-1}$ . The identity of RR samples was confirmed by UV-vis spectroscopy before and after exposure to the laser beam using a modified Cary 50 spectrophotometer.

A JASCO J-820 CD spectrometer was used for circular dichroism (CD) spectrum measurement. The spectra (200–250 nm) of the McCP- $\beta$  and McP460 proteins (20  $\mu\text{M}$ ) in 25 mM sodium acetate buffer pH 5.0 at 25 °C were measured. The components of the protein secondary structures were estimated from the CD spectra using BeStSel program.<sup>49,50</sup>

Thermal stability of McCP- $\beta$  and McP460 was also measured by CD spectra in a pressure-proof cell compartment (Jasco, Tokyo) attached to the CD spectrometer. The protein solutions (20  $\mu\text{M}$ ) were dialysed against 25 mM sodium acetate buffer pH 5.0 at 25 °C. The temperature-dependent CD ellipticity changes at 222 nm were monitored in a cuvette of 1 mm path length. The CD values were recorded from 25 to 120 °C at intervals of 0.5 °C at a heating rate of 1.0 °C min<sup>-1</sup> under the 0.9 MPa pressure conditions. The raw data were subjected to nonlinear least-squares fitting as described previously.<sup>51</sup> The data points were corrected for the slopes of the baselines for the native and denatured forms and were normalized to calculate the fraction of protein denatured. The value for fraction denatured was plotted as a function of temperature, and the resulting thermal denaturation curves were used to determine the temperature at the midpoint of the transition ( $T_m$ ) during protein denaturation.

## Bioinformatics

Protein BLAST<sup>52</sup> was performed May 2018 against NCBI GenBank using as query the sequences *Methylococcus capsulatus* (Bath) cytochrome c' (WP\_010961620.1), *Methylococcus capsulatus* (Bath) P460 (WP\_010959872.1), *Nitrosomonas europaea* P460 (WP\_011110663.1) and *Alcaligenes xylosoxidans* cytochrome c' (WP\_088139370.1). After removal of redundant sequences, full length protein sequences were aligned using ClustalW.<sup>53</sup> Any sequences that did not have the cross-linking Lys residue were removed from the list of cys P460. A total of 214 identified cytochrome c' (69 sequences) and P460 (145 sequences) protein sequences were used in the alignments. Maximum likelihood trees were constructed using RAxML version 8.2.10<sup>54</sup> with the evolutionary model LG + Gamma + Invariants, and the support for each was assessed with 1000 rapid bootstrap replicates. The resulting tree was annotated using MEGA7.<sup>55</sup>

## Normal-coordinate structural decomposition

Out of plane distortions of the heme were analysed using an online normal coordinate structural decomposition script

(<http://mliptak.w3.uvm.edu/nsd.html>) derived from the normal-coordinate structural decomposition procedure.<sup>34,56</sup>

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

Synchrotron data were measured on beamlines I03 and I04-1 at Diamond Light Source under proposal MX7461. We acknowledge use of the JCSG Quality Control Server. CRA acknowledges funding from the National Science Foundation (grant MCB 1411963). TMC was supported by Leverhulme Trust award RPG-2014-355 to MAH. We thank Professor Pierre Moënne-Loccoz (Oregon Health & Science University) for helpful discussions and for assistance with resonance Raman measurements. Professor Michael Wilson and Dr Corinne Whitby of the University of Essex are thanked for insightful discussions of the manuscript.

## References

- 1 L. Nazaries, J. C. Murrell, P. Millard, L. Baggs and B. K. Singh, *Environ. Microbiol.*, 2013, **15**, 2395–2417.
- 2 N. Lehnert, T. C. Berto, M. G. I. Galinato and L. E. Goodrich, in *Handbook of porphyrin science*, ed. K. M. Kadish, K. M. Smith and R. Guilard, World Scientific, 2011, vol. 14, pp. 1–247.
- 3 Y. A. Trotsenko and J. C. Murrell, *Adv. Appl. Microbiol.*, 2008, **63**, 183–229.
- 4 D. J. Bergmann, J. A. Zahn and A. A. DiSpirito, *Arch. Microbiol.*, 2000, **173**, 29–34.
- 5 B. O. Elmore, D. J. Bergmann, M. G. Klotz and A. B. Hooper, *FEBS Lett.*, 2007, **581**, 911–916.
- 6 J. A. Zahn, D. M. Arciero, A. B. Hooper and A. A. Dispirito, *Eur. J. Biochem.*, 1996, **240**, 684–691.
- 7 J. D. Caranto and K. M. Lancaster, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 8217–8222.
- 8 L. Y. Stein, *Methods Enzymol.*, 2011, **486**, 131–152.
- 9 N. Igarashi, H. Moriyama, T. Fujiwara, Y. Fukumori and N. Tanaka, *Nat. Struct. Biol.*, 1997, **4**, 276–284.
- 10 P. Cedervall, A. B. Hooper and C. M. Wilmot, *Biochemistry*, 2013, **52**, 6211–6218.
- 11 P. E. Cedervall, A. B. Hooper and C. M. Wilmot, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2009, **65**, 1296–1298.
- 12 A. R. Pearson, B. O. Elmore, C. Yang, J. D. Ferrara, A. B. Hooper and C. M. Wilmot, *Biochemistry*, 2007, **46**, 8340–8349.
- 13 M. A. Smith and K. M. Lancaster, *Biochemistry*, 2018, **57**, 334–343.
- 14 J. D. Caranto, A. C. Vilbert and K. M. Lancaster, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 14704–14709.
- 15 A. C. Vilbert, J. D. Caranto and K. M. Lancaster, *Chem. Sci.*, 2018, **9**, 368–379.



16 M. A. Hough and C. R. Andrew, *Adv. Microb. Physiol.*, 2015, **67**, 1–84.

17 G. R. Moore and G. W. Pettigrew, *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*, Springer-Verlag, Berlin, 1990.

18 S. Antonyuk, N. Rustage, C. A. Petersen, J. L. Arnst, D. J. Heyes, R. Sharma, N. Berry, N. S. Scrutton, R. R. Eady, C. R. Andrew and S. S. Hasnain, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 15780–15785.

19 E. M. Garton, D. A. Pixton, C. A. Petersen, R. R. Eady, S. S. Hasnain and C. R. Andrew, *J. Am. Chem. Soc.*, 2012, **134**, 1461–1463.

20 D. Kekilli, C. A. Petersen, D. A. Pixton, D. D. Ghafoor, G. H. Abdullah, F. S. N. Dworkowski, M. T. Wilson, D. J. Heyes, S. J. O. Hardman, L. M. Murphy, R. W. Strange, N. S. Scrutton, C. R. Andrew and M. A. Hough, *Chem. Sci.*, 2017, **8**, 1986–1994.

21 E. Martin, V. Berka, I. Sharina and A. L. Tsai, *Biochemistry*, 2012, **51**, 2737–2746.

22 R. P. Ambler, R. G. Bartsch, M. Daniel, M. D. Kamen, L. McLellan, T. E. Meyer and J. Van Beeumen, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 6854–6857.

23 A. T. Poret-Peterson, J. E. Graham, J. Gulleedge and M. G. Klotz, *ISME J.*, 2008, **2**, 1213–1220.

24 M. A. Campbell, G. Nyerges, J. A. Kozlowski, A. T. Poret-Peterson, L. Y. Stein and M. G. Klotz, *FEMS Microbiol. Lett.*, 2011, **322**, 82–89.

25 D. J. Bergmann and A. B. Hooper, *Eur. J. Biochem.*, 2003, **270**, 1935–1941.

26 K. K. Andersson, G. T. Babcock and A. B. Hooper, *Biochem. Biophys. Res. Commun.*, 1991, **174**, 358–363.

27 M. Numata, T. Saito, T. Yamazaki, Y. Fukumori and T. Yamanaka, *J. Biochem.*, 1990, **108**, 1016–1021.

28 M. A. Hough and C. R. Andrew, in *Advances in Microbial Physiology. Recent advances in microbial oxygen-binding proteins*, ed. R. K. Poole, Elsevier, UK, 2015, vol. 67, ch. 1, pp. 1–84.

29 J. A. Zahn, C. Duncan and A. A. DiSpirito, *J. Bacteriol.*, 1994, **176**, 5879–5887.

30 C. R. Andrew, E. L. Green, D. M. Lawson and R. R. Eady, *Biochemistry*, 2001, **40**, 4115–4122.

31 S. Fujii, H. Oki, K. Kawahara, D. Yamane, M. Yamanaka, T. Maruno, Y. Kobayashi, M. Masanari, S. Wakai, H. Nishihara, T. Ohkubo and Y. Sambongi, *Protein Sci.*, 2017, **26**, 737–748.

32 D. Yamane-Koshizawa, S. Fujii, T. Maruno, Y. Kobayashi, M. Yamanaka, S. Wakai and Y. Sambongi, *Biosci., Biotechnol., Biochem.*, 2018, **82**, 304–311.

33 J. Lund, M. P. Woodland and H. Dalton, *Eur. J. Biochem.*, 1985, **147**, 297–305.

34 A. B. Graves, M. T. Graves and M. D. Liptak, *J. Phys. Chem. B*, 2016, **120**, 3844–3853.

35 M. D. Liptak, X. Wen and K. L. Bren, *J. Am. Chem. Soc.*, 2010, **132**, 9753–9763.

36 J. G. Kleingardner, S. E. Bowman and K. L. Bren, *Inorg. Chem.*, 2013, **52**, 12933–12946.

37 M. L. Fernandez, D. A. Estrin and S. E. Bari, *J. Inorg. Biochem.*, 2008, **102**, 1523–1530.

38 L. O. Michel, M. Sandkvist and M. Bagdasarian, *Gene*, 1995, **152**, 41–45.

39 E. Arslan, H. Schilz, R. Zufferey, P. Künzler and L. Thöny-Meyer, *Biochem. Biophys. Res. Commun.*, 1998, **251**, 744–747.

40 S. Barbieri, L. M. Murphy, R. G. Sawers, R. R. Eady and S. S. Hasnain, *J. Biol. Inorg. Chem.*, 2008, **13**, 531–540.

41 W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125–132.

42 P. R. Evans and G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2013, **69**, 1204–1214.

43 K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 1002–1011.

44 G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 355–367.

45 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 486–501.

46 I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375–W383.

47 S. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss and P. A. Tucker, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2005, **61**, 449–457.

48 S. X. Cohen, M. Ben Jelloul, F. Long, A. Vagin, P. Knipscheer, J. Lebbink, T. K. Sixma, V. S. Lamzin, G. N. Murshudov and A. Perrakis, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2008, **64**, 49–60.

49 A. Micsonai, F. Wien, L. Kurnya, Y. H. Lee, Y. Goto, M. Refregiers and J. Kardos, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E3095–E3103.

50 A. Micsonai, F. Wien, E. Bulyaki, J. Kun, E. Moussong, Y. H. Lee, Y. Goto, M. Refregiers and J. Kardos, *Nucleic Acids Res.*, 2018, **46**, W315–W322.

51 Uchiyama, *et al.*, *J. Am. Chem. Soc.*, 2004, **126**, 14684–14685.

52 S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403–410.

53 M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins, *Bioinformatics*, 2007, **23**, 2947–2948.

54 A. Stamatakis, *Bioinformatics*, 2014, **30**, 1312–1313.

55 S. Kumar, G. Stecher and K. Tamura, *Mol. Biol. Evol.*, 2016, **33**, 1870–1874.

56 W. Jentzen, X. Z. Song and J. A. Shelnutt, *J. Phys. Chem. B*, 1997, **101**, 1684–1699.

