Structural flexibility of the heme cavity in the cold-adapted truncated hemoglobin from the Antarctic marine bacterium *Pseudoalteromonas haloplanktis* TAC125

Daniela Giordano[1]*, Alessandra Pesce[2]*, Leonardo Boechi[3], Juan Pablo Bustamante[3], Elena Caldelli[4], Barry D. Howes[4], Alessia Riccio[1], Guido di Prisco[1], Marco Nardini[5], Dario Estrin[3], Giulietta Smulevich[4,§], Martino Bolognesi[5,6,§], Cinzia Verde[1,7,§]

[1]Institute of Biosciences and BioResources, National Research Council, Via Pietro Castellino 111, I-80131 Napoli, Italy; [2]Department of Physics, University of Genova, Via Dodecaneso 33, I-16146 Genova, Italy; [3]Departamento de Química Inorgánica, Analítica y Química Física, Universidad de Buenos Aires, Ciudad Universitaria, C1428EGA, Argentina; [4]Department of Chemistry "Ugo Schiff", University of Firenze, Via della Lastruccia 3-13, I-50019 Sesto Fiorentino (Fi), Italy; [5]Department of Biosciences, University of Milano, Via Celoria, 26, I-20133 Milano, Italy; [6]CNR-Institute of Biophysics and CIMAINA, University of Milano, I-20133 Milano, Italy; [7]Roma 3 University, Department of Biology, Viale Marconi 446, I-00146 Roma, Italy.

*These authors contributed equally to this work

[§]Corresponding authors:
G. Smulevich, Department of Chemistry "Ugo Schiff", University of Firenze, Via della Lastruccia 3-13, 50019 Sesto Fiorentino (Fi), Italy; Tel. +39 0554573083; Fax +39 0554573077; e-mail: giulietta.smulevich@unifi.it;
M. Bolognesi, Department of Biosciences, University of Milano, Via Celoria, 26, 20133 Milano (Italy); Tel. +39 0250314893; Fax +39 0250314895; e-mail: martino.bolognesi@unimi.it
C. Verde, Institute of Biosciences and BioResources, National Research Council, Via Pietro Castellino 111, 80131 Napoli, Italy; Tel./Fax +39 0816132710; e-mail: cinzia.verde@ibbr.cnr.it.

Running title: Heme-cavity flexibility in an Antarctic bacterial hemoglobin

Abbreviations

Truncated hemoglobin, TrHb; TrHbII from *Pseudoalteromonas haloplanktis* TAC125, *Ph*-2/2HbO; myoglobin, Mb; flavohemoglobins FHb; single-domain globin, SDHb; TrHbII from *Mycobacterium tuberculosis*, *Mt*-2/2HbO; TrHbII from *Thermobifida fusca*, *Tf*-2/2HbO; TrHbII from *Bacillus subtilis*, *Bs*-2/2HbO; TrHbII from *Geobacillus stearothermophilus*, *Gs*-2/2HbO; TrHbII from *Agrobacterium tumefaciens*, *At*-2/2HbO; TrHbII from *Arabidopsis thaliana*, *Ath*-2/2HbO; *Pseudoalteromonas haloplanktis* TAC125, *Ph*TAC125; Resonance Raman spectroscopy, RR; Electron Paramagnetic Resonance spectroscopy, EPR; Molecular

Dynamics, MD; Quantum Mechanics, QM; Density Functional Theory, DFT; Perdew-Burke-Ernzerhof, PBE; High-Spin, HS; Low-Spin, LS; Root Mean Square Deviation, RMSD; Implicit Ligand Sampling, ILS; hexa-coordinated heme state, 6c; penta-coordinated heme state, 5c; Horseradish Peroxidase Isoenzyme C, HRPC.

Database

Structural data have been submitted to the Protein Data Bank under accession numbers 4UUR and R4UURSF.

Keywords

Adaptation, Bacterial hemoglobin, Resonance Raman, X-ray structure, Molecular dynamics

Abstract

Truncated hemoglobins build one of the three branches of the globin protein superfamily. They display a characteristic two-on-two α-helical sandwich fold and are clustered into three groups (I, II, and III) based on distinct structural features. Truncated hemoglobins are present in eubacteria, cyanobacteria, protozoa and plants. Here we present a structural, spectroscopic and molecular dynamics characterization of a group-II truncated hemoglobin, encoded by the *PSHAa0030* gene from *Pseudoalteromonas haloplanktis* TAC125 (*Ph*-2/2HbO), a cold-adapted Antarctic marine bacterium hosting one flavohemoglobin and three distinct truncated hemoglobins. *Ph*-2/2HbO aquo-met crystal structure (at 2.21 Å resolution) shows typical features of group-II truncated hemoglobins, namely the two-on-two α-helical sandwich fold, a helix Φ preceding the proximal helix F, and a heme distal-site hydrogen bonded network that includes water molecules and several distal-site residues, **including His(58)CD1**. Analysis of *Ph*-2/2HbO by electron paramagnetic resonance, resonance Raman and electronic absorption spectra, under varied solution conditions, shows that *Ph*-2/2HbO can access diverse heme ligation states. Among these, detection of a low-spin heme hexa-coordinated species suggests that residue Tyr(42)B10 can undergo large conformational changes in order to act as the sixth heme-Fe ligand. Altogether, the results show that *Ph*-2/2HbO maintains the general structural features of group-II truncated hemoglobins, but displays enhanced conformational flexibility in the proximity of the heme cavity, a property likely related to the functional challenges, such as low temperature, high O_2 concentration and low kinetic energy of molecules, experienced by organisms living in the Antarctic environment.

Introduction

The most recent bioinformatics search for globin-like sequences identified putative globins in more than one half of the 2200 sequenced bacterial genomes. A comprehensive nomenclature, including prokaryotic and eukaryotic globins, assigns globins to three distinct families: (*i*) the myoglobin (Mb)-like family displaying the classical three-on-three (3/3) α-helical sandwich fold, and including flavohemoglobins (FHbs) and single-domain globins (SDHbs); (*ii*) the sensor globin family; and (*iii*) the truncated hemoglobin (TrHb) family, that displays the two-on-two (2/2) α-helical sandwich fold [1]. Members of the TrHb family have been found in eubacteria, cyanobacteria, protozoa and plants, but not in animals [2]. On the basis of phylogenetic analyses, the TrHb family can be further divided into three distinct sub-families: TrHbI (or N), TrHbII (or O), and TrHbIII (or P); specific structural features related to sequence motifs, heme cavity, helices in their folds, and protein matrix tunnel(s) distinguish each group [2,3].

Several sequence and 3D structure analyses of the TrHbII group have been reported. To date, six TrHbII crystal structures have been described, of *Mycobacterium tuberculosis* (*Mt*-2/2HbO, **PDB ID: 1IDR,** [4]), *Thermobifida fusca* (*Tf*-2/2HbO, **PDB ID: 2BMM,** [5]), *Bacillus subtilis* (*Bs*-2/2HbO, **PDB ID: 1UX8,** [6]), *Geobacillus stearothermophilus* (*Gs*-2/2HbO, **PDB ID: 2BKM,** [7]), *Agrobacterium tumefaciens* (*At*-2/2HbO, **PDB ID: 2XYK,** [8]), and of the first plant TrHb from *Arabidopsis thaliana* (*Ath*-2/2HbO, also called AtGLB3 or AHb3, **PDB ID: 4C0N**, [9]). Besides sequence signatures that translate into tertiary structure features, the most distinctive feature of TrHbIIs is the presence of specific heme distal residues at the B10, CD1, E7, E11, and G8 topological sites, **which are implicated in modulation** of heme ligand binding [2,3] .

TrHbs have also been identified in the cold-adapted Antarctic marine bacterium *Pseudoalteromonas haloplanktis* TAC125 (*Ph*TAC125): one TrHbI (encoded by the *PSHAa0458* gene), and two distinct TrHbIIs (encoded by *PSHAa0030* and *PSHAa2217*) [10].

Amino-acid sequence identity between the two TrHbIIs is only 24%, suggesting that these proteins may play different function(s) in the bacterial metabolism. Moreover, *Ph*TAC125 hosts one FHb, annotated as *PSHAa2880* [10]. To our knowledge, *Ph*TAC125 provides the first example of coexistence of genes encoding one FHb and three TrHbs [11]. The presence of multiple globin genes supports the idea of their involvement in a protection mechanism against oxidative and nitrosative stress in a cold and O_2-rich environment such as Antarctica [12].

In particular, TrHbII encoded by the *PSHAa0030* gene, named *Ph*-2/2HbO, has been thoroughly investigated from the biochemical and functional viewpoints [10,12-16]. The inactivation of the *PSHAa0030* gene encoding *Ph*-2/2HbO makes the mutant bacterial strain sensitive to high O_2 levels, hydrogen peroxide, and nitrosating agents [12]. In contrast, overexpression of the *PSHAa0030* gene in a mutant *Escherichia coli* defective of FHb and hypersensitive to nitrosative stress indicates that *Ph*-2/2HbO protects growth and cellular respiration of the heterologous host against the toxic effect of NO-donors. Moreover, the ferric form of *Ph*-2/2HbO was shown to catalyse peroxynitrite isomerisation *in vitro*, confirming its potential role in reactive nitrogen species scavenging [16].

Here we report a comprehensive study of *Ph*-2/2HbO, combining its crystal structure (solved at 2.21 Å resolution) with a spectroscopic analysis based on Resonance Raman (RR) and Electron Paramagnetic Resonance (EPR), and complemented by molecular dynamic (MD) simulations. The RR solution data highlight spin population differences between the solution and the crystalline states, suggesting a flexible architecture of the protein structure in proximity of the heme cavity, not previously observed in other TrHbs. The structural and dynamic features relevant for determining the ligand-binding properties and flexibility of *Ph*-2/2HbO are considered.

Results

Structure of Ph-2/2HbO

The 3D structure of aquo-met *Ph*-2/2HbO (devoid of the first nineteen residues, see Materials and Methods) was solved using crystals belonging to the orthorhombic space group $P2_12_12_1$, hosting two protein chains (A and B) per asymmetric unit (**PDB ID: 4UUR** and **R4UURSF**). The structure was refined at 2.21 Å resolution to a final R-factor and R-free values of 18.3% and 24.6%, respectively, with ideal stereochemical parameters (Table 1). Conformational disorder affects the N-terminal 20-21 residues of both *Ph*-2/2HbO chains, whose electron density was lost to the solvent.

Superposition of 124 C_α atoms of the two independent chains yields a Root Mean Square Deviation (RMSD) of 0.38 Å. The results discussed below apply equally to both the A and B chains unless otherwise stated. Analysis of the *Ph*-2/2HbO dimeric assembly, by means of the program PISA (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver) [18], did not highlight a significant association interface between the A and B chains in the crystal asymmetric unit that could justify a stable quaternary assembly (Fig. S1). This finding is in agreement with the results of size-exclusion chromatography experiments that identified the protein as a monomer in solution [10].

Structural comparison with other TrHbIIs

The *Ph*-2/2HbO tertiary structure displays the typical features of the α-helical 2/2Hb-fold [2,3,19], with some structural modifications that are specific for TrHbIIs, such as an additional α-helix (Φ helix, residues 82-87 in *Ph*-2/2HbO) located between helices E and F (Figs. 1 and 2) [4]. Structural superposition of *Ph*-2/2HbO on *At*-2/2HbO, *Mt*-2/2HbO, *Tf*-2/2HbO, *Bs*-2/2HbO, *Gs*-2/2HbO and *Ath*-2/2HbO yields RMSD values in the 0.9-1.2 Å range, the best match being with *At*-2/2HbO. The main differences relative to other TrHbIIs (taking *Mt*-2/2HbO as reference, see alignment in Fig. 2) are located in the N-terminal region (where full-length *Ph*-2/2HbO hosts a sequence extension of nineteen residues, proteolytically cleaved during protein purification), in the BC (three-residue insertion in *Ph*-2/2HbO), CE (one-residue

deletion in *Ph*-2/2HbO) and GH (three-residue deletion in *Ph*-2/2HbO) loops, and in the C-terminal region. **The two Gly-Gly motifs present at the C-terminal end of helices A and E, typical of TrHbI and TrHbII, are conserved in *Ph*-2/2HbO*;* here, in particular, they help to stabilize the short A helix in a conformation locked onto helices B and E** (Fig. 1).

Analysis of the crystal structure reveals that in *Ph*-2/2HbO the CD1 topological position is occupied by His58, not by Tyr55 as previously proposed [13]. CD1 His is also present in *At*-2/2HbO, whereas in TrHbs this site typically hosts Phe, with the exception of TrHbIIs from *M. tuberculosis*, *M. avium*, *M. leprae*, *M. smegmatis*, *Streptomyces coelicolor*, *Corynebacterium diphtheriae* and *T. fusca*, which host Tyr [5,20]. **Bustamante and co-workers (unpublished results) have recently demonstrated that, by comparing all the TrHb sequences available to date (~ 1100), the residues found at position CD1 are: His (~ 20 %), Phe (~ 63 %) and Tyr (~ 15 %).**

A medium polarity distal cavity of about 50 $Å^3$ volume is defined by the porphyrin ring and by residues Phe(41)B9, Tyr(42)B10, His(58)CD1, Ile(65)E7, Phe(69)E11 and Trp(109)G8 (Fig. 3); the cavity hosts two water molecules that form a H-bonded network with the heme distal residues (see below). Analysis of the whole *Ph*-2/2HbO crystal structure does not highlight a proper protein matrix tunnel(s) connecting the heme distal cavity to the solvent, such as those observed in *Mt*-TrHbI [20,21]. This finding is in keeping with the absence of a protein-matrix tunnel previously reported for TrHbIIs. TrHbIIs also tend to host a small residue at E7 (typically Ala, Ser, or Thr), thus suggesting the an E7 route entry path to facilitate the accessibility of diatomic ligands to the heme distal site [2,3,19]. In E7, the residue in *Ph*-2/2HbO is Ile(65), whose side chain separates the heme distal cavity from the solvent region (Fig. 3A). An alternative access to the heme distal cavity might be located between the B, C and E helices, however the bulky side chains of Tyr(42)B10, Tyr(55)C5, His(58)CD1 and Leu(62)E2 appear to hinder such an entry site in the *Ph*-2/2HbO crystal structure (Fig. 3A).

Heme pocket

Both *Ph*-2/2HbO A and B chains display the heme porphyrin ring rotated by 180° around the methinic α–γ meso axis, relative to the common orientation found in (non)vertebrate globins. In *Ph*-2/2HbO, the bound heme is stabilized through direct Fe-coordination to proximal His(96)F8, electrostatic interactions with the heme propionates, and van der Waals contacts (< 4.0 Å) with 23 residues. In particular, in both chains, propionate-D is stabilized by a H-bonded salt bridge with Arg(95)F7, and propionate-A is electrostatically coupled to Arg64 slightly differently in the two chains (Fig. 3B); moreover, a water molecule provides a H-bond bridge between propionates A and D in the B chain. It is worth noting that in *Ph*-2/2HbO site 83 (within helix Φ), which generally hosts a highly conserved Tyr in TrHbIIs, is occupied by Phe (Fig. 2), thus impairing the formation of an additional H bond with propionate D. **The residues found in the Φ helix region, particularly in position 83 (*Ph*-2/2HbO numbering) are: Phe (~ 13 %) and Tyr (~ 84 %) (Bustamante et al., unpublished results).**

In the proximal site, Leu92, Phe99, Ile101 and Met141 are mainly involved in van der Waals contacts with the porphyrin ring. Proximal His(96)F8 is coordinated to the heme iron atom (coordination-bond length of 2.2 Å in chain A and 2.1 Å in chain B), with the F8 imidazole ring lying in a staggered azimuthal orientation relative to the heme pyrrole N atoms. Accordingly, the ferrous form showed a ν(Fe–His) stretching mode at 222 cm^{-1} [14], higher than that of human HbA (214 cm^{-1}), but similar to that of other TrHbs characterized by a staggered orientation of the proximal His imidazole ring [22]. In fact, the staggered orientation, in contrast to the eclipsed orientation observed in HbA [22], is expected to increase the iron–His bond strength.

In the heme distal pocket, Phe(41)B9, Tyr(42)B10, His(58)CD1, Ile(65)E7, Phe(69)E11 and Trp(109)G8 surround the heme-ligated water molecule (Fig. 3B). The His(58)CD1 imidazole ring is parallel to the heme, its orientation being set by a H-bond (2.9 Å) between the ND1 atom and the **carbonyl O** atom of Leu(54)C4. At the rear end of the distal pocket,

Trp(109)G8, conserved in TrHbIIs and TrHbIIIs, fills the inner part of the heme distal site, preventing further diffusion of ligands away from the heme-Fe coordination site. The Trp(109)G8 indole ring is parallel and in contact with the porphyrin ring at the B and C pyrroles (Fig. 3B).

The heme distal site is characterized by a highly intertwined H-bonded network, involving Tyr(42)B10, His(58)CD1, Trp(109)G8, and two water molecules (W13 and W32 in the A chain, W16 and W23 in the B chain; Fig. 3B). In particular, the heme-coordinated water molecules (W32 in A and W16 in B chains, with Fe-O 2.2 Å coordination bonds), are H bonded to a second distal-site water molecule (W13 in the A chain and W23 in the B chain; 2.6 Å and 2.5 Å distances, respectively), and to the indole NE1 atom of Trp(109)G8 (3.0-Å and 3.1 Å distances, respectively). The W13 and W23 water molecules are fully buried in the distal site, and are in turn H bonded to the Tyr(42)B10 hydroxyl group (2.8 Å and 2.7 Å, respectively) and to the NE2 atom of His(58)CD1 (2.5 Å and 2.6 Å, respectively).

Spectroscopic measurements in solution

The electronic absorption spectrum of ferric *Ph*-2/2HbO and its second derivative spectrum at pH 7.6 (Fig. 4A), characterized by a Soret band at 408 nm, Q bands at 503, 541 and 570 nm, and a CT1 band at 635 nm, suggests a mixture of high-spin (HS) and low-spin (LS) forms. This is confirmed by the corresponding high-frequency RR spectrum (Fig. 4D), which clearly indicates a mixture of two forms; a predominant aquo hexa-coordinated HS state (6cHS; ν_3 1480, ν_2 1558, ν_{10} 1608 cm^{-1}) and a less populated hexa-coordinated LS state (6cLS; ν_3 1503 cm^{-1}). These spectra are different from those previously reported for ferric *Ph*-2/2HbO at pH 7.6 [13] where, likely due to different preparation procedures, multiple LS forms were observed (see Materials and Methods). However, as reported in that case, the absorption maxima of the LS form are quite unusual, and reminiscent of those of ferric *Chlamydomonas* chloroplast Hb [23] and of the hemophore HasA proteins from *Serratia marcescens* and *Pseudomonas*

aeruginosa [24,25]. Nevertheless, they are very different from either a LS His–Fe–His (that exhibits well-defined absorption bands at about 535 and 565 nm), or a His–Fe–OH heme complex as observed at alkaline pH (see below) [26].

As suggested by the crystal structure, two distal-site residues may be involved in generating the 6cLS state, i.e. His(58)CD1 and Tyr(42)B10 (Fig. 3B). In order to check whether a Tyr might be coordinated to the heme-Fe atom, RR excitation in the visible region was performed since the characteristic vibrational frequencies of bound phenolate can often be identified by excitation in the tyrosinate–Fe(III) charge transfer band (near 500 nm) [27]. As shown in Fig. 5, upon excitation in the visible region at pH 7.6, three bands at 590, 1312 and 1509 cm^{-1} (which are polarized, data not shown) are enhanced, becoming particularly evident upon excitation at 514.5 nm. These bands are assigned to the ν(Fe-O_{Tyr}), ν_{Tyr}(C-O), and ν_{Tyr}(C=C) tyrosinate modes, respectively, which are not present in the spectrum upon excitation at 413.1 nm. The band around 590 cm^{-1}, due to the stretching mode of the covalent bond between the heme-Fe atom and the Tyr phenolic O, ν(Fe–O_{Tyr}), is of special interest, since its frequency is highly sensitive to the Tyr protonation state. Table 2 reports the frequencies of the phenolate modes observed in a number of heme-Fe tyrosinate complexes in different proteins, for comparison with those presented here; such frequencies differ slightly from those reported in a previous study on *Ph*-2/2HbO [13]. Based on previously reported UV and visible RR spectra (Table 2, [29]), the frequency at 590 cm^{-1} observed in *Ph*-2/2HbO is consistent with a deprotonated Tyr.

In agreement with the electronic absorption and RR spectra, the *Ph*-2/2HbO X-band EPR spectrum at pH 7.6 displays an axial HS signal ($g_{\perp}$~6.0, $g_{//}$~ 2.0) and only one rhombic LS form (g = 2.95, 2.26, 1.49) (Fig. 4B). The LS g values differ from those normally observed for His–Fe–Tyr axial coordination, where the tyrosinate is not H-bonded, but also from the values reported for His-Fe-His and His-Fe-OH LS signals (Table S1). However, they are similar to those of the human serum albumin–ibuprofen complex (2.93, 2.27, 1.55) [32] and the HasA

proteins from *S. marcescens* and *P. aeruginosa* (g = 2.86, 2.21, 1.71) [24,25]. In both cases, a nearby His serves as H-bond acceptor to stabilize the phenolate ligand. Hence, by analogy with HasA, and in agreement with the electronic absorption and RR spectra, the *Ph*-2/2HbO LS form is proposed to host His–Fe–Tyr axial coordination, with a deprotonated Tyr strongly H-bonded to a nearby residue that stabilizes the phenolate. **This assignment is discussed within the context of the Blumberg–Peisach diagrams in the supplementary materials.**

Based on sequence alignments (Fig. 2) and the location of heme distal residues resulting from the crystal structure (Fig. 3B), Tyr(42)B10 is proposed as the sixth heme-Fe ligand in the *Ph*-2/2HbO LS form. Notably, analysis of the aquo-met *Ph*-2/2HbO crystal structure (which displays a heme-coordinated water molecule) does not allow identification of a distal residue acting as the sixth heme-Fe ligand, as observed in the protein 6cLS form in solution. In this respect, attempts to produce a *Ph*-2/2HbO variant mutated at the B10 site (replacing Tyr42 with Phe) yielded a purified greenish globin, likely due to the presence of heme-d [33], which compromised the spectroscopic analyses. Attempts to generate the apoprotein followed by reconstitution *in vitro* with hemin proved unsuccessful in this mutant, but also in a mutant bearing the Trp(109)G8→Phe mutation.

The RR bands assigned to tyrosinate modes disappear at alkaline pH (Fig. 5), as at this pH (>10) the protein is characterized by a hydroxyl ion bound to the heme-Fe atom. The absorption spectrum is characteristic of a mixture of 6cHS and 6cLS forms (Soret band at 414 nm, Q bands at 545 and 587 nm, and a weak CT1 band at 610 nm, Fig. 4A), consistent with the observed high-frequency RR spectrum that displays a predominant 6cLS form (ν_3 1501, ν_2 1579, ν_{37} 1604, ν_{10} 1638 cm^{-1}) and a weak 6cHS form (ν_3 1477 cm^{-1}). It is worth noting that the high-frequency RR spectrum also reveals the presence of a 5cHS form at alkaline pH (ν_3 1491 cm^{-1}) (Fig. 4D) (see below). At alkaline pH the relative intensities of the γ_6 (at 336 cm^{-1}), propionyl (368 and 380 cm^{-1}) and vinyl bending (410 and 422 cm^{-1}) modes are significantly reduced as compared to pH 7.6 (Fig. 4C), suggesting that at pH > 10 the H-bonding interactions

with the propionyl groups (observed in the X-ray structure) are weakened [34]. Figs. S2-S5 and Tables S2-S5 in the supplemental materials report the curve fitting analysis of the RR spectra shown in Fig. 5 and the complete vibrational assignment.

In agreement with the RR data, in the EPR spectrum at alkaline pH the LS signal corresponding to His–Fe–Tyr axial coordination is almost absent (g = 2.95, 2.26, 1.49), being replaced by a new LS signal (g = 2.65, 2.19, 1.82) (Fig. 4B), typical of hydroxyl heme LS forms (Table S1, [35]).

Fig. 6 compares the low-frequency RR spectra of *Ph*-2/2HbO at alkaline pH in H_2O, D_2O and $H_2{}^{18}O$ buffered solutions, at 12 K (Fig. 6A-B) and 298 K (Fig. 6C-D). The 6cHS and 6cLS forms exist in a thermal spin-state equilibrium at room temperature, being converted to almost pure 6cLS heme at 12 K with a small population of 5cHS still present (Fig. S6). At 12 K the resolution of the bands is higher, enabling the weak ν(Fe-OH) bands to be more easily identified through their sensitivity to ^{2}H and ^{18}O isotopic substitution. The two isotope-sensitive bands observed at 528 and 568 cm^{-1} (at 12 K), (525 and 570 cm^{-1} at 298 K) are assigned to two ν(Fe-OH) stretching modes. The band at 528 cm^{-1} is assigned to a His-Fe-OH^- 6cLS form. Its frequency, about 20-cm^{-1} lower than that observed in met-Mb and met-Hb [36], is indicative of strong H bonds between the OH^- ligand and distal residues. In fact, with an increase of the H-bond strength, a decrease of the force constant of the Fe-OH bond, with concomitant decrease of the ν(Fe-OH) stretching frequency, is expected. Upon isotopic substitution in D_2O buffer, this band up-shifts by 6 cm^{-1} rather than displaying the expected down-shift due to the increased mass of the Fe-OD oscillator relative to Fe-OH. A similar behaviour has been previously observed in horseradish peroxidase (HRPC) [36] (Table 3), where distal Arg38 appears essential for stabilization of the OH^- ligand and distal His(42)E7 acts as a H-bond acceptor [43]. Moreover, similar to HRPC [36], upon isotopic substitution in D_2O at room temperature (Fig. S7), a decrease in intensity and a small frequency change are observed for the bending mode of the propionyl group at 367 cm^{-1}, indicating an alteration in

the H-bonded network induced by the different strengths of H- and deuterium-bonds interacting with the heme ligand. The corresponding band of the 6cHS (observed only at room temperature) has not been identified.

The band at 568 cm^{-1} down-shifts to 554 and 549 cm^{-1} in D_2O and $H_2{}^{18}O$, respectively. It is noteworthy that the frequency of this band is much higher than those typical of 6cLS and 6cHS Fe-OH forms. In analogy with the His93Gly Mb mutant (ν(Fe-OH) at 575 cm^{-1}) and *Scapharca inaequivalvis* Hb [41,42], it is assigned to a 5cHS Fe-OH form, uncommon in heme proteins. Such an assignment implies that the proximal His-Fe bond must be ruptured at alkaline pH. **The alkaline transition is fully reversible, since upon returning to pH 7.6, the initial spectra of a typically His-Fe containing protein are obtained, thus ruling out irreversible denaturation**. Table 3 lists the ν(Fe-OH) frequencies observed in a number of reference heme proteins at alkaline pH. The 6cLS form, whose frequency is similar to that observed in *Mt-2/2*HbO [37], is characterized by strong H bonds between OH^- and distal residues. However, similar to TrHb-I of the cyanobacterium *Synechocystis* [40], at alkaline pH *Ph*-2/2HbO shows the presence also of a 5cHS characterized by the rupture of the Fe-His proximal ligand.

Fig. 7 compares the RR spectra as a function of pH between neutral and alkaline pH, including the intermediate pH 8.5 value, which corresponds to the pH used in the crystal-structure analysis. While the electronic absorption spectrum at pH 8.5 is almost identical (data not shown) to that at pH 7.6 (Fig. 4A), at increasing pH a general reduction in the low-frequency RR spectrum intensity (normalized on the ν_7 mode (678 cm^{-1})) takes place (Fig. 7A). The most evident changes are the marked decrease in relative intensity of the bending modes in the 330-430 cm^{-1} region and, in the high frequency region, the up-shift of the ν_4 mode by 1 cm^{-1} from pH 7.6 to 8.5 (Fig. 7B). Therefore, we can conclude that at pH 8.5 the spectrum is intermediate between those at pH 7.6 and at pH 10.9, thus displaying a small Fe-OH LS population. This result is in agreement with the results from the X-ray structure, where a heme-

Fe coordinated water molecule (W13 in chain A, and W23 in chain B) is involved in a distal H-bonded network, suggesting that this molecule may assume a degree of OH^- 6cLS character.

Molecular Dynamic Simulations

MD simulations of ferrous *Ph*-2/2HbO, in the absence of any heme distal ligand, show that water molecules can access the heme distal cavity and be stabilized through a H-bond network based on the distal residues Tyr(42)B10, His(58)CD1 and Trp(109)G8 (Fig. S8), in agreement with the X-ray crystal structure (Fig. 3B). Besides mapping the water dynamics in the heme distal site, during MD simulation it was possible to analyze the dynamic behaviour of the residues surrounding the heme. In particular, we focused our attention on Tyr(42)B10 that, on the basis of the RR experiments, is suggested to be coordinated to the heme-Fe atom. In our MD simulation, Tyr(42)B10 is too far from the iron to be considered a candidate to occupy the sixth coordination site. This is similar to other well studied members of the TrHbII family, such as *Bs*-2/2HbO*, Tf*-2/2HbO and *Mt*-2/2HbO [44-46], where no endogenous hexa-coordination was observed (Fig. S9). As the currently accessible timescales of MD simulations inherently limit the extent of conformational space that can be sampled, an alternative conformation, with Tyr(42)B10 coordinated to the heme-Fe atom cannot be ruled out at longer simulation times. A large structural rearrangement would however be required to attain a conformation compatible with endogenous heme hexa-coordination based on Tyr(42)B10.

MD simulations of the *Ph*-2/2HbO CO (Fig. 8A) and O_2 (Fig. 8B) complexes show that both ligands are stabilized by a strong H bond to Trp(109)G8, and a weaker one to Tyr(42)B10 (Figs. 8C-D) with the interaction between O_2 and Trp(109)G8 (Fig. 8D) more stable than that between CO and Trp(109)G8 (Fig. 8C). The important role of Trp(109)G8 in ligand stabilization has been previously observed in other TrHbIIs, e.g. *Bs*-2/2HbO [45,47], *Tf*-2/2HbO [46], and *Mt*-2/2HbO [37,44]. No significant interactions of the CO ligand with His(58)CD1 are apparent from the simulations (data not shown).

A detailed analysis of possible ligand migration pathways and associated tunnel/cavities within the protein matrix was performed using the Implicit Ligand Sampling (ILS) method. A comparison between cavities identified by the X-ray and the MD structures is provided in Fig. 9. Even though no widely open tunnels were identified, a possible pathway for ligand migration at the E7 gate was shown by MD simulations (Fig. 9). In particular, two on-pathway docking sites were found for the CO ligand: the first cavity is located over the heme-Fe atom $(CO)_1$, as suggested by the crystal structure, while the second, $(CO)_2$, originates from rearrangements of a few residues with respect to the X-ray structure, sampled along the MD simulation. This structural rearrangement involves residues Ile(65)E7, Val(68)E10 and Phe(69)E11 that, together with Tyr(42)B10, define a possible ligand-entry pathway located between helices E and F.

Discussion

Since only about half of the 2200 bacterial genomes sequenced so far include genes encoding globins, it appears that these proteins may not always be required to sustain bacterial physiology; on the other hand, many bacterial genomes comprise more than one globin gene [1]. The occurrence of the TrHbII + FHb combination in 1162 genomes poses the question whether the function of TrHbIIs is necessarily linked to the NO-dioxygenase activity of FHbs. Different double and ternary combinations of TrHbI, TrHbII, and TrHbIII with FHb are uncommon. *Ph*TAC125 is the only example described so far in which genes encoding a FHb and three TrHbs coexist [11], a feature potentially related to cold environments. In fact, gene redundancy has been shown to entail mechanisms that warrant expression of an essential function [48]; such redundancy may be perceived as a systemic adaptation to extreme environments.

The ability of an organism to survive and grow in the cold depends on a number of adaptive strategies, finalized to maintaining vital cellular functions at low temperatures [49]. In

addition to adaptations at the cellular level, to preserve their functions proteins must maintain a balance between conformational stability and flexibility under the selected environmental conditions. The apparent flexibility of proteins from cold-adapted organisms would compensate for the low kinetic energy of solvent and molecules in their environment. A reduction of surface, of inter-domain or inter-subunit ionic interactions, a decreased number of H bonds and salt bridges are all key mechanisms that can lead to increased conformational flexibility [50]. Several studies indicate that minor structural modifications are sufficient to modify the intrinsic stability of cold-adapted proteins, and that local rather than global flexibility may play an important role in cold adaptation [51].

A hint of enhanced plasticity in cold-adapted *Ph*-2/2HbO was provided by comparative analysis with other members of the family, e.g. *Tf*-2/2HbO and *Mt*-2/2HbO. RMSF analysis over 100 ns of MD simulations at 300 K on all three proteins highlighted significant conformational fluctuations in the *Ph*-2/2HbO EF loop (residues 79-83) compared to the other two TrHbIIs (Fig. S10). Moreover, comparative analysis of structural adaptation in many cold-adapted globins, such as Antarctic fish Hbs and neuroglobins, shows higher flexibility in the CD and EF regions [52-55]. Such findings suggest that helix F and its surroundings may be involved in conferring localized flexibility in cold-adapted globins. The RR data at alkaline pH also strongly support the evidence of an unusually flexible proximal heme pocket in *Ph*-2/2HbO. In fact, the observation of the uncommon 5cHS Fe-OH form at alkaline pH implies that the proximal HisF8-Fe bond must be severed at this pH without denaturing the protein, since the alkaline transition is completely reversible and the protein fully regains its native state upon return to neutral pH.

The comparison of the spectroscopic and crystallographic properties of *Ph*-2/2HbO, here reported, provides additional information regarding protein flexibility. In fact, the spectroscopic observation of a mixture of an aquo-met 6cHS species and a 6cLS component, the latter due to ligation of Tyr(42)B10 to the heme-Fe atom but not detected by the crystal structure, indicates

that heme coordination is affected by the physical and/or environmental conditions experienced by the protein. Such coexistence of two different heme coordination states is in keeping with the proposed conformational flexibility of the heme pocket of this cold-adapted TrHbII. Differences in heme coordination between solution and crystalline states have been previously described in various hemoproteins [56,57]. In addition, analogous results have been reported in *B. subtilis* dye-decolourising peroxidase. The 6cLS state of this protein is the most populated in solution, whereas a 6cHS state is mainly detected in the crystal, such differences being related to a highly flexible heme pocket [58]. In such a general context, it can be proposed that, in the *Ph*-2/2HbO case, the crystallization conditions promote attainment of the aquo-met 6cHS species, or select it for crystal growth as the most populated/stable of the two forms.

In terms of ligand accessibility to the active site, the E7 pathway located between helices E and F was already observed in related TrHbIIs, e.g. *Tf*-2/2HbO and *Mt*-2/2HbO, suggesting similar k_{on} values (10^4-10^5 M^{-1} s^{-1}) for the predominant CO-rebinding process in these proteins [14,15,59,60]. Nevertheless, the crystallographic and MD analyses presented in this work do not provide an explanation for the biphasic nature of CO-rebinding kinetics of *Ph*-2/2HbO [14,15]. Moreover, the RR of the *Ph*-2/2HbO-CO complex indicates the presence of a CO conformer in which polar interactions with the surrounding residues of the distal cavity are absent or very weak [14].

Taken together, these data depict a flexible architecture in the heme pocket of *Ph*-2/2HbO, hardly reported for other globins. The results highlight the unique adaptive structural properties of the protein that enhance its overall flexibility, as recently demonstrated by its resistance to pressure-induced stress [15]. Although some of these adaptive mechanisms are undoubtedly species-specific, the remarkable access of *Ph*-2/2HbO to unusual conformations may represent a molecular property encoded in the TrHbII structure to face the dynamic and functional requirements posed by the Antarctic environment.

Materials and Methods

Protein expression and purification

The *PSHAa0030* gene encoding *Ph*-2/2HbO was cloned as previously described [10]. The previously used fermentor [13] was replaced by small-scale (200 ml) cultures in shake flasks to improve the quality of expression. In fact, production of the protein under strong oxygenation conditions may have given rise to the multiple LS forms observed previously [13]. For overexpression of *Ph*-2/2HbO in *E. coli* BL21(DE3), a single colony carrying the plasmid construct (pET28a-*Ph*-2/2HbO) was inoculated in LB medium supplemented with kanamycin (50 μg/ml), in the presence of 0.3 mM D-aminolevulinic acid, and allowed to grow at 25 °C until A_{600} reached 0.6 OD. The culture was then induced with 1 mM isopropyl-β-D-thiogalactopyranoside, further incubated for 16 h at 25 °C and shaken at 150 rpm. The expression of the cloned protein was monitored after running the cell lysate of recombinant strains in 15% SDS-PAGE followed by Coomassie Brilliant Blue staining.

To purify *Ph*-2/2HbO, the frozen cells, previously harvested by centrifugation at 4 °C, were thawed, suspended in 50 mM Tris–HCl pH 7.6, 1.0 mM EDTA, 1.0 mM phenyl-methyl-sulfonyl-fluoride and protease-inhibitor cocktail (SIGMA P8465), and disrupted in a French press until the supernatant was reddish and clear. The cell debris was removed by centrifugation at 30,000 rpm for 1 h at 4 °C. The supernatant was loaded onto an anion-exchange column (Q Sepharose Fast Flow, GE Healthcare Biosciences), equilibrated with 20 mM Tris–HCl pH 7.6 and 1.0 mM EDTA (Akta Explorer system, GE Healthcare Biosciences, Amersham Biosciences Ltd, UK). *Ph*-2/2HbO was eluted with a NaCl gradient (from 0 to 0.25 M) in 20 mM Tris–HCl pH 7.6, 1.0 mM EDTA. The eluted fractions were collected on the basis of the heme and protein absorbance at 407 and 280 nm, respectively. The pooled fractions were concentrated and loaded onto a second anion-exchange column (HiTrap Q Sepharose XL, GE Healthcare Biosciences) equilibrated with 20 mM Tris–HCl pH 7.6 and 1.0 mM EDTA. *Ph*-2/2HbO was eluted with a NaCl gradient (from 0 to 0.25 M) in 20 mM Tris–HCl pH 7.6,

1.0 mM EDTA. The collected protein was concentrated, dialysed against 50 mM MES pH 6.0, and further purified with a SP (sulfopropyl) Sepharose Fast Flow cation-exchange column (GE Healthcare Biosciences) equilibrated with 50 mM MES pH 6.0. The protein was eluted with a NaCl gradient from 0 to 0.50 M. A fourth purification step (strong anion-exchanger, MonoQ, GE Healthcare Biosciences) was necessary to obtain the protein at 98% purity, as judged from SDS-PAGE.

Ph-2/2HbO was isolated as a cleaved form, as confirmed by mass spectrometry (data not shown), devoid of the first 19 N-terminal residues. This cleaved form is likely the result of proteolysis of the full-length protein by trace amounts of proteases co-purified with the cold-adapted globin [10]. It should be recalled that previous *in vivo* experiments demonstrated that deletion of the N-terminal segment does not impair the NO scavenging activity of *Ph*-2/2HbO [16]. Throughout the following, the truncated form of the protein (i.e. devoid of the first 19 N-terminal residues) will simply be referred to as *Ph*-2/2HbO.

Crystallization and data collection

Crystals of aquo-met *Ph*-2/2HbO (protein concentration 30 mg/ml) were grown at 4 °C by the vapour-diffusion technique, using 8% PEG 8000 and 0.1 M Tris–HCl pH 8.5 as precipitant solution. Crystals grew after several months as bunches of rods from which single elongated prismatic crystals (of about $0.15 \times 0.15 \times 0.4$ mm^3) could be isolated; these were cryoprotected with 15% PEG 8000, 0.1 M Tris–HCl pH 8.5, and 20% glycerol prior to data collection at 100 K. The crystals diffracted up to 2.21 Å resolution on beamline ID23-1, at the ESRF synchrotron (Grenoble, France), and were shown to belong to the orthorhombic space group $P2_12_12_1$, with unit cell parameters a = 42.9 Å, b = 72.3 Å, c = 88.3 Å, $\alpha = \beta = \gamma = 90°$ (Table 1). Calculation of the packing parameter (V_M = 2.41 Å^3/Da, 49% solvent content) indicated the presence of two *Ph*-2/2HbO molecules in the crystal asymmetric unit. Raw diffraction data were processed using Mosflm [61] and SCALA [62].

Structure determination and refinement

Solution of the 3D structure was achieved through molecular-replacement techniques, using the web-server CaspR (http://www.igs.cnrs-mrs.fr/Caspr2/; [63]. The crystal structure of *At*-2/2HbO (PDB code: 2XYK) was used as the search model. Several cycles of manual rebuilding using the program COOT [64], and refinement (rigid body and restrained refinement) using the program REFMAC5 [65], were carried out to improve the initial electron density for the two independent *Ph*-2/2HbO molecules (A and B chains). At the end of refinement, no electron density was present for residues 20-21 of *Ph*-2/2HbO in both A and B chains, and 62 water molecules were located through difference Fourier maps (Table 1). The programs Procheck [17] and Surfnet [66] were used to assess the model stereochemical quality and to search for protein matrix cavities. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (**PDB ID: 4UUR** and **R4UURSF**, respectively).

Spectroscopic measurements

The *Ph*-2/2HbO samples at pH 7.6 and 8.5 were prepared in 0.1 M Tris–HCl, and those at pH 10.9 in 0.1 M CAPS. Protein concentrations in the range 10-70 μM were used for electronic absorption and RR samples. Sample concentration for low-temperature RR was between 30 and 100 μM. The concentration of the EPR samples was 90 μM. The protein concentration was estimated on the basis of the molar absorptivity, $\varepsilon = 131\ mM^{-1}\ cm^{-1}$ at 408 nm.

The hydroxyl complex in isotopically enriched water was prepared by adding 5 μL of *Ph*-2/2HbO in 0.02 M natural abundance buffer, to 50 μL of 0.1 M CAPS buffer prepared with D_2O or $H_2{}^{18}O$ to obtain a final pD 11.0 and pH 10.9, respectively. Isotopically enriched water ($H_2{}^{18}O$) (95%) and D_2O (99.8%) was purchased from Cambridge Isotope Laboratories (USA) and Merck AG (Darmstadt, Germany), respectively. All chemicals were of analytical or reagent grade and were used without further purification.

Electronic absorption spectra were recorded with a double-beam spectrophotometer (Varian Cary 5) using a 1-mm quartz cuvette and a 5-mm NMR tube at a scan rate of 600 nm min^{-1}. The RR spectra were obtained using a 5-mm NMR tube and by excitation with the 413.1 nm line of a Kr^+ laser (Innova 300 C, Coherent, Santa Clara, CA, USA), and the 496.5 mm and 514.5 nm lines of an Ar+ laser (Innova 90/5, Coherent). The backscattered light from a slowly rotating NMR tube was collected and focused into a triple spectrometer (consisting of two Acton Research SpectraPro 2300i working in the subtractive mode, and a SpectraPro 2500i in the final stage), equipped with a liquid nitrogen-cooled CCD detector (Roper Scientific Princeton Instruments). Absorption spectra were measured both prior to and after RR measurements to ensure that no degradation had taken place. The low-temperature experiments were carried out using an Air Products Displex closed-cycle He refrigerator with automatic temperature control. For the low-temperature RR measurements, 20 μL of the protein solution was placed on the copper cold finger of the refrigerator at 90 K, under nitrogen flow, and the RR spectra were recorded at 12 K. The RR spectra were calibrated with indene, n-pentane, dimethylsulfoxide, acetonitrile, and carbon tetrachloride as standards to an accuracy of 1 cm^{-1} for intense isolated bands.

All RR measurements were performed several times under the same conditions to ensure reproducibility. To improve the signal-to-noise ratio, a number of spectra were accumulated and summed only if no spectral differences were noted. All spectra were baseline corrected.

To determine peak bandwidth and positions, a curve-fitting program (Lab Calc; Galactic) was used to simulate the spectra using a Lorentzian line shape. The frequencies of the bands were optimized with an accuracy of 1 cm^{-1} and the bandwidths with an accuracy of 0.5 cm^{-1}.

EPR spectra were recorded with a Bruker Elexys E500, equipped with an NMR gaussmeter and a microwave frequency counter. An Oxford Instruments ESR 900 cryostat was used to obtain low temperatures. Spectra were recorded under non-saturating conditions at 5 K,

1 mW microwave power, and 1 mT modulation amplitude. The g-values were determined by careful visual inspection of the spectra.

Molecular Dynamic Simulations

Simulations were performed starting from the crystal structure of *Ph*-2/2HbO reported herein. Three protein-ligand complexes, displaying CO, hydroxide anion, or molecular O_2 ligands bound to the heme-Fe atom, were built and simulated. The ligands were added at the distal site, bound to Fe(III) (hydroxide) and to Fe(II) (carbon monoxide and O_2), according to the equilibrium structure in an isolated model system. Quantum Mechanics (QM) calculations were performed at the Density Functional Theory (DFT) level (with the Perdew-Burke-Ernzerhof functional, PBE, and 6-31G** basis sets), taking into account structural features of the distal-site environment. The charges and parameters of the Fe(III) heme-hydroxide and the Fe(II) heme-CO and heme-O_2 adducts were determined by the standard procedure, i.e. partial charges were computed using the restricted electrostatic potential (RESP) recipe and DFT electronic-structure calculations with the PBE functional and 6-31G** basis sets. The calculations were performed in the high-spin (HS) state for the hydroxide adduct, and in the low-spin (LS) state for the oxygenated species. Equilibrium distances and angles, as well as force constants, were computed using the same methods and basis set used for the computed charges. The parm99 force field implemented in AMBER12 was used to describe the protein [67]. The protein complexes were then immersed in a pre-equilibrated octahedral box of ~4910 TIP3P water molecules using the tLEaP module of the AMBER12 package [67]. Periodic boundary conditions with a 9 Å cutoff and Ewald sums were used for treating long-range electrostatic interactions. The SHAKE algorithm [68] was used to keep bonds involving H atoms at their equilibrium length, allowing the use of a 2-fs time step for the integration of Newton's equations. The temperature and pressure were kept constant with a Berendsen thermostat and barostat, respectively, as implemented in the AMBER12 package [67]. The His

tautomeric state and protonation (Nε-H, Nδ-H, His^+) were carefully analyzed for each of the six His (His58, His87, His89, His96, His140 and His145), and set to favour the H-bond network suggested by the experimental crystal structure. The equilibration protocol consisted of (*i*) slowly heating the whole system from 0 to 300 K for 20 ps at constant volume, with harmonic restraints of 80 Kcal/ mol $Å^2$ for all Cα atoms, (*ii*) pressure equilibration of the entire system simulated for 1 ns at 300 K with the same restrained atoms. After these two steps, unconstrained 200 ns of MD simulations at constant temperature (300 K) were performed. All structures were stable during the simulations, as determined by RMSD calculation, depicted in Figs. S11 and S12.

Analysis of the small ligand migration free-energy profiles along the protein tunnel

The free energy of the CO migration process within the protein matrix was computed by the ILS approach that uses pre-computed MD simulations in the absence of the ligand and incorporates it afterwards [69]. This method has been thoroughly tested in heme proteins [70]. ILS calculations were performed in a rectangular grid (0.5 Å resolution) that includes the whole simulation box (i.e. protein and solvent), using the CO molecule as a probe. Calculations were performed on 5,000 frames taken from the last 180 ns of simulation time. The values for grid size, resolution and frame numbers have been thoroughly tested in a previous study [70].

Acknowledgements

The authors wish to thank the Centre de Ressources Biologiques de l'Institut Pasteur, Paris, France (http://www.crbip.pasteur.fr) for supplying the *P. haloplanktis* CIP 108707 strain. This study was carried out in the framework of the SCAR programme "Antarctic Thresholds–Ecosystem Resilience and Adaptation" (AnT-ERA). It was financially supported by the Italian National Programme for Antarctic Research (PNRA), the Italian Ministero dell'Istruzione,

dell'Università e della Ricerca (MIUR) (PRIN 2007SFZXZ7, “Structure, function and evolution of heme proteins from Arctic and Antarctic marine organisms: cold adaptation mechanisms and acquisition of new functions”) (to G.S and C.V.). J.P.B. holds a CONICET PhD fellowship. L.B. is a Pew Latin American Fellow. D.A.E. and L.B. are members of CONICET. We thank the ESRF and EMBL-Grenoble staff for assistance and support using the beam line ID23-1, and Dr Maria Fittipaldi for provision of EPR facilities and assistance in recording the spectra.

Author contributions

All authors conceived and designed the experiments. DG, AP, JPB, LB, EC, BDH, AR performed the experiments. All authors analyzed the data, and contributed reagents/materials/analysis tools. DG, AP, DE, MN, GS, MB, CV wrote the paper.

References

1. Vinogradov S, Tinajero-Trejo M, Poole RK & Hoogewijs D (2013) Bacterial and archaeal globins - a revised perspective. *Biochim Biophys Acta* **1834**, 1789–1800.

2. Vuletich DA & Lecomte JT (2006) A phylogenetic and structural analysis of truncated hemoglobins. *J Mol Evol* **62**, 196–210.

3. Pesce A, Bolognesi M & Nardini M (2013) The diversity of 2/2 (truncated) globins. *Adv Microb Physiol* **63**, 49–78.

4. Milani M, Savard PY, Ouellet H, Ascenzi P, Guertin M & Bolognesi M (2003) A TyrCD1/TrpG8 hydrogen bond network and a TyrB10TyrCD1 covalent link shape the heme distal site of *Mycobacterium tuberculosis* hemoglobin O. *Proc Natl Acad Sci USA* **100**, 5766–5771.

5. Bonamore A, Ilari A, Giangiacomo L, Bellelli A, Morea V & Boffi A (2005) A novel thermostable hemoglobin from the actinobacterium *Thermobifida fusca*. *FEBS J* **272**, 4189–4201.

6. Giangiacomo A, Ilari L, Boffi A, Morea V & Chiancone E (2005) The truncated oxygen-avid hemoglobin from *Bacillus subtilis*: X-ray structure and ligand binding properties. *J Biol Chem* **280**, 9192–9202.

7. Ilari A, Kjelgaard P, von Wachenfeldt C, Catacchio B, Chiancone E & Boffi A (2007) Crystal structure and ligand binding properties of the truncated hemoglobin from *Geobacillus stearothermophilus*. *Arch Biochem Biophys* **457**, 85–94.

8. Pesce A, Nardini M, Labarre M, Richard C, Wittenberg JB, Wittenberg BA, Guertin M & Bolognesi M (2011) Structural characterization of a group II 2/2 hemoglobin from the plant pathogen *Agrobacterium tumefaciens*. *Biochim Biophys Acta* **1814**, 810–816.

9. Reeder BJ & Hough MA (2014) The structure of a class 3 nonsymbiotic plant haemoglobin from *Arabidopsis thaliana* reveals a novel N-terminal helical extension. *Acta Crystallogr D Biol Crystallogr* **70**, 1411–1418.

10. Giordano D, Parrilli E, Dettaï A, Russo R, Barbiero G, Marino G, Lecointre G, di Prisco G, Tutino L & Verde C (2007) The truncated hemoglobins in the Antarctic psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Gene* **398**, 69–77.

11. Giordano D, Coppola D, Russo R, Tinajero-Trejo M, di Prisco G, Lauro F, Ascenzi P & Verde C (2013) The globins of cold-adapted *Pseudoalteromonas haloplanktis* TAC125: from the structure to the physiological functions. *Adv Microb Physiol* **63**, 329–89.

12. Parrilli E, Giuliani M, Giordano D, Russo R, Marino G, Verde C & Tutino ML (2010) The role of a 2-on-2 haemoglobin in oxidative and nitrosative stress resistance of Antarctic *Pseudoalteromonas haloplanktis* TAC125. *Biochimie* **92**, 1003–1009.

13. Howes BD, Giordano D, Boechi L, Russo R, Mucciacciaro S, Ciaccio C, Sinibaldi F, Fittipaldi M, Martı MA, Estrin DA, di Prisco G, Coletta M, Verde C & Smulevich G (2011) The peculiar heme pocket of the 2/2 hemoglobin of cold-adapted *Pseudoalteromonas haloplanktis* TAC125. *J Biol Inorg Chem* **16**, 299–311.

14. Giordano D, Russo R, Ciaccio C, Howes BD, di Prisco G, Smulevich G, Coletta M & Verde C (2011) Ligand- and proton-linked conformational changes of the ferrous 2/2 hemoglobin of *Pseudoalteromonas haloplanktis* TAC125. *IUBMB Life* **63**, 566–573.

15. Russo R, Giordano D, di Prisco G, Hui Bon Hoa G, Marden MC, Verde C & Kiger L (2013) Ligand-rebinding kinetics of 2/2 hemoglobin from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Biochim Biophys Acta* **1834**, 1932–1938.

16. Coppola D, Giordano D, Tinajero-Trejo M, di Prisco G, Ascenzi P, Poole RK & Verde C (2013) Antarctic bacterial hemoglobin and its role in the protection against nitrogen reactive species. *Biochim Biophys Acta* **1834**, 1923–1931.

17. Laskowski RA, MacArthur MW, Moss DS & Thornton JM (1993) PROCHECK, a program to check the stereochemical quality of protein structure. *J Appl Crystallogr* **26**, 283–291.

18. Krissinel E & Henrick K (2005) Detection of Protein Assemblies in Crystals. In: *Computational Life Science*, (Berthold MR, Glen RC, Diederichs K, Kohlbacher O & Fischer I ed), pp. 163–174. Springer Berlin Heidelberg.

19. Nardini M, Pesce A, Milani M & Bolognesi M (2007) Protein fold and structure in the truncated (2/2) globin family. *Gene* **398**, 2–11.

20. Milani M, Pesce A, Nardini M, Ouellet H, Ouellet Y, Dewilde S, Bocedi A, Ascenzi P, Guertin M, Moens L, Friedman JM, Wittenberg JB & Bolognesi M (2005) Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins. *J Inorg Biochemistry* **99**, 97–109.

21. Milani M, Pesce A, Ouellet H, Guertin M & Bolognesi M (2003) Truncated hemoglobins and nitric oxide action. *IUBMB Life* **55**, 623–627.

22. Egawa T & Yeh SR (2005) Structural and functional properties of hemoglobins from unicellular organisms as revealed by resonance Raman spectroscopy. *J Inorg Biochem* **99**, 72–96.

23. Das TK, Couture M, Lee HC, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB & Guertin M (1999) Identification of the ligands to the ferric heme of *Chlamydomonas* chloroplast hemoglobin: evidence for ligation of tyrosine-63 (B10) to the heme. *Biochemistry* **38**, 15360–15368.

24. Caillet-Saguy C, Turano P, Piccioli M, Lukat-Rodgers GS, Czjzek M, Guigliarelli B, Izadi-Pruneyre N, Rodgers KR, Delepierre M & Lecroisey A (2008) Deciphering the structural role of histidine 83 for heme binding in hemophore HasA. *J Biol Chem* **283**, 5960–5970.

25. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P & Rivera M (2009) Structural characterization of the hemophore HasAp from *Pseudomonas aeruginosa*: NMR spectroscopy reveals protein-protein interactions between Holo-HasAp and hemoglobin. *Biochemistry* **48**, 96–109.

26. Smulevich G, Miller MA, Kraut J & Spiro TG (1991) Conformational change and histidine control of heme chemistry in cytochrome c peroxidase: resonance Raman evidence from Leu-52 and Gly-181 mutants of cytochrome c peroxidase. *Biochemistry* **30**, 9546–9558.

27. Que L (1988) Metal-tyrosinate proteins. In: *Biological applications of Raman spectroscopy*, (Spiro TG, ed), pp. 491–521. Wiley, NY.

28. Nagai M, Yoneyama Y & Kitagawa T (1989) Characteristics in tyrosine coordinations of four hemoglobins M probed by resonance Raman spectroscopy. *Biochemistry* **28**, 2418–2422.

29. Aki Y, Nagai M, Nagai Y, Imai K, Aki M, Sato A, Kubo M, Nagatomo S & Kitagawa T (2010) Differences in coordination states of substituted tyrosine residues and quaternary structures among hemoglobin M probed by resonance Raman spectroscopy. *J Biol Inorg Chem* **15**, 147–158.

30. Sharma KD, Andersson LA, Loehr TM, Terner J & Goff HM (1989) Comparative spectral analysis of mammalian, fungal, and bacterial catalases. Resonance Raman evidence for iron-tyrosinate coordination. *J Biol Chem* **264**, 12772–12779.

31. Eakanunkul S, Lukat-Rodgers GS, Sumithran S, Ghosh A, Rodgers KR, Dawson JH & Wilks A (2005) Characterization of the periplasmic heme-binding protein shut from the heme uptake system of *Shigella dysenteriae*. *Biochemistry* **44**, 13179–13191.

32. Nicoletti FP, Howes BD, Fittipaldi M, Fanali G, Fasano M, Ascenzi P & Smulevich G (2008) Ibuprofen induces an allosteric conformational transition in the heme complex of human serum albumin with significant effects on heme ligation. *J Am Chem Soc* **130**, 11677–11688.

33. Reeder BJ, Svistunenko DA & Wilson MT (2011) Lipid binding to cytoglobin leads to a change in haem co-ordination: a role for cytoglobin in lipid signalling of oxidative stress. *Biochemistry* **434**, 483–492.

34. Cerda-Colon JF, Silfa E & Lopez-Garriga J (1998) Unusual rocking freedom of the heme in the hydrogen sulfide-binding hemoglobin from *Lucina pectinata*. *J Am Chem Soc* **120**, 9312–9317.

35. Kraus DW, Wittenberg JB, Lu JF & Peisach J (1990) Hemoglobins of the *Lucina pectinata*/bacteria symbiosis. II. An electron paramagnetic resonance and optical spectral study of the ferric proteins. *J Biol Chem* 265, 16054–16059.

36. Feis A, Marzocchi MP, Paoli M & Smulevich G (1994) Spin state and axial ligand bonding in the hydroxide complexes of metmyoglobin, methemoglobin, and horseradish peroxidase at room and low temperatures. *Biochemistry* **33**, 4577–4583.

37. Mukai M, Savard PY, Ouellet H, Guertin M & Yeh SR (2002) Unique ligand-protein interactions in a new truncated hemoglobin from *Mycobacterium tuberculosis*. *Biochemistry* **41**, 3897–3905.

38. Nicoletti FP, Bustamante JP, Droghetti E, Howes BD, Fittipaldi M, Bonamore A, Baiocco P, Feis A, Boffi A, Estrin DA & Smulevich G (2014) Interplay of the H-Bond Donor−Acceptor Role of the Distal Residues in Hydroxyl Ligand Stabilization of Thermobifida fusca Truncated Hemoglobin. *Biochemistry* 53, 8021−8030.

39. Yeh SR, Couture M, Ouellet Y, Guertin M & Rousseau DL (2000) A cooperative oxygen binding hemoglobin from *Mycobacterium tuberculosis*. Stabilization of heme ligands by a distal tyrosine residue. *J Biol Chem* **275**, 1679–1684.

40. Couture M, Das TK, Savard P-Y, Ouellet Y, Wittenberg JB, Wittenberg BA, Rousseau DL & Guertin M (2000) Structural investigations of the hemoglobin of the cyanobacterium Synechocystis PCC6803 reveal a unique distal heme pocket. Eur. J. Biochem. 267, 4770-4780.

41. Das TK, Franzen S, Pond A, Dawson JH & Rousseau DL (1999) Formation of a five-coordinate hydroxide-bound heme in the His93Gly mutant of sperm whale myoglobin. *Inorg Chem* **38**, 1952–1953.

42. Das TK, Boffi A, Chiancone E & Rousseau DL (1999) Hydroxide rather than histidine is coordinated to the heme in five-coordinate ferric *Scapharca inaequivalvis* hemoglobin. *J Biol Chem* **274**, 2916–2919.

43. Howes BD, Rodriguez-Lopez JN, Smith AT & Smulevich G (1997) Mutation of distal residues of horseradish peroxidase: influence on substrate binding and cavity properties. *Biochemistry* **36***,* 1532–1543.

44. Boechi L, Martí MA, Milani M, Bolognesi M, Luque FJ & Estrin DA (2008) Structural determinants of ligand migration in *Mycobacterium tuberculosis* truncated hemoglobin O. *Proteins* **73**, 372–379.

45. Boechi L, Mañez PA, Luque FJ, Marti MA & Estrin DA (2010) Unraveling the molecular basis for ligand binding in truncated hemoglobins: the trHbO *Bacillus subtilis* case. *Proteins* **78(4)**, 962–70.

46. Droghetti E, Nicoletti FP, Bonamore A, Boechi L, Arroyo Manez P, Estrin DA, Boffi A, Smulevich G & Feis A (2010) Heme pocket structural properties of a bacterial truncated hemoglobin from *Thermobifida fusca*. *Biochemistry* **49**, 10394–10402.

47. Feis A, Lapini B, Catacchio B, Brogioni S, Foggi P, Chiancone E, Boffi A & Smulevich G (2008) Unusually strong H-bonding to the heme ligand and fast geminate recombination dynamics of the carbon monoxide complex of *Bacillus subtilis* truncated hemoglobin. *Biochemistry* **47**, 902–910.

48. Chen Z, Cheng CH, Zhang J, Cao L, Chen L, Zhou L, Jin Y, Ye H, Deng C, Dai Z, Xu Q, Hu P, Sun S, Shen Y & Chen L (2008) Transcriptomic and genomic evolution under constant cold in Antarctic notothenioid fish. *Proc Natl Acad Sci USA* **105**, 12944–12949.

49. Rodrigues DF & Tiedje JM (2008) Coping with our cold planet. *Appl Eviron Microbiol* **74**, 1677–1686.

50. D'Amico S, Collins T, Marx JC, Feller G & Gerday C (2006) Psychrophilic microorganisms: challenges for life. *EMBO Rep* **7**, 385–389.

51. Siddiqui KS & Cavicchioli R (2006) Cold-adapted enzymes. *Annu Rev Biochem* **75**, 403–433.

52. Riccio A, Vitagliano L, di Prisco G, Zagari A & Mazzarella L (2002) The crystal structure of a tetrameric hemoglobin in a partial hemichrome state. *Proc Natl Acad Sci USA* **99**, 9801–9806.

53. Giordano D, Boechi L, Vergara A, Martí MA, Samuni U, Dantsker D, Grassi L, Estrin DA, Friedman JM, Mazzarella L, di Prisco G & Verde C (2009) The hemoglobins of the sub-Antarctic fish *Cottoperca gobio*, a phyletically basal species--oxygen-binding equilibria, kinetics and molecular dynamics. *FEBS J* **276**, 2266–2277.

54. Boron I, Russo R, Boechi L, Cheng CH, di Prisco G, Estrin DA, Verde C & Nadra AD (2011) Structure and dynamics of Antarctic fish neuroglobin assessed by computer simulations. *IUBMB Life* 63, 206–213.

55. Balsamo A, Sannino F, Merlino A, Parrilli E, Tutino ML, Mazzarella L & Vergara A (2012) Role of the tertiary and quaternary structure in the formation of bis-histidyl adducts in cold-adapted hemoglobins. *Biochimie* **94**, 953–960.

56. Smulevich G, Wang Y, Mauro JM, Wang JM, Fishel LA, Kraut J & Spiro TG (1990) Single-crystal resonance Raman spectroscopy of site-directed mutants of cytochrome c peroxidase. *Biochemistry* **29**, 7174-7180.

57. Indiani C, Santoni E, Becucci M, Boffi A, Fukuyama F & Smulevich G (2003) New insight into the peroxidase-hydroxamic acid interaction revealed by the combination of spectroscopic and crystallographic studies. *Biochemistry* **42***,* 14066-14074.

58. Sezer M, Santos A, Kielb P, Pinto T, Martins LO & Todorovic S (2013) Distinct structural and redox properties of the heme active site in bacterial dye decolorizing peroxidase-type peroxidases from two subfamilies: resonance Raman and electrochemical study. *Biochemistry* **52**, 3074−3084.

59. Ouellet H, Juszczak L, Dantsker D, Samuni U, Ouellet YH, Savard PY, Wittenberg JB, Wittenberg BA, Friedman JM & Guertin M (2003) Reactions of *Mycobacterium tuberculosis* truncated hemoglobin O with ligands reveal a novel ligand-inclusive hydrogen bond network. *Biochemistry* **42**, 5764–5774.

60. Marcelli A, Abbruzzetti S, Bustamante JP, Feis A, Bonamore A, Boffi A, Gellini C, Salvi PR, Estrin DA, Bruno S, Viappiani C & Foggi P (2012) Following ligand migration pathways from picoseconds to milliseconds in type II truncated hemoglobin from *Thermobifida fusca*. *PloS One* **7**, e39884.

61. Leslie AGM (2003) *MOSFLM User Guide, Mosflm Version 6.2.3*, MRC Laboratory of Molecular Biology, Cambridge, UK.

62. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82.

63. Claude JB, Suhre K, Notredame C, Claverie JM & Abergel C (2004) CaspR: a web-server for automated molecular replacement using homology modelling. *Nucleic Acids Res* **32**, 606–609.

64. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.

65. Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.

66. Laskowski RA (1995) SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* **13**, 323–330.

67. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham III TE, DeBolt S, Ferguson D, Seibel G & Kollman P (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations

to simulate the structural and energetic properties of molecules. *Comput Phys Commun* **91**, 1–41.

68. Ryckaert JP, Ciccotti G & Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **23**, 327–341.

69. Cohen J, Olsen KW & Schulten K (2008) Finding gas migration pathways in proteins using implicit ligand sampling. *Methods Enzymol* **437**, 439–457.

70. Forti F, Boechi L, Estrin DA & Marti MA (2011) Comparing and combining implicit ligand sampling with multiple steered molecular dynamics to study ligand migration processes in heme proteins. *J Comput Chem* **32(10)**, 2219–2231.

Supporting information

Table S1. Comparison of the EPR spectral parameters of various LS hemoproteins.

Table S2. Vibrational assignment and RR frequencies (cm^{-1}) of *Ph*-2/2HbO at pH 7.6 in the high-frequency region, based on the curve fitting analysis of the RR spectra in Fig. S2.

Table S3. Vibrational assignment and RR frequencies (cm^{-1}) of *Ph*-2/2HbO at pH 7.6 in the low-frequency region, based on the curve fitting analysis of the RR spectra in Fig. S3.

Table S4. Vibrational assignment and RR frequencies (cm^{-1}) of *Ph*-2/2HbO at pH 10.7 in the high-frequency region, based on the curve fitting analysis of the RR spectra in Fig. S4.

Table S5. Vibrational assignment and RR frequencies (cm^{-1}) of *Ph*-2/2HbO at pH 10.7 in the low-frequency region, based on the curve fitting analysis of the RR spectra in Fig. S5.

Fig. S1. The quaternary assembly of two *Ph*-2/2HbO chains observed in the crystal asymmetric unit.

Fig. S2. Curve fittings of the *Ph*-2/2HbO RR spectra in the high-frequency region shown in Fig. 5 obtained with excitation at 413.1, 496.5 and 514.5 nm (pH 7.6).

Fig. S3. Curve fittings of the *Ph*-2/2HbO RR spectra in the low-frequency region shown in Figure 5 obtained with excitation at 413.1, 496.5 and 514.5 nm (pH 7.6).

Fig. S4. Curve fittings of the *Ph*-2/2HbO RR spectra in the high-frequency region shown in Figure 5 obtained with excitation at 514.5 nm (pH 10.7).

Fig. S5. Curve fittings of the *Ph*-2/2HbO RR spectra in the low-frequency region shown in Figure 5 obtained with excitation at 514.5 nm (pH 10.7).

Fig. S6. Comparison of the RR spectra of *Ph*-2/2HbO at pH 7.6 and 10.7 at 298 K and 12 K.

Fig. S7. Low-frequency region RR spectra of alkaline *Ph*-2/2HbO at 298 K, in H_2O, D_2O, and $H_2{}^{18}O$ buffered solutions.

Fig. S8. Schematic representation of the X-ray distal heme cavity of *Ph*-2/2HbO (light orange) and the average structure obtained by MD simulations in the absence of a distal heme ligand.

Fig. S9. Time evolution of Fe-O_{Tyr}(42)B10 distances for *Ph*-2/2HbO (black), *Mt*-2/2HbO (red) and *Tf*-2/2HbO (green).

Fig. S10. RMSF (Å) for backbone atoms in WT deoxy forms of *Ph*-2/2HbO (black), *Mt*-2/2HbO (green) and *Tf*-2/2HbO (red) at 300K. *Ph*-2/2HbO shows major fluctuations in the EF loop (residues 79 to 83).

Fig. S11. Time evolution (ns) of RMSD (Å) for the backbone atoms in the deoxy form of *Ph*-2/2HbO (black), and for OH^- heme liganded ferric species (green).

Fig. S12. Time evolution (ns) of RMSD (Å) for backbone atoms of *Ph*-2/2HbO hosting CO (blue) and O_2 (red) as ligands bound to the heme.

Table 1. Data collection and crystallographic refinement statistics for *Ph*-2/2HbO.

Data collection	
Space group	$P2_12_12_1$
Cell dimensions:	
a, *b*, *c* (Å)	42.9, 72.3, 88.3
α, β, γ (°)	90, 90, 90
Wavelength (Å)	0.979
Resolution (Å)	72.25-2.21 (2.33-2.21)[a]
No. reflections	63042
Unique reflections	14322
R_{merge}[b]	0.102 (0.451)
$I/\sigma(I)$	8.7 (3.1)
Completeness (%)	99.5 (99.7)
Multiplicity	4.4 (4.6)
Refinement	
Resolution (Å)	44.15-2.21
R_{factor}/R_{free} (%)	18.3/24.6
No. of residues/protein atoms	248 (2x124, from residue 22 to 145)/2132
No. of heme groups	2
No. of water molecules	62
B-factors (Å^2):	
Protein	40.1
Heme group	31.3
Water molecules	38.3
RMSD from ideality:	
bond lengths (Å)	0.015
bond angles (°)	1.7
Ramachandran plot[c]:	
most favored regions (%)	94.8
additional allowed regions (%)	5.2

[a]Values in parentheses are for highest-resolution shell

[b]$R_{merge} = \Sigma_h\Sigma_i| I_{hi} - \langle I_h \rangle| / \Sigma_h\Sigma_i I_{hi}$.

[c]Data produced using the program PROCHECK [17]

Table 2. Tyrosinate vibrational bands (cm^{-1}) of ferric-heme proteins with tyrosinate ligation.

Protein	Spin state	ν_{Tyr}(C=C)	ν_{Tyr}(C=C)	ν_{Tyr}(C-O)	ν(Fe-O_{Tyr})	Reference
Ph-2/2HbO	6cLS	*	1509	1312	590	This work
Chlamydomonas Hb	6cLS	1595	1500	1308	502	[23]
Hb M Saskatoon	6cHS	1607	1504	1300	581/598	[28,29]
Hb M Boston	5cHS	1603	1504	1279	603	[28]
Hb M Iwate	5cHS	1607	1504	1308	588	[28]
Hb M Hyde Park	5cHS	1609	1502	1300	588	[28]
Aspergillus niger catalase	5cHS	1615	-	1245	-	[30]
Shigella dysenteriae ShuT	5cHS	1601	1502	1301/1265	613	[31]

* This band is not observed, as it overlaps those at 1602 cm^{-1} and 1609 cm^{-1} (ν_{37} and ν_{10}) determined by a curve-fitting analysis (Fig. S2, Table S2).

Table 3. RR frequencies (cm^{-1}) of the *v*(Fe-OH) modes of various heme proteins at alkaline pH. The observed isotopic shifts are in parentheses.

Protein	6cHS (D_2O, $H_2{}^{18}O$)	6cLS (D_2O, $H_2{}^{18}O$)	5cHS (D_2O, $H_2{}^{18}O$)	Reference
Ph-2/2HbO		528 (+6,-18)	568 (-14,-19)	This work
Mt-2/2HbO	446 (nr, 22)	533 (nr, -30)		[37]
Tf-2/2HbO		485 (+4, -14)		[38]
Mt-2/2HbN	454 (nr,-31)	561 (nr,-28)		[39]
2/2Hb *Synechocystis*			573 (nr, -24)	[40]
Mb (horse heart)	491 (-14,-23)	550 (-12,-25)		[36]
Hb (human)	492 (-13,nr)	553 (-9,nr)		[36]
HRPC		516 (+6,-19)		[36]
H93G Mb (sperm whale)			575 (-13,-24)	[41]
Scapharca inaequivalvis Hb			578 (nr,-25)	[42]

Figure legends

Fig. 1. Ribbon representation of the three-dimensional fold of *Ph*-2/2HbO. The heme group is in red and the helices are labelled; proximal His(96)F8 is shown as stick model. Helix A is just a one-turn helical segment following the N terminus. Note that the presence of the additional helix Φ, between helices E and F, is specific for TrHbIIs.

Fig. 2. Structure-based sequence alignment of *Ph*-2/2HbO with members of the TrHbII family whose crystal structures are available. Helical regions are shaded in grey and the TrHbII-specific helix Φ is indicated; residues (B10, CD1 and G8) specific for TrHbIIs are in yellow; HisF8, conserved in the globin superfamily, is in magenta; the Gly-Gly motifs are in cyan. The main differences (in sequence and structure) are located in the N- and C-terminal regions, and in the BC, CE, and GH loops.

Fig. 3. (A): Location of distal and exposed cavities throughout the protein matrix in *Ph*-2/2HbO, separated by the Ile(65)E7 side chain. The heme group is in red and the cavity system is highlighted as a pink mesh. The residues defining an alternative access to the heme distal site are labelled. (B): Heme pocket structure in *Ph*-2/2HbO. The Figure shows some of the residues involved in stabilization of the heme through Fe coordination, i.e. His(96)F8, salt bridges and/or H bonds with Arg(64)E6 and Arg(95)F7. The Figure also displays the protein environment surrounding the heme-Fe atom coordinated water molecule (W32, shown as a magenta sphere) in the A subunit. The side chains of residues involved in interactions with the coordinated water molecule are shown and labelled. The additional water molecule W13 (magenta sphere), that is an integral part of the distal site H-bonded network stabilising the heme ligand, is also shown.

Fig. 4. Ferric *Ph*-2/2HbO at pH 7.6 and 10.9. (A): UV–vis absorption (continuous line) and **second derivative** (dotted line) spectra. The visible region has been expanded five-fold. Spectra have been shifted along the ordinate axis to allow better visualization. (B): X-band EPR spectrum. Spectra were recorded at 5 K, 9.39-GHz microwave frequency, 1-mW microwave power and 10 G modulation amplitude. (C, D): RR spectra were recorded with the 413.1-nm excitation wavelength. Experimental conditions: 15-mW laser power at the sample, 1-cm^{-1} spectral resolution, average of twelve spectra (C) and four spectra (D) with 600-s integration time. The intensities are normalized to those of the ν_7 band (678 cm^{-1}) (C) and the ν_4 band (1371 cm^{-1}) (D).

Fig. 5. RR spectra of *Ph*-2/2HbO at pH 7.6 and 10.7 at the excitation wavelengths shown. Experimental conditions: pH 7.6: 413.1 nm, see Fig. 4; 496.5 nm 70-mW laser power at the sample, 2.4-cm^{-1} spectral resolution, average of 9 spectra with 600-s integration time (high frequency), average of 36 spectra with 600-s integration time (low frequency); 514.5 nm, 140-mW laser power at the sample, 2.2-cm^{-1} spectral resolution, average of 4 spectra with 600-s integration time (high frequency); average of 18 spectra with 600-s integration time (low frequency); pH 10.7: 514.5 nm, average of 15 spectra with 600-s integration time (high frequency); average of 36 spectra with 600-s integration time (low frequency). The tyrosinate modes are in blue. The asterisks indicate bands due to the alkaline buffer CAPS. The δ(H-O-H) water bending mode has been subtracted from the spectra taken in the visible region.

Fig. 6. (A and C): low-frequency region RR spectra of alkaline *Ph*-2/2HbO, at 12 and 298 K, respectively, obtained in H_2O, D_2O, and $H_2{}^{18}O$ buffered solutions. (B and D): difference spectra, H_2O-D_2O, D_2O-$H_2{}^{18}O$, and H_2O-$H_2{}^{18}O$, at 12 and 298 K, respectively. Experimental conditions as in Figure 4; 298 K: average of 12 spectra with 600 s integration time (H_2O, D_2O), average of 4 spectra with 600 s integration time ($H_2{}^{18}O$); 12 K: average of 12 spectra with 600 s integration time.

Fig. 7. RR spectra of ferric *Ph*-2/2HbO at pH 7.6 (blue), 8.5 (magenta) and 10.9 (black). Experimental conditions as in Fig. 4. Average of twelve spectra (A) and four spectra (B) with 600-s integration time. The intensities are normalized to those of the ν_7 band (678 cm^{-1}) (A) and of the ν_4 band (1371 cm^{-1}) (B). The red arrows indicate the changes in relative intensity or frequency upon raising pH.

Figure 8 Schematic representation of the distal heme cavity of *Ph*-2/2HbO with CO (A) and O_2 (B) ligands stabilized by H-bonds. Time evolution of X-N_{TrpG8} (black) and X-O_{TyrB10} (red) distances showing H-bond network interactions during the timescale of simulation for X=CO (C) and X=O_2 (D) are also shown. Distances are in Å.

Figure 9 Schematic representation of the heme distal pocket, the open E7 tunnel, and the cavity system of *Ph*-2/2HbO along the MD simulation (violet surface), superimposed on the crystal structure (light orange atoms and yellow cavity).

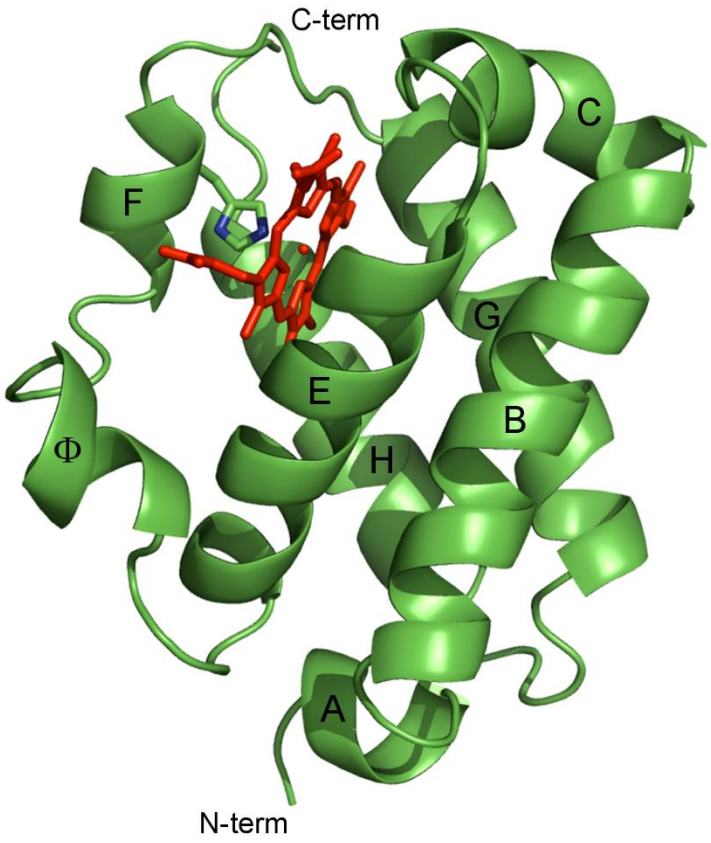

Fig. 1

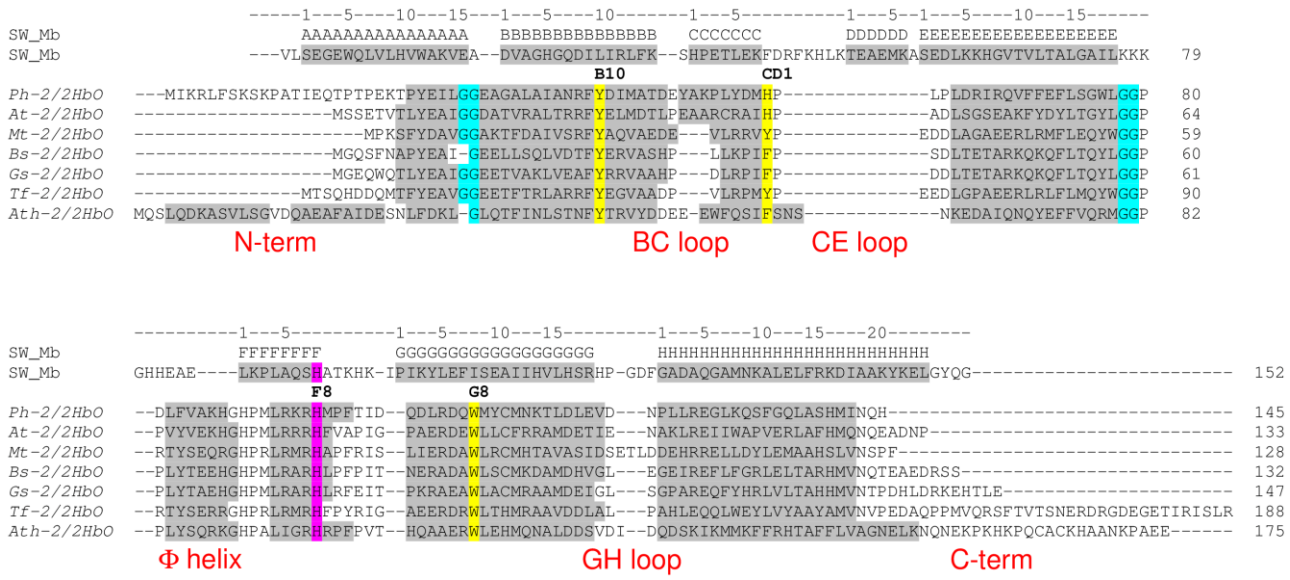

Fig. 2

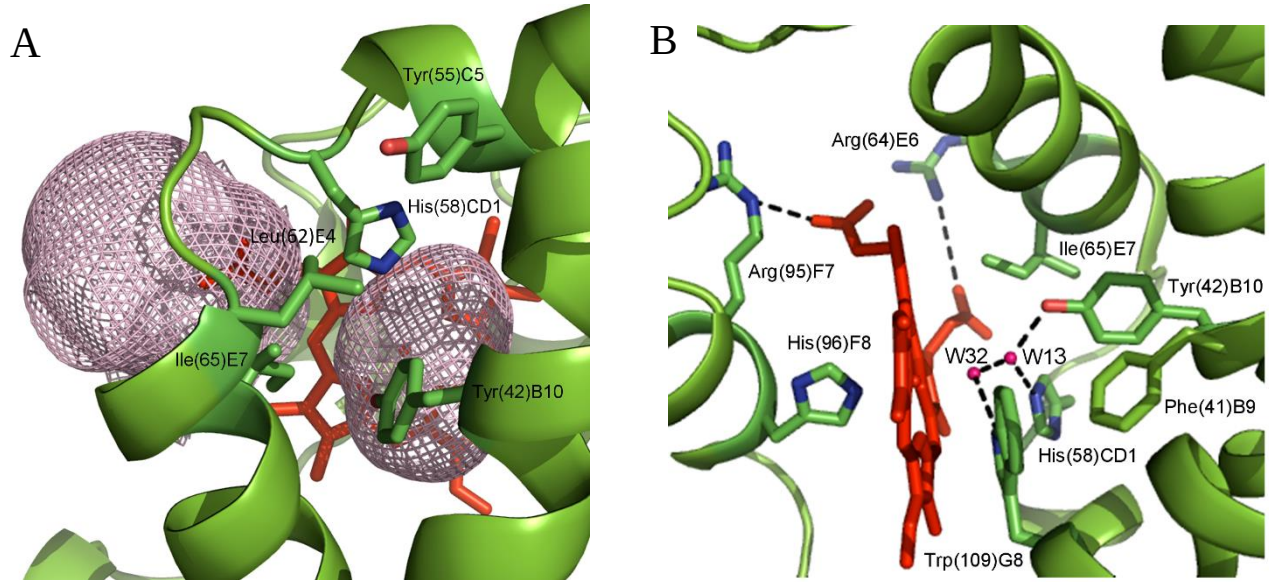
A
Tyr(55)C5
His(58)CD1
Leu(62)E4
Ile(65)E7
Tyr(42)B10
B
Arg(64)E6
Ile(65)E7
Arg(95)F7
Tyr(42)B10
His(96)F8
W32
W13
Phe(41)B9
His(58)CD1
Trp(109)G8

Fig. 3

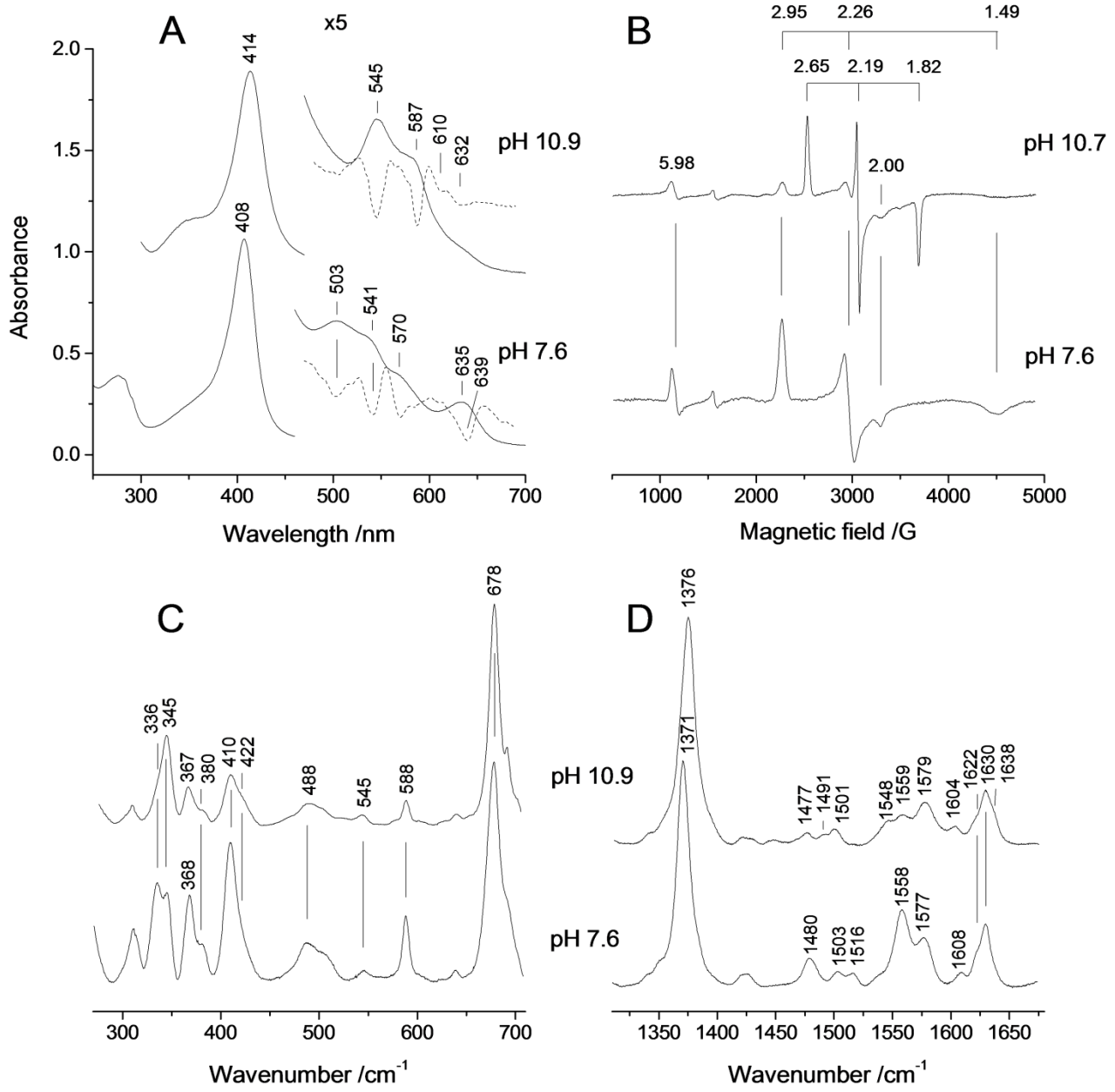
A
x5
414
545
587
610
632
pH 10.9
408
503
541
570
635
639
pH 7.6
Absorbance
Wavelength /nm
B
2.95
2.26
1.49
2.65
2.19
1.82
5.98
2.00
pH 10.7
pH 7.6
Magnetic field /G
C
678
336
345
367
380
410
422
488
545
588
368
Wavenumber /cm-1
D
1376
1371
pH 10.9
1477
1491
1501
1548
1559
1579
1604
1622
1630
1638
1558
1577
1480
1503
1516
1608
pH 7.6
Wavenumber /cm-1

Fig. 4

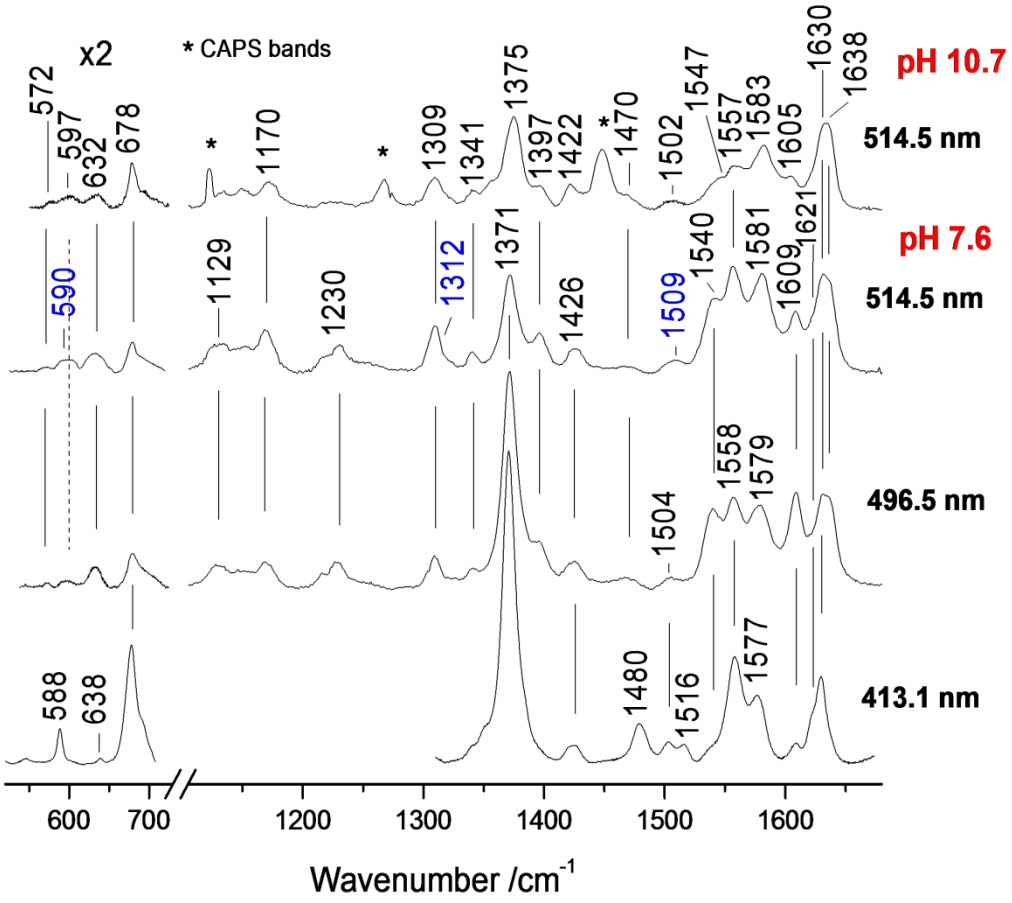
x2
* CAPS bands
572
597
632
678
*
1170
*
1309
1341
1375
1397
1422
*
1470
1502
1547
1557
1583
1605
1630
1638
pH 10.7
514.5 nm
590
1129
1230
1312
1371
1426
1509
1540
1581
1609
1621
pH 7.6
514.5 nm
1504
1558
1579
496.5 nm
588
638
1480
1516
1577
413.1 nm
600
700
1200
1300
1400
1500
1600
Wavenumber /cm-1

Fig. 5

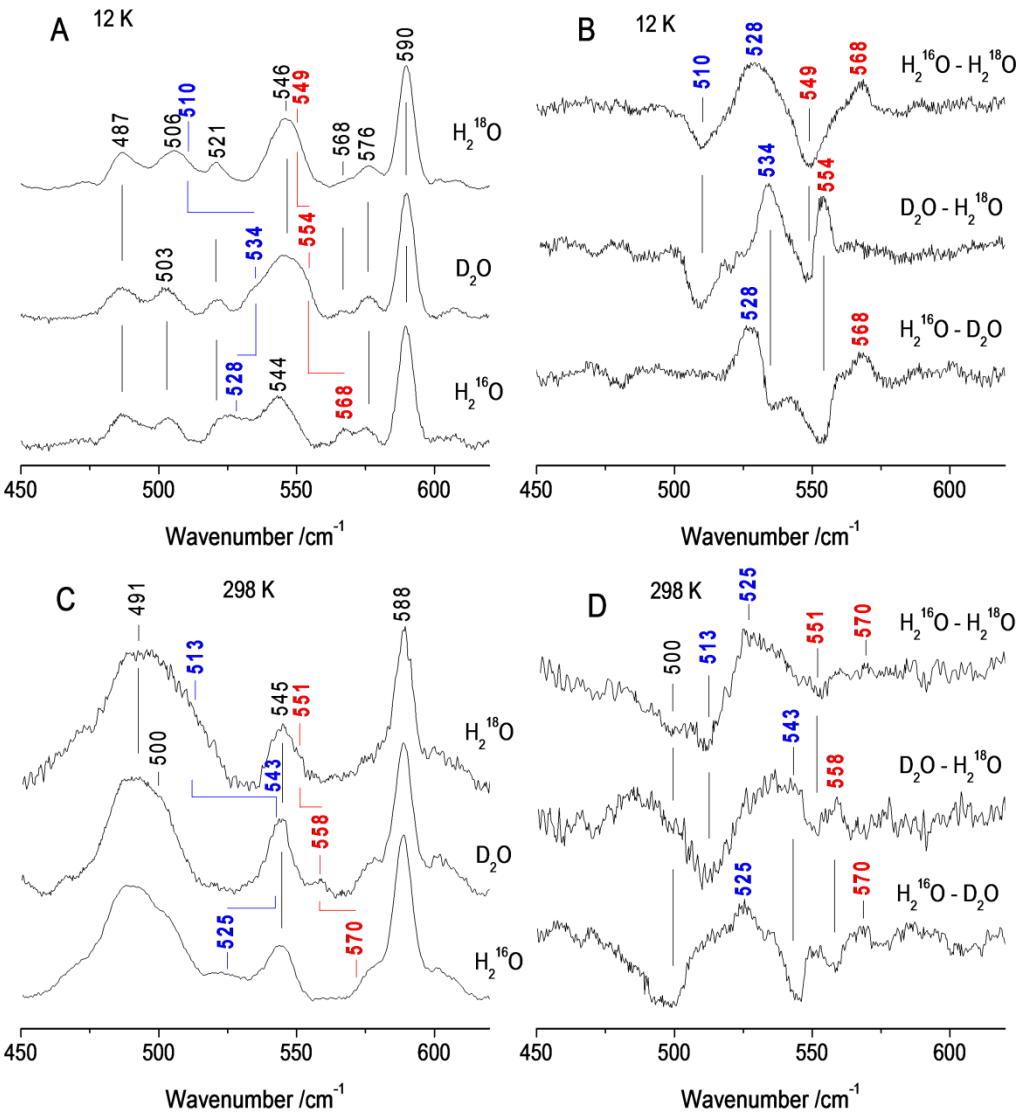
A
12 K
487
506
510
521
546
549
568
576
590
$H_2^{18}O$
503
534
554
D_2O
528
544
568
$H_2^{16}O$
450
500
550
600
Wavenumber /cm^{-1}
B
12 K
510
528
549
568
$H_2^{16}O - H_2^{18}O$
534
554
$D_2O - H_2^{18}O$
528
568
$H_2^{16}O - D_2O$
C
298 K
491
513
545
551
588
$H_2^{18}O$
500
543
558
D_2O
525
570
$H_2^{16}O$
D
298 K
525
500
513
551
570
$H_2^{16}O - H_2^{18}O$
543
558
$D_2O - H_2^{18}O$
525
570
$H_2^{16}O - D_2O$

Fig. 6

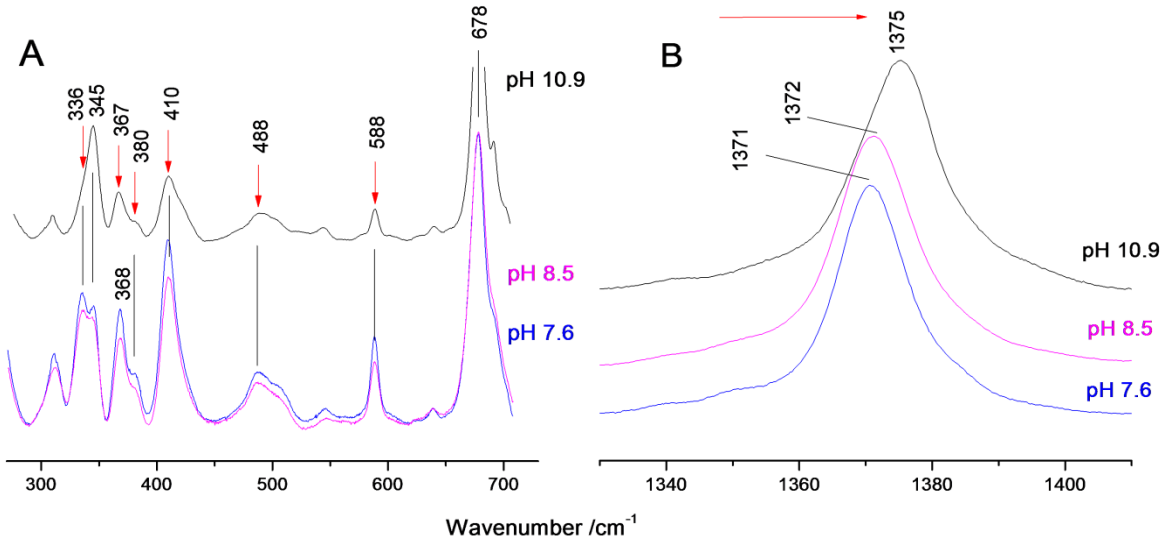
A
336
345
367
380
410
488
588
678
pH 10.9
368
pH 8.5
pH 7.6
300
400
500
600
700
B
1375
1372
1371
pH 10.9
pH 8.5
pH 7.6
1340
1360
1380
1400
Wavenumber /cm-1

Fig. 7

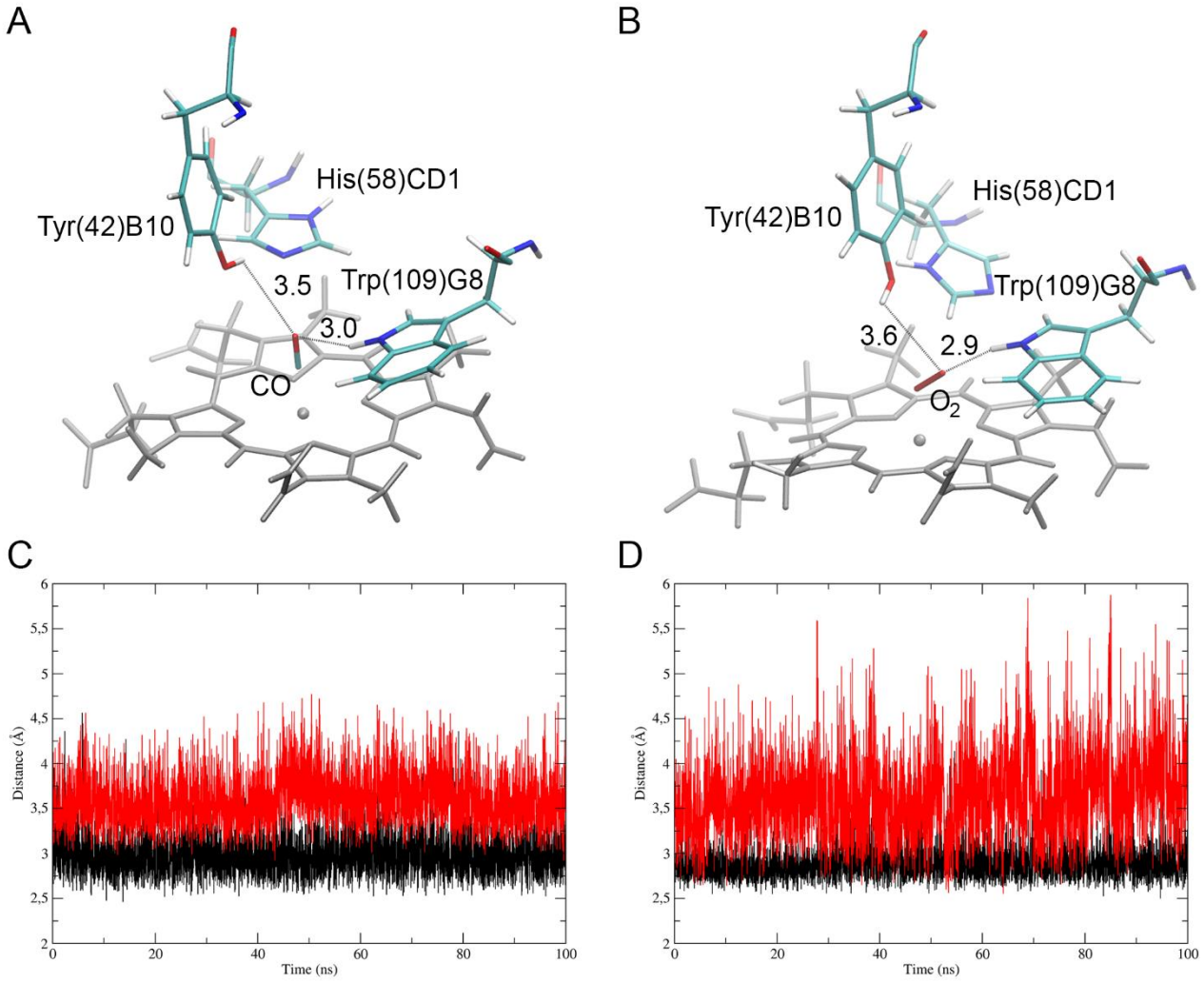
A
His(58)CD1
Tyr(42)B10
3.5
Trp(109)G8
3.0
CO
B
Tyr(42)B10
His(58)CD1
Trp(109)G8
3.6
2.9
O2
C
Distance (Å)
Time (ns)
D
Distance (Å)
Time (ns)

Fig. 8

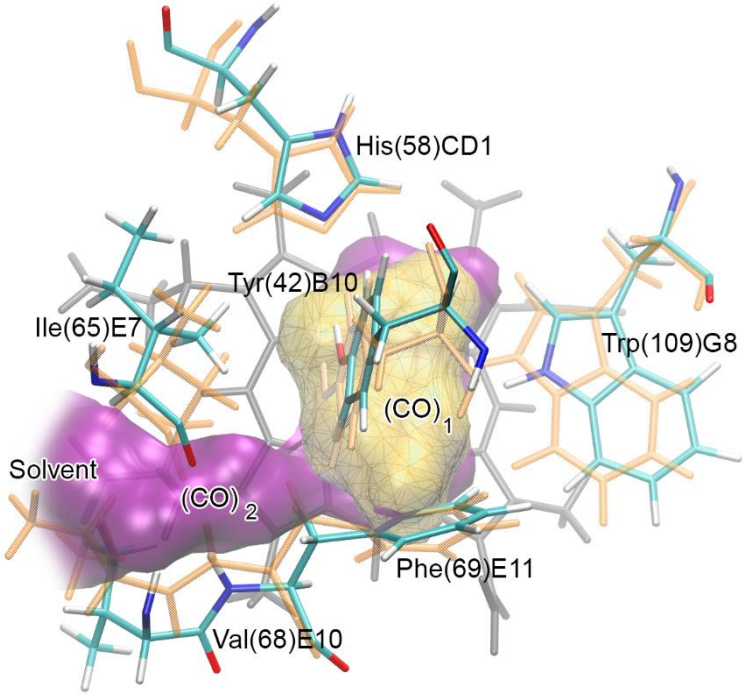
His(58)CD1
Tyr(42)B10
Ile(65)E7
Trp(109)G8
(CO)1
Solvent
(CO)2
Phe(69)E11
Val(68)E10

Fig. 9