doi:10.1016/j.jmb.2008.12.023 J. Mol. Biol. (2009) 386, 246–260

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

HisE11 and HisF8 Provide Bis-histidyl Heme Hexa-coordination in the Globin Domain of *Geobacter sulfurreducens* Globin-coupled Sensor

Alessandra Pesce[1]†, Liesbet Thijs[2]†, Marco Nardini[3], Filip Desmet[4], Lorenza Sisinni[3], Louise Gourlay[3], Alessandro Bolli[5], Massimiliano Coletta[6], Sabine Van Doorslaer[4], Xuehua Wan[7], Maqsudul Alam[7], Paolo Ascenzi[5], Luc Moens[2], Martino Bolognesi[3] and Sylvia Dewilde[2]*

[1]*Department of Physics, CNISM and Center for Excellence in Biomedical Research, University of Genova, Via Dodecaneso, 33, I-16146 Genova, Italy*

[2]*Department of Biomedical Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium*

[3]*Department of Biomolecular Sciences and Biotechnology and CIMAINA, University of Milano, Via Celoria 26, I-20133-Milano, Italy*

[4]*Department of Physics, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium*

[5]*Department of Biology and Interdepartmental Laboratory for Electron Microscopy, University Roma Tre, Viale Guglielmo Marconi 446, I-00146 Roma, Italy*

Among heme-based sensors, recent phylogenomic and sequence analyses have identified 34 globin coupled sensors (GCS), to which an aerotactic or gene-regulating function has been tentatively ascribed. Here, the structural and biochemical characterization of the globin domain of the GCS from *Geobacter sulfurreducens* (*Gs*GCS^{162}) is reported. A combination of X-ray crystallography (crystal structure at 1.5 Å resolution), UV-vis and resonance Raman spectroscopy reveals the ferric *Gs*GCS^{162} as an example of bis-histidyl hexa-coordinated GCS. In contrast to the known hexa-coordinated globins, the distal heme-coordination in ferric *Gs*GCS^{162} is provided by a His residue unexpectedly located at the E11 topological site. Furthermore, UV-vis and resonance Raman spectroscopy indicated that ferrous deoxygenated *Gs*GCS^{162} is a penta-/hexa-coordinated mixture, and the heme hexa-to-penta-coordination transition does not represent a rate-limiting step for carbonylation kinetics. Lastly, electron paramagnetic resonance indicates that ferrous nitrosylated *Gs*GCS^{162} is a penta-coordinated species, where the proximal HisF8-Fe bond is severed.

**Corresponding author.* E-mail address: sylvia.dewilde@ua.ac.be.

† A.P. and L.T. contributed equally to this work.

Present address: L. Sisinni, Department of Chemical Sciences, Via Marzolo 1, I-35131 Padova, Italy.

Abbreviations used: GCS, globin-coupled sensor; GsGCS^{162}, globin domain of GCS from *Geobacter sulfurreducens*; PAS, an acronym formed from the names of the first three proteins (the Period clock protein of *Drosophila*, the Aryl hydrocarbon receptor nuclear translocator of vertebrates, and the Single-minded protein of *Drosophila*) recognized as sharing such sensor motif; CooA, transcriptional activator of coo operons encoding proteins required to metabolize CO as a sole energy source; HemAT, heme-based aerotaxis transducers; *Bs*HemAT, HemAT from *Bacillus subtilis*; Pgb, protoglobin; *Gs*GCS, globin-coupled sensor from *Geobacter sulfurreducens*; EPR, electron paramagnetic resonance; RR, resonance Raman; met-*Gs*GCS^{162}, ferric form of *Gs*GCS^{162}; MAD, multiple wavelength anomalous dispersion; CW-EPR, continuous-wave EPR; r.m.s., root-mean-square; swMb, sperm whale myoglobin; *Ma*Pgb, protoglobin from *Methanosarcina acetivorans* C2A; Hb, hemoglobin; LS, low spin; HS, high spin; *Av*GReg, globin coupled sensor of *Azotobacter vinelandii*.

[6]*Department of Experimental Medicine and Biological Sciences, University of Roma "Tor Vergata", Via Montpellier 1, I-00133 Roma, Italy*

[7]*Department of Microbiology, Snyder Hall 207, 2538 The Mall, University of Hawaii, Honolulu, HI 96822, USA*

Received 29 September 2008; received in revised form 5 December 2008; accepted 9 December 2008
Available online 16 December 2008

Edited by R. Huber

Keywords: globin coupled sensor; 3D structure; ligand binding properties; hexacoordination; *Geobacter sulfurreducens*

Introduction

Heme-based sensor proteins regulate adaptive responses to fluctuations of gaseous physiological messengers (i.e., O_2, CO, and NO) by coupling in the same protein a regulatory heme-binding domain (or independent subunit) to a neighbouring transmitter region, often endowed with enzymatic functions. The known heme-based sensors feature four different types of heme-binding modules: the heme-binding PAS domain, the modified globin domain, the CO-sensing transcription factor CooA, and the heme-NO-binding domain of guanylyl cyclise.[1] Recent phylogenomic and sequence analyses have identified 34 heme-based sensors displaying a globin-like heme-binding domain (globin-coupled sensor, or GCS), in approximately 30 bacteria and one archaea, each sensor being tentatively classified either as aerotactic or gene regulating.[2] So far, the best characterized GCS proteins are the heme-based aerotaxis transducers (HemATs), originally discovered in *Halobacterium salinarum* and in *Bacillus subtilis*, which couple the N-terminal sensing globin domain to a C-terminal signalling domain typical of methyl-accepting chemotaxis proteins.[3] Crystal structures of the unliganded and liganded HemAT sensor heme-binding domain from *B. subtilis* (*Bs*HemAT) revealed that its fold is a variant of the canonical globin fold, where an extra N-terminal helix (called the Z-helix) is present and the D-helix is deleted.[4] Apart from the chimeric GCS, several single domain globins have been identified within the GCS lineage,[5] protoglobins (Pgbs) being the best characterized example up to now.[5–7]

The biochemical characterization of a novel GCS identified in the δ-Proteobacterium *Geobacter sulfurreducens* (*Gs*GCS) has been reported recently.[8] Although *G. sulfurreducens* has been considered in the past to be an obligate anaerobe, complete genome sequencing revealed the presence of 111 putative *c*-type cytochromes, and a high proportion of proteins involved in environmental sensing, with evidence for aerobic metabolism.[9] Indeed, *G. sulfurreducens* has been reported to survive after exposure to atmospheric O_2 and to grow in low concentrations of O_2.[10]

In the context of our ongoing studies on the role of GCSs and on the molecular mechanisms regulating heme–ligand binding and globin action, we report here the results of a biochemical, biophysical, and crystallographic investigation on the sensor globin domain (162 amino acids) of the GCS from *G. sulfurreducens* (*Gs*GCS^{162}), the first globin domain coupled to a transmembrane signal-transduction domain (of yet unknown function) so far characterized. The high-resolution (1.5 Å) crystal structure, together with UV-vis and resonance Raman (RR) spectroscopic results, shows that the ferric protein (met-*Gs*GCS^{162}) displays a hexa-coordinated heme, where residue HisE11 is unexpectedly the endogenous sixth axial ligand. Furthermore, we show that the ferrous deoxygenated *Gs*GCS^{162} is a penta-/hexa-coordinated mixture, where the hexa-to-penta-coordination transition does not represent a rate-limiting step for CO binding kinetics. Lastly, electron paramagnetic resonance (EPR) indicates that ferrous nitrosylated *Gs*GCS^{162} is penta-coordinated, due to cleavage of the proximal HisF8-Fe bond.

Results

Crystal structure

The crystal structure of ferric *Gs*GCS^{162} (two molecules per asymmetric unit, referred to as A and B) was solved by multiple wavelength anomalous dispersion (MAD) methods, based on the anomalous signal of the heme Fe atom. Refinement of the structure converged at an *R*-factor value of 16.4% (R_{free} 20.7%), at 1.50 Å resolution, with ideal stereochemical parameters (Table 1). The final model

Table 1. Data collection, phasing (MAD: Fe) and crystallographic refinement statistics for *Gs*GCS[162]

	Native	MAD:Fe		
A. Data collection				
Space group	$P2_12_12_1$	$P2_12_12_1$		
Cell dimensions:				
a, *b*, *c* (Å)	56.6, 73.8, 89.9	56.9, 73.7, 89.7		
α, β, γ (°)	90, 90, 90	90, 90, 90		
		Peak	Inflection	Remote
Wavelength (Å)	0.873	1.739	1.740	0.954
Resolution (Å)	38.4–1.5 (1.58–1.50)	89.4–2.5 (2.64–2.50)	89.4–2.5 (2.64–2.50)	57.0–2.1 (2.21–2.10)
No. reflections	227,555	174,632	175,146	230,854
Unique reflections	60,673	13,209	13,254	22,232
R_{merge}[a]	0.056 (0.238)	0.074 (0.208)	0.074 (0.291)	0.071 (0.316)
$I/\sigma(I)$	14.9 (4.6)	32.2 (10.5)	32.2 (8.2)	25.8 (6.8)
Completeness (%)	99.4 (100.0)	98.2 (97.1)	98.2 (96.8)	98.6 (98.8)
Multiplicity	3.8 (3.8)	13.2 (13.7)	13.2 (13.7)	10.4 (10.3)
B. Refinement				
Resolution (Å)	35.2–1.5			
R_{factor}/R_{free} (%)	16.4/20.7			
No. of residues/protein atoms	304 (2×152, 1–152)/2,631[b]			
No. heme groups	2			
No. glycerol molecules	3			
No. water molecules	378			
B-factors (Å^2):				
Protein	18.5			
Heme group	17.7			
Glycerol	58.1			
Water molecules	35.4			
r.m.s.d. from ideality:				
Bond lengths (Å)	0.016			
Bond angles (°)	1.4			
Ramachandran plot[c]				
Most favored regions (%)	96.5			
Additionally allowed regions (%)	3.5			

Values in parentheses are for the highest-resolution shell.
[a] $R_{merge} = \Sigma_h \Sigma_i | I_{hi} - \langle I_h \rangle | / \Sigma_h \Sigma_i I_{hi}$.
[b] Met(4), Arg(34), Ser(72), Ser(77), Ile(83), Ile(89), His(93), Arg(132), and Ile(145) side-chains, and the heme A- and D-propionates display a double conformation in the *Gs*GCS[162] A chain, while Met(4), Phe(39), Leu(62), His(66), Arg(81), Arg(95), Ile(118), and Arg(132) side-chains, and the heme A-propionate display a double conformation in the *Gs*GCS[162] B chain.
[c] Data produced using the program PROCHECK.[47]

consists of 2×152 amino acids (residues 1–152 in both *Gs*GCS[162] A and B chains; no electron density was observed for residues 153–162), two heme groups (one per chain), three glycerol molecules, and 378 water molecules. A total of 17 residues and the heme propionates display double conformations in *Gs*GCS[162] A and B chains (Table 1).

*Gs*GCS[162] displays a modified globin fold[11,12] consisting of eight helices that, according to the classical globin nomenclature, have been named Z (3–12), A(14–30), B(31–45), C(46–60 in the A chain, 46–54 in the B chain; in both chains residues 52–54 are in the 3_{10} conformation), E(61–73), F(79–97), G(99–122), and H(124–152) (Figs. 1 and 2a). Superposition of 150 Cα atoms of the two independent *Gs*GCS[162] chains (excluding the first and the last residue) yields a root-mean-square deviation of 0.55 Å. Major structural differences are localized at the CE (residues 51–60) and the EF (residues 75–79) interhelical regions. In particular, in the B subunit the C-helix is nine residues long (with the last three residues in the 3_{10} conformation), while in the A subunit the C-helix gains about two C-terminal turns (six residues) after the 3_{10} stretch, thus connecting directly the C- and the E-helices (Fig. 1; Supplementary Data Fig. S1) (for details, see the legend to Fig. 1). When compared to sperm whale Mb (swMb), taken as the prototype globin fold (9% sequence identity with *Gs*GCS[162]), the overall r.m.s.d. is 2.50 Å, calculated over 115 C_α pairs, the best match being between helices A, F, G, and H. The largest structural deviations are localized at the N-terminus, where *Gs*GCS[162] displays the additional helix (Z) at the C-terminus of the B-helix, and in the CD region, where the swMb D-helix is absent from *Gs*GCS[162], and the C-helix is connected directly to the E-helix through a two-turn elongation. Other structural deviations are evident at the E-helix (markedly divergent in the two globins), at the F-helix (which is nine residues longer in *Gs*GCS[162]), in the EF and FG hinges, at the G- and H-helices (both four residues longer in *Gs*GCS[162]), and at the GH hinge (which is reduced to one Pro residue) (Fig. 2a; Supplementary Data Fig. S1).

Quaternary structure

The *Gs*GCS[162] A and B chains assemble in a dimer with a subunit contact interface of 1684 Å^2 (~18.5%

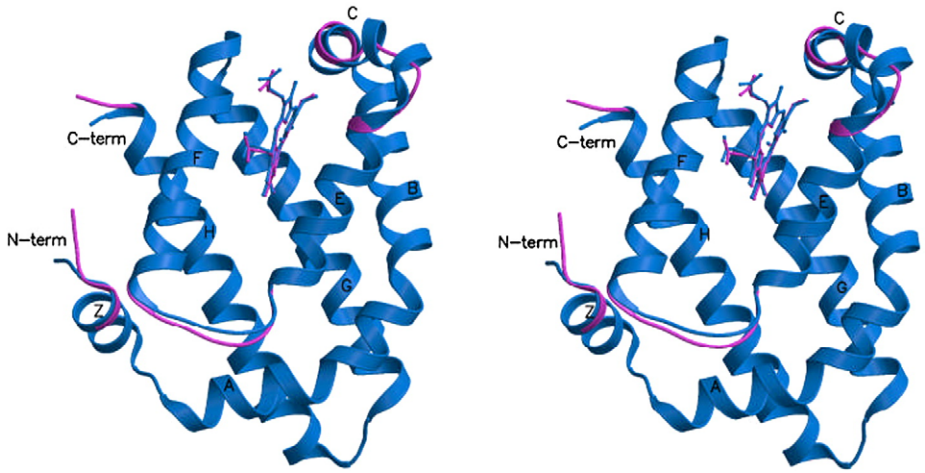

Fig. 1. The *Gs*GCS[162] fold. Stereo view of the structural overlay of the *Gs*GCS[162] A (blue ribbon) and B (violet ribbon) chains. The structural differences located at the CE interhelical region are related to the presence in the B chain of a salt bridge between Glu(51)C5 OE1 and Lys(54) NZ atoms, absent from the A chain. Thus, the C-helix of the A chain extends for two additional turns, stabilized by a hydrogen-bond between Leu(57) O and Gln(60) NE2 atoms. Similarly, structural differences at the EF interhelical region are related to the presence/absence of a salt bridge. In the B chain, the Arg(81) side chain displays two alternative conformations, both with the NH1 atom salt-bridged to the Asp(79) OE2 atom. Such a salt bridge is absent from the A chain, resulting in a more mobile EF-F region, with Ile(83), Ile(89), His(93) side-chains in double conformations.

of the total surface of each subunit). The interface contacts are provided mostly by residues from the G- and H-helices (26 residues), partly from the Z-helix (four residues), and from the BC interhelical region (four residues) of the A and B chains. Thus, the dimerization interface has a four α-helical bundle as the core region (Fig. 3a). Dimer association is determined mostly by polar interactions that include 31 hydrogen bonds, six salt bridges, and 24 water molecules bridging side chains from the facing dimer subunits. Parts of these water molecules are localized in two major cavities (120 Å^3 and 108 Å^3 volume) at the centre of the HG-bundle.

Moreover, gel-filtration chromatography has shown that the protein elutes as a dimer (our unpublished data), suggesting that the dimer is a functional quaternary assembly.

Bis-histidyl hexa-coordination and heme distal site

One of the striking features displayed by the *Gs*GCS[162] structure is the bis-histidyl hexa-coordination of the heme Fe atom observed in its ferric form, which is based on the distal His(66)E11‡ and the proximal His(93)F8 residues (Fig. 4). Such a pair of Fe axial ligands is unprecedented, since the endogenous heme distal ligand in known hexa-coordinated Hbs is regularly HisE7.[11,13] Coordination to the heme Fe by His(66)E11 results in an anticlockwise rotation of about 30° toward the heme of the E-helix, pivoting around residue His(66)E11, relative to the orientation adopted in swMb. As a consequence, the E-helix is almost parallel with the heme and orthogonal to the C-helix (Fig. 2a).

Different azimuthal orientations of the heme-coordinated His(66)E11 imidazole ring were observed in the *Gs*GCS[162] A and B chains. In the A chain, the His(66)E11 imidazole ring is eclipsed relative to the NB-ND pyrrole nitrogen atoms of the heme group (heme Fe to His(66)E11 NE2, a bond length of 2.10 Å) (Fig. 4a). On the contrary, in the B chain, the distal His (66)E11 side chain shows two coexisting orientations (each refined at 0.50 occupancy) (Fig. 4b). One conformation is eclipsed and equivalent to that of the A chain (heme Fe coordination bond length of 2.16 Å). The second conformation shows the His(66) E11 side chain staggered relative to the heme pyrrole nitrogen atoms (heme Fe coordination bond length of 2.10 Å). The B chain His(66)E11 double conformation correlates with the double conformation observed for Phe(39)B9 and Leu(62)E7. When His(66)E11 is staggered, and the Phe(39)B9 aromatic ring is roughly parallel with the His(66)E11 imidazole ring, a water molecule (the "distal water molecule") is hydrogen bonded to the His(66)E11 ND1 atom (2.84 Å). A second water molecule is nearby, and hydrogen bonded to the "distal water molecule" (2.76 Å), and at hydrogen bonding distance from the carbonyl oxygen atoms of Ala(36)B6 (2.95 Å) and Lys(63)E8 (3.06 Å) (Fig. 4b). Both water molecules in the B chain are within an hourglass-shaped distal cavity of about 42 Å^3 in volume.

In the A chain, the corresponding distal cavity is L-shaped and larger (about 85 Å^3), due to the eclipsed orientation of the His(66)E11 imidazole ring and to the correlated positioning of the Phe(39)B9 aromatic ring roughly parallel with the His(66)E11 side chain.

‡Amino acid residues have been labelled using the three-letter codes, the sequence numbering (in parentheses), and the topological site they occupy within the globin fold; topological site identification has been omitted for most interhelical residues.

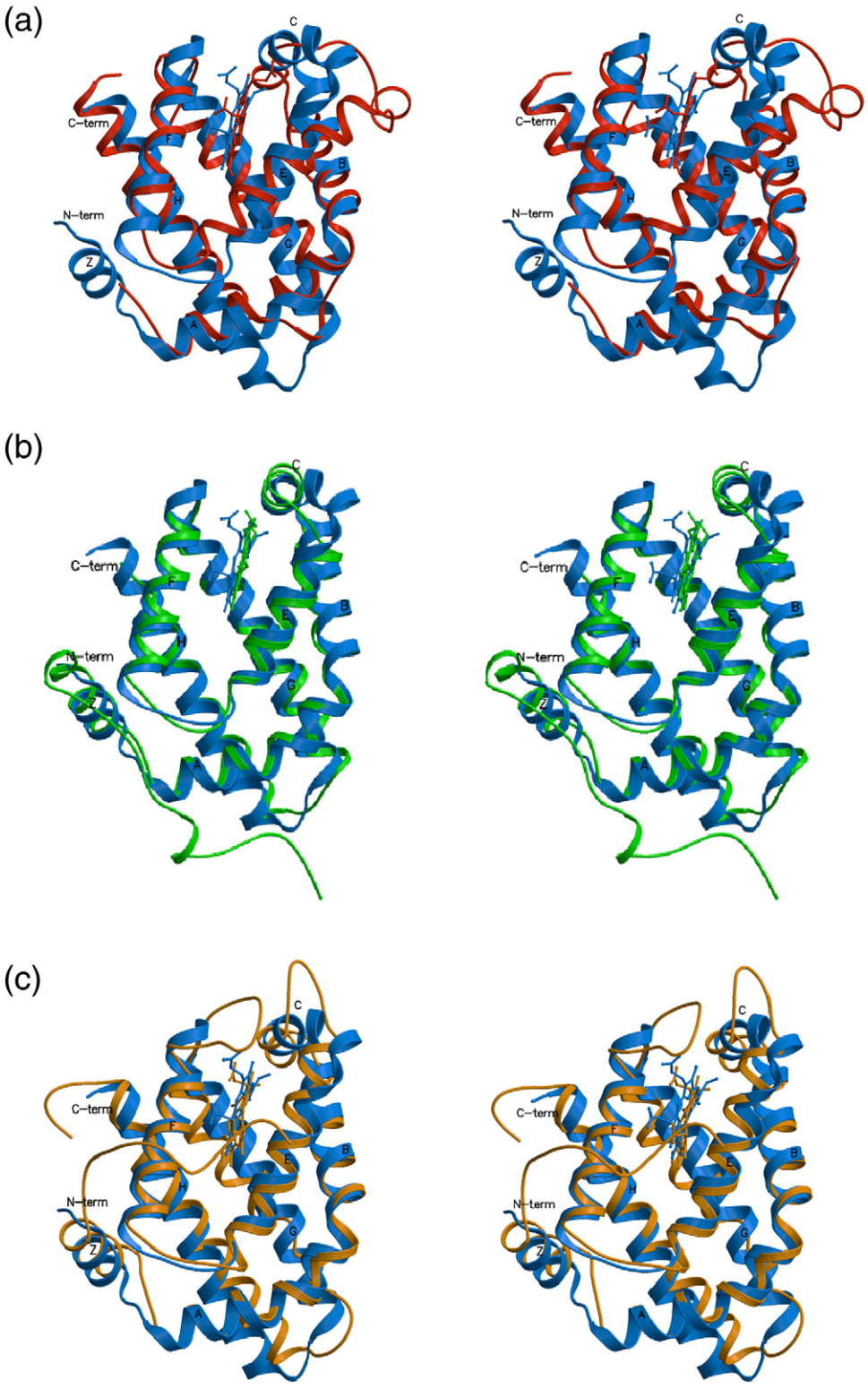

Fig. 2. The *Gs*GCS^{162} fold *versus* structurally related proteins. a, Stereo view of the structural overlay of *Gs*GCS^{162} (blue ribbon) and swMb (red ribbon), taken as a prototype reference 3-on-3 globin. b, Fold overlay of *Gs*GCS^{162} (blue ribbons) on *Bs*HemAT (green ribbons). c, Fold overlay of *Gs*GCS^{162} (blue ribbons) on *Ma*Pgb (orange ribbons). The heme group is shown with the colour code of the corresponding protein. Note that the heme group in *Ma*Pgb is markedly more deeply located within the protein matrix than in GCS proteins.

Despite the increased cavity volume, the eclipsed orientation of His(66)E11 prevents binding of the distal water molecule; only the second water molecule is present in the A chain heme distal site, at hydrogen bonding distance from the carbonyl oxygen atoms of Ala(36)B6 (3.04 Å) and Lys(63)E8 (3.25 Å). In both

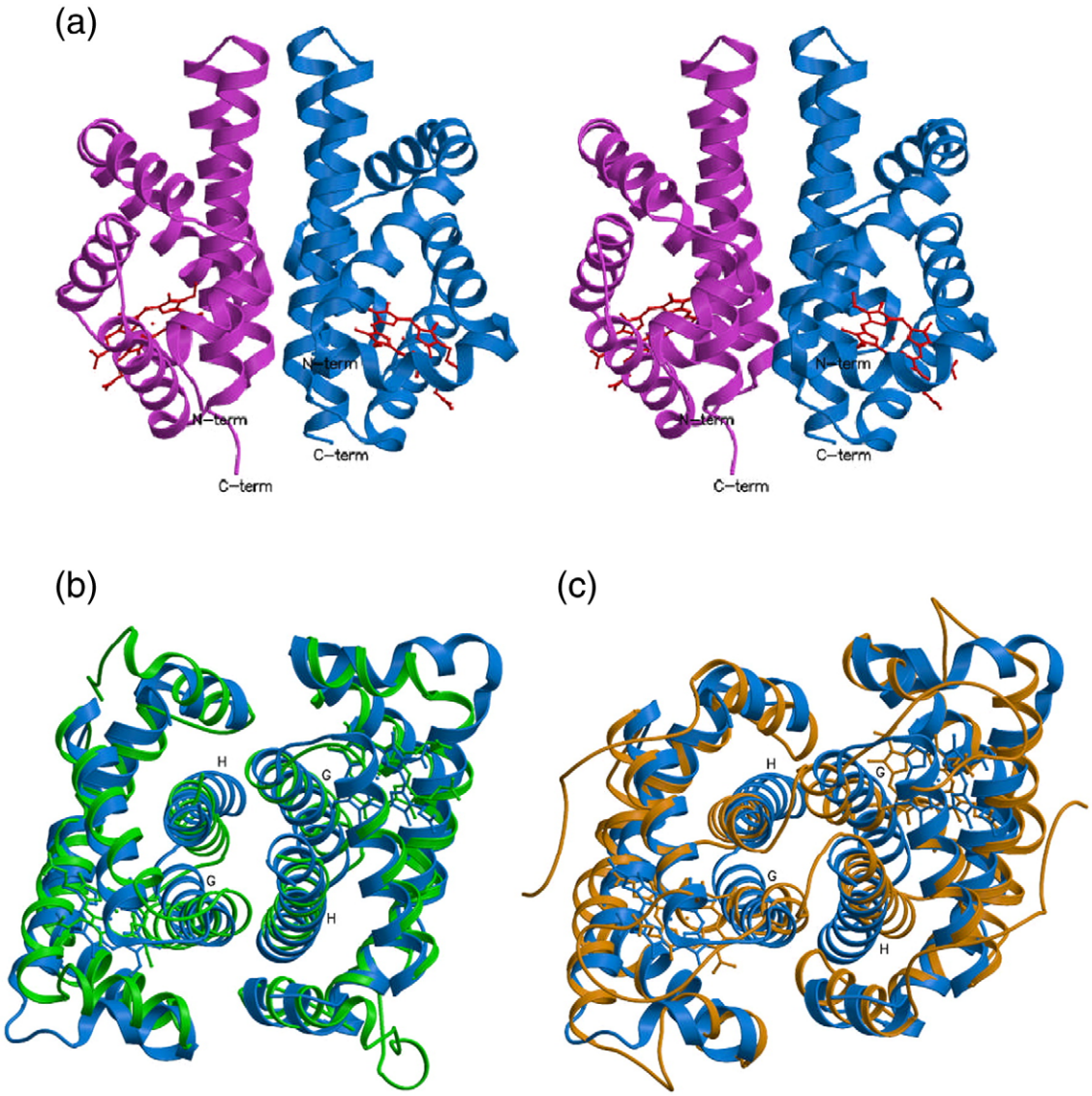

Fig. 3. The *Gs*GCS^{162} dimer. a, Stereo view of the *Gs*GCS^{162} dimer, with the two subunits, interacting mainly through the G- and H-helices (helical bundle) and through the Z-helices, shown as blue and violet ribbon models. *b*, Top view showing the structural overlay of the *Gs*GCS^{162} dimer (blue ribbons) on the *Bs*HemAT dimer (green ribbons). c, Overlay of the *Gs*GCS^{162} dimer (blue ribbons) on the *Ma*Pgb dimer (orange ribbons).

subunits, direct access for water molecules from the solvent region to the heme distal site may be provided by a narrow aperture localized between the B- and E-helices (lined by Ala(36)B6, Asp(37)B7, Lys(63)E8 and Gln(67)E12), where the Lys(63)E8 side chain may act as a polar gate. Moreover, it should be noted that the heme A- and D-propionates are solvent-exposed and salt-linked to residues located at the heme-proximal site in both *Gs*GCS^{162} subunits (see below). As a consequence, the distal site cavity may be partly solvent-accessible from the heme propionate side, due to the presence of small/medium size residues in the His(66)E11 surroundings.

Analysis of the *Gs*GCS^{162} distal site shows that residues Phe(52)CD1 and Tyr(40)B10 do not point towards the inner distal site, a highly unusual structural feature among globins. The Phe(52)CD1 side chain is rotated by about 90° compared to the orientation regularly found in globins, essentially because in *Gs*GCS^{162} Phe(52)CD1 is not part of the CD hinge, but belongs to the extended C-helix (Fig. 4 and Supplementary Data Fig. S1). Similarly, the unique structural organization of the C-helix in the hexa-coordinated *Gs*GCS^{162} promotes a hydrogen bond between Asp(56) OD2 and Tyr(40)B10 OH atoms, forcing the Tyr(40)B10 side chain away from the heme-distal site. In other globins, TyrB10 (when present) is normally involved in heme–ligand stabilization, thus well oriented toward the distal ligand coordination site.[14,15]

Heme stabilization and the heme proximal site

Stabilization of the *Gs*GCS^{162} bound heme is achieved through bis-histidyl heme-Fe coordination, and through electrostatic interaction of the heme

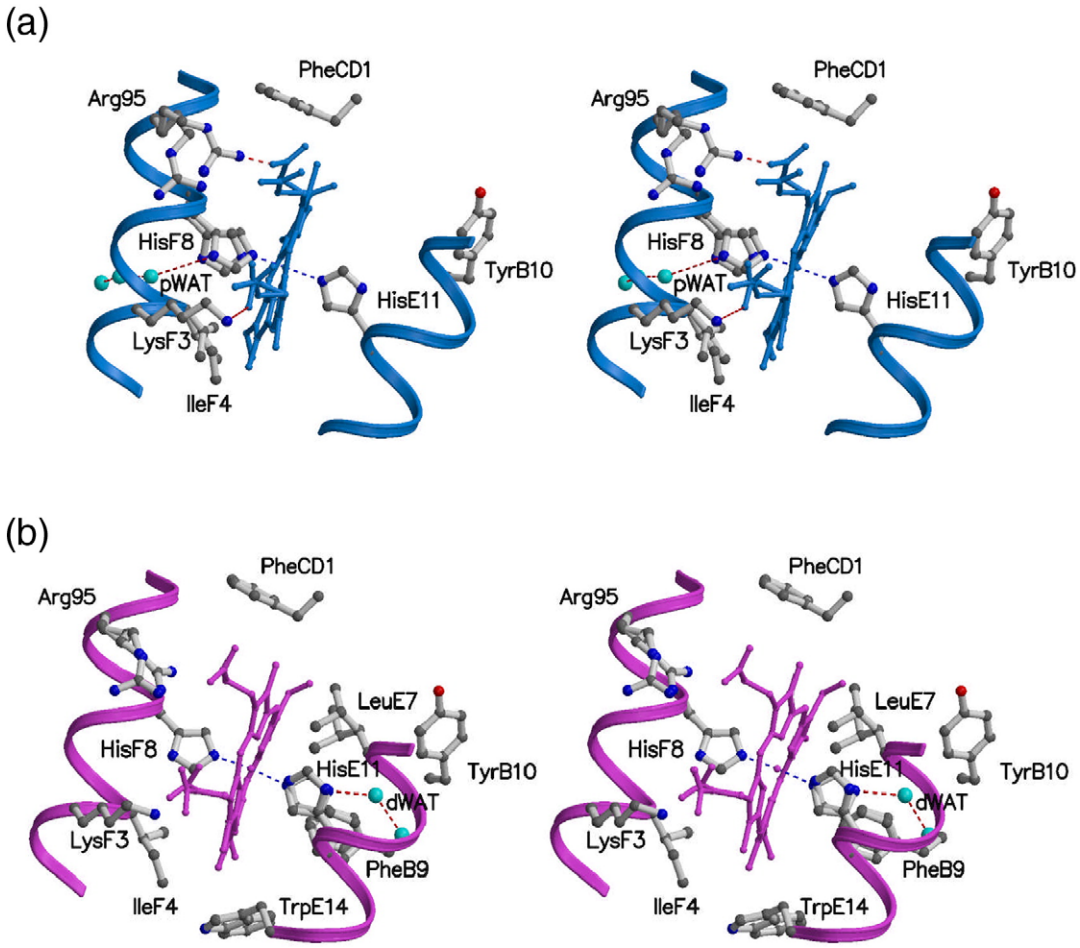

Fig. 4. Structure in the heme pocket of *Gs*GCS[162]. A stereo view of the heme, of the proximal His(93)F8, of the distal His(66)E11 and of the surrounding residues contacting the heme. When appropriate, residue topological positions are indicated. Hydrogen bonds are shown as broken lines and water molecules are shown as cyan spheres. a, In the A subunit, the heme A- and D-propionates and the side chains of Ile(89)F4, His(93)F8, and Arg(95) are shown in two coexisting orientations. The "proximal water molecule" is indicated by pWAT. b, In the B subunit, the heme A-propionate and the side chains of Phe(39)B9, Leu(62)E7, His(66)E11, and Arg(95) are shown in two coexisting orientations. The "distal water molecule" is indicated by dWAT. Note the unusual orientation of residues Phe(52)CD1 and Tyr(40)B10 that point away from the distal site cavity in both subunits.

propionates and van der Waals contacts (<4.0 Å) with 15 and 16 protein residues for the A and B subunits, respectively. In both subunits, two salt bridges to Lys (88)F3 and Arg(95) are provided by the heme A- and D-propionates, respectively (Fig. 4).

The architecture of the heme pocket in the proximal region is dominated by Phe(52)CD1, Ile (89)F4, Ala(92)F7, Val(96), Leu(98), Val(103)G4, Met (146)H17, and Tyr(150)H21, which surround and mostly contact the porphyrin ring. The proximal His (93)F8 residue provides a coordination bond of 1.89 Å and of 2.05 Å to the heme-iron atom in the A and B chains, respectively. In both subunits, the His (93)F8 side chain is staggered relative to the heme pyrrole nitrogen atoms, being rotated by about 24° anticlockwise relative to the heme methinic CHB-CHD atoms direction. Such an orientation is typical of an unstrained imidazole ring that facilitates the heme in-plane location of the iron atom (Fig. 4). Surprisingly, a second conformation of the His(93)F8 side chain (eclipsed along the direction of the heme pyrrole NB and ND atoms) is present in the A chain, the imidazole ring being rotated clockwise by about 65° relative to the other azimuthal orientation (refined to 0.5 occupancy) (Fig. 4a). The His(93)F8 alternate conformation correlates with a double conformation of the Ile89 side chain. The azimuthal orientation of the His(93)F8 imidazole ring is further stabilized, as often in globins,[11] by a hydrogen bond between the histidyl ND1 atom and the Ile(89)F4 O atom (2.90 Å and 2.76 Å in the A and B chain, respectively). In the A chain, the second conformation of the His(93)F8 imidazole ring (eclipsed) breaks the hydrogen bond between the histidyl ND1 atom and the Ile(89)F4 O atom, thus allowing the His(93)F8 ND1 atom to hydrogen bond with the proximal water molecule.

The back of the proximal heme-pocket of both A and B chains is connected to a solvated cavity of about 33 $Å^3$, located at the dimer interface, lined by

Ala(100)G1, Asn(104)G5, Ser(147)H18, Tyr(150)H21, and Arg(151)H22 from both subunits. The proximal water molecule found in the A chain is the inner component of a chain composed of four water molecules (mutually hydrogen bonded) connecting the heme proximal site to the dimer interface cavity described above (Fig. 4a). In the B chain, where the eclipsed orientation of the proximal His(93)F8 is not present, the proximal water molecule is absent.

Structurally related proteins

The fold and the dimeric assembly observed in *Gs*GCS^{162} are similar to those found in *Bs*HemAT (PDB accession codes **1OR4** and **1OR6**)[4] and in *Methanosarcina acetivorans* (*Ma*Pgb; PDB accession codes **2VEB** and **2VEE**)[7] When the *Gs*GCS^{162} C_α backbone is superimposed on *Bs*HemAT (18.5% sequence identity with *Gs*GCS^{162}), the matching is satisfactory for all α-helices (Fig. 2b; Supplementary Data S1), and best when superimposing residues 12–51 and 60–146 of *Gs*GCS^{162} on 42–81 and 90–176 of *Bs*HemAT (r.m.s.d. 1.05–1.18 Å, over 127 C^α pairs). Major structural differences are evident at the N-terminus, where *Gs*GCS^{162} is 30 residues shorter than *Bs*HemAT, and where the Z-helices do not match exactly, at the C-CE region (residues 52–59), which has a greater helical content in *Gs*GCS^{162}, at the GH loop (due to the shorter length of the *Bs*HemAT G- and H-helices), and at the C-terminus of the H-helix (Fig. 2b). Such structural differences have two important implications. On one hand, the different structure of the G-H region results in a rotation of about 20° of the two dimeric subunits in *Gs*GCS^{162} relative to the *Bs*HemAT dimer (Fig. 3b). On the other hand, the *Gs*GCS^{162} *versus* *Bs*HemAT conformational differences at the C-CE regions result in an average shift of the *Gs*GCS^{162} E-helix of about 1.2 Å (calculated at the C^α atom of residue E11) towards the heme. Such a small shift is sufficient to allow bis-histidyl heme hexa-coordination by the distal His(66)E11 residue. Interestingly, the *Gs*GCS^{162} and *Bs*HemAT E-helices are almost parallel with the heme; however, such an orientation results in bis-histidyl coordination of the *Gs*GCS^{162} heme-Fe atom, *versus* penta-coordination of the *Bs*HemAT heme and stabilization of the heme-bound exogenous ligand by TyrB10.[4] A detailed structural analysis of the heme cavity and conformational freedom of TyrB10 in the two proteins is provided as Supplementary Data.

When *Gs*GCS^{162} A and B subunits are superimposed on *Ma*Pgb (16.0% sequence identity with *Gs*GCS^{162}), the matching is good for all α-helices, and best when superimposing residues 3–47, 61–95, 101–117, and 131–151 of *Gs*GCS^{162} on residues 24–68, 88–122, 140–156, and 170–190 of *Ma*Pgb, with r.m.s.d. of 1.67 Å and 1.76 Å, respectively, over 121 C_α pairs (Fig. 2c). A detailed analysis of the observed structural differences characterizing *Gs*GCS^{162} relative to *Ma*Pgb is provided as Supplementary Data.

Because the heme group in *Gs*GCS^{162} displays bis-histidyl hexa-coordination, it is of particular interest to examine the protein structure in relation to other hexa-coordinated hemoglobins (Hbs), all displaying heme Fe distal coordination at topological position E7. While details of the structural comparisons are provided as Supplementary Data, the general conclusions that can be drawn from such comparison are as follows. A good structural overlay of E11-hexa-coordinated *Gs*GCS^{162} on E7-hexa-coordinated globins[13] is usually found for the A-, F-, G- and H-helices, while the B-helix often cannot be superimposed accurately. Large structural mismatches are found in the C–D region, where the C-helix is about two turns longer, the D-helix is absent, and the CD connection is 6–13 residues shorter in *Gs*GCS^{162}. As a consequence, the *Gs*GCS^{162} E-helix, as described above, runs almost parallel with the heme, while usually in E7-hexa-coordinated Hb it is at an angle of 22–26° divergent from the heme plane (Supplementary Data Fig. S2A). Such rotation of the E-helix functions to promote distal E11-hexa-coordination. It is worth noting that, to date, the only Hb structure reported to have a bis-histidyl coordination at a topological position other than E7 is the 2/2 (truncated) Hb from *Synechocystis* sp. (*Ss*2/2Hb), where HisE10 provides the heme Fe distal ligation.[16–18] The short connection (two residues) between the C- and E-helices, and the location of the E-helix relative to the heme and to the F-helix make the E10-hexa-coordinated *Ss*2/2Hb distal environment structurally more similar to that of E11-hexa-coordinated *Gs*GCS^{162} than to that of E7-hexa-coordinated Hbs (Supplementary Data Fig. S2B).[13] Interestingly, as for *Gs*GCS^{162} Tyr(40) B10, *Ss*2/2Hb TyrB10 points outside of the distal cavity in the endogenous hexa-coordinated protein; however, TyrB10 comes into close contact with the distal ligand in the *Ss*2/2Hb cyano-met form, upon swinging the E-helix away from the heme.[18]

UV-visible spectroscopy

The ferric form of *Gs*GCS^{162} has an optical absorption spectrum with a Soret band maximum at 410 nm, and the Q-bands at 567 nm and 536 nm (Supplementary Data Fig. S3), indicative of a bis-histidyl coordination of the heme iron.[19] The deoxygenated ferrous form of *Gs*GCS^{162} has maxima at 426 nm, 560 nm and 533 nm (Supplementary Data Fig. S3), characteristic of a hexa-coordinated low-spin ferrous form.[19,20] Note that the Q-bands of deoxygenated ferrous *Gs*GCS^{162} are less sharp than those of other bis-histidyl hexa-coordinated globins,[19,20] indicating the presence of a fraction of the ferrous penta-coordinated form.

The absorption maxima of *Gs*GCS^{162}-CO (421 nm, 563 nm and 545 nm) (Supplementary Data Fig. S3) are similar to those of the carbonylated derivative of *Bs*HemAT[21] and *Paramecium* 2/2Hb.[22] The maximum of the Soret band of ferrous *Gs*GCS^{162}-NO at pH 8.5 is located at 412 nm, and the Q-bands are found at 537 nm and 570 nm (Supplementary Data Fig. S3), with a shift of the Soret band to 417 nm at pH 7.0 (not shown), all these values being comparable to those reported for most penta-coordinated

ferrous nitrosylated heme-proteins.[23–25] Note that lowering and increasing the pH (i.e., pH <5.5 and >8.3) induces stabilization of the penta-coordinated derivative of ferrous nitrosylated heme-proteins; the maximum stability of the hexa-coordinated derivative of ferrous nitrosylated heme-proteins occurs at pH ~7.[24,25] The X-band (CW-EPR) spectrum of *Gs*GCS^{162}-NO confirmed the penta-coordination of the heme Fe, where the bond between the proximal His residue and the heme Fe is cleaved upon NO binding to the distal side of the heme (as detailed in the Supplementary Data, including Fig. S4).

Resonance Raman spectroscopy

The assignment of the oxidation, spin and coordination marker bands in the RR spectra is based on previous work on Mb.[26] The ν_4, ν_3, and ν_2 bands of ferric *Gs*GCS^{162} are located at 1374 cm^{-1}, 1502 cm^{-1}, and 1579 cm^{-1}, respectively (Fig. 5a, trace a). These values are typical of hexa-coordinated low-spin (LS) ferric heme proteins.[27]

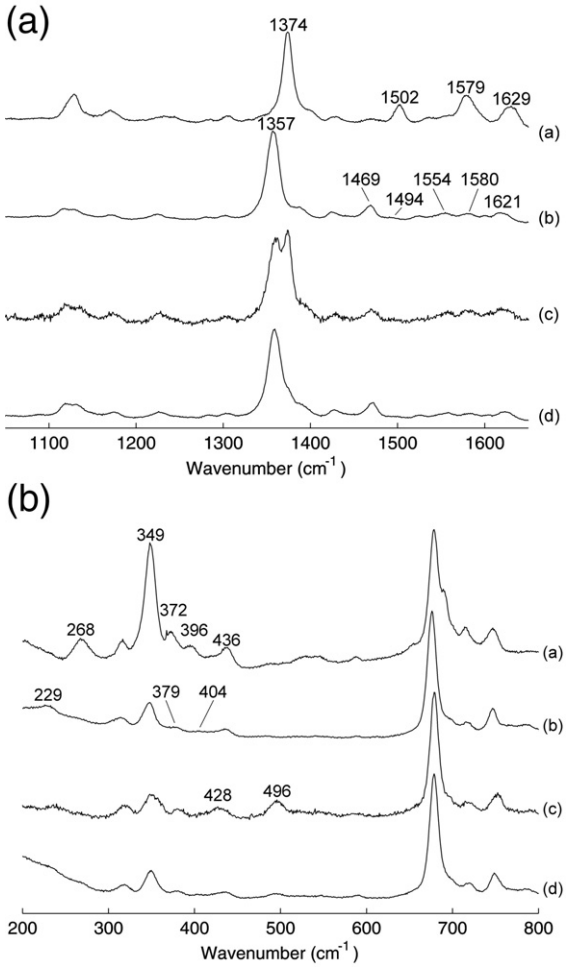

Fig. 5. Resonance Raman spectra. a, High-frequency data; b, low-frequency data for ferric *Gs*GCS^{162} (line a), deoxy ferrous *Gs*GCS^{162} (line b) and ferrous *Gs*GCS^{162}-CO (lines c and d). The laser power was 50 mW (lines a and b), 0.5 mW (line c) and 10 mW (line d), respectively.

The RR spectrum of ferrous deoxygenated *Gs*GCS^{162} has clear marker lines at 1357 cm^{-1} (ν_4), 1469 cm^{-1} (ν_3), and 1554 cm^{-1} (ν_2) (Fig. 5a, trace b), typical for a penta-coordinated high-spin (HS) ferrous state. Furthermore, the marker lines at ~1494 cm^{-1} (ν_3), and 1580 cm^{-1} (ν_2) indicate the presence of a hexa-coordinated LS ferrous species. Since the intensity of the ν_3 line of penta-coordinated species is known to be very high in comparison to that of hexa-coordinated species, the intensity ratio of lines at 1469 cm^{-1} and 1494 cm^{-1} does not imply that the HS species is the dominant form. In fact, the ν_2 lines of the two species are of similar intensity, suggesting that the two fractions are present in similar concentrations. Note that, since the maximum extinction coefficient in the Q-bands is generally higher for the hexa-coordinated LS ferrous form than the penta-coordinated HS ferrous species, the absorption spectrum of ferrous deoxygenated *Gs*GCS^{162} (Supplementary Data Fig. S3) is dominated by the characteristic features of the hexa-coordinated LS form.

At low laser power (<1 mW), two peaks are visible in the ν_4 region of the RR spectrum of ferrous *Gs*GCS^{162}-CO (Fig. 5a, trace c). The relative intensity of these two bands depends on the laser power. At higher laser power, the intensity of the band at 1374 cm^{-1} (CO-ligated species) decreases, while that of the ν_4 band at 1357 cm^{-1} (ferrous deoxy heme) increases (Fig. 5a, trace d), which corresponds to the photolysis of the heme-bound CO.

The propionate and vinyl modes, observable in the RR spectra, reveal details about the stabilization of the heme by the protein surroundings. In the RR spectrum of met-*Gs*GCS^{162}, the band at 372 cm^{-1} is assigned to the heme propionate bending mode $\delta(C_\beta C_c C_d)$ (Fig. 5B, trace a).[26] A frequency of 372 cm^{-1} for this mode has been related to a strong hydrogen bond between the A propionate and surrounding amino acids.[28] In the spectra of ferrous deoxygenated and carbonylated *Gs*GCS^{162} species, the propionate bending mode shifts to 379 cm^{-1} (Fig. 5b, traces b and c). An increase of the frequency of the $\delta(C_\beta C_c C_d)$ mode has been attributed to a stronger interaction between the heme propionate and nearby amino acid residues.[29] Thus, the hydrogen bonds between the heme A propionate and nearby residues for deoxy *Gs*GCS^{162} seem to be stronger than those seen in deoxygenated Mb, and the carbonylated form of *Gs*GCS^{162} has hydrogen bonds on the heme propionate of strength similar to those of Mb-CO, while those of the ferric form of the protein are weaker than those of met-Mb.

The band at 436 cm^{-1} in the RR spectra of ferric and ferrous *Gs*GCS^{162} and the band at 428 cm^{-1} in the *Gs*GCS^{162}-CO are assigned to a vinyl bending mode $\delta(C_\beta C_a C_b)$. The band at 396 cm^{-1} in the spectrum of the ferric form of the protein falls between the frequency regions of the vinyl and the propionate bending modes and cannot be assigned without isotopic labelling of the heme. Although vinyl modes with a comparable low frequency have been observed for isotopically labelled cytochrome *c*

peroxidase[30] and Mb,[26] the difference in frequency between the 396 cm^{-1} and the 436 cm^{-1} bands is, to our knowledge, higher than normally observed for the splitting of the vinyl modes of other heme-proteins.[31] In the RR spectrum of the ferrous form of *Gs*GCS^{162} a band appears at 404 cm^{-1}. This band is assigned to a vinyl bending mode together with the band at 436 cm^{-1}. Similar vinyl modes were observed in Mb and in several variants of Hb M.[31] A change of the vinyl modes between the various forms of *Gs*GCS^{162} indicates a modification of the heme pocket. The fact that the two vinyl bending modes of ferrous deoxygenated *Gs*GCS^{162} have different frequencies suggests that each of the pyrrole rings is affected differently by the out-of-plane distortions of the heme.[22]

In the low-frequency region of the RR spectrum of ferrous *Gs*GCS^{162}-CO, the Fe-CO stretching mode ($\nu_{Fe\text{-}CO}$) is visible at 496 cm^{-1} when low laser power is used (Fig. 5b, trace c). This band disappears when the power is increased and CO is photodissociated (Fig. 5b, trace d). A similar frequency of the Fe-CO mode is indicative of an open conformation of the heme pocket.[21,32] In the RR spectra of some carbonylated heme-proteins, the Fe-C-O bending mode ($\delta_{Fe\text{-}C\text{-}O}$) can be seen near 570 cm^{-1}.[33] The absence of a strong Fe-C-O bending mode is typical of an open conformation of the heme pocket.

*Gs*GCS^{162} carbonylation

Over the whole concentration range of CO explored, the time-course of *Gs*GCS^{162} carbonylation conforms to a single-exponential for more than the 90% of its course. The values of the pseudo first-order rate constant for *Gs*GCS^{162} carbonylation (k_{obsCO}) and of the molar fraction of the carbonylated protein (Y_{CO}) obtained from the kinetic and equilibrium experiments are independent of wavelength at a fixed concentration of CO ([CO]). The analysis of the hyperbolic dependence of Y_{CO} on [CO] according to Eq. (1) (Supplementary Data, Experimental Procedures) allowed the determination of K_{CO}=4.6±0.6 μM (at pH 7.0 and 20 °C); *Gs*GCS^{162} carbonylation appears to be a non-cooperative process (n_{50}=1.01±0.09) (Fig. 6a). The analysis of the linear dependence of k_{obsCO} on [CO] according to Eq. (2) (Supplementary Data, Experimental Procedures) allowed the

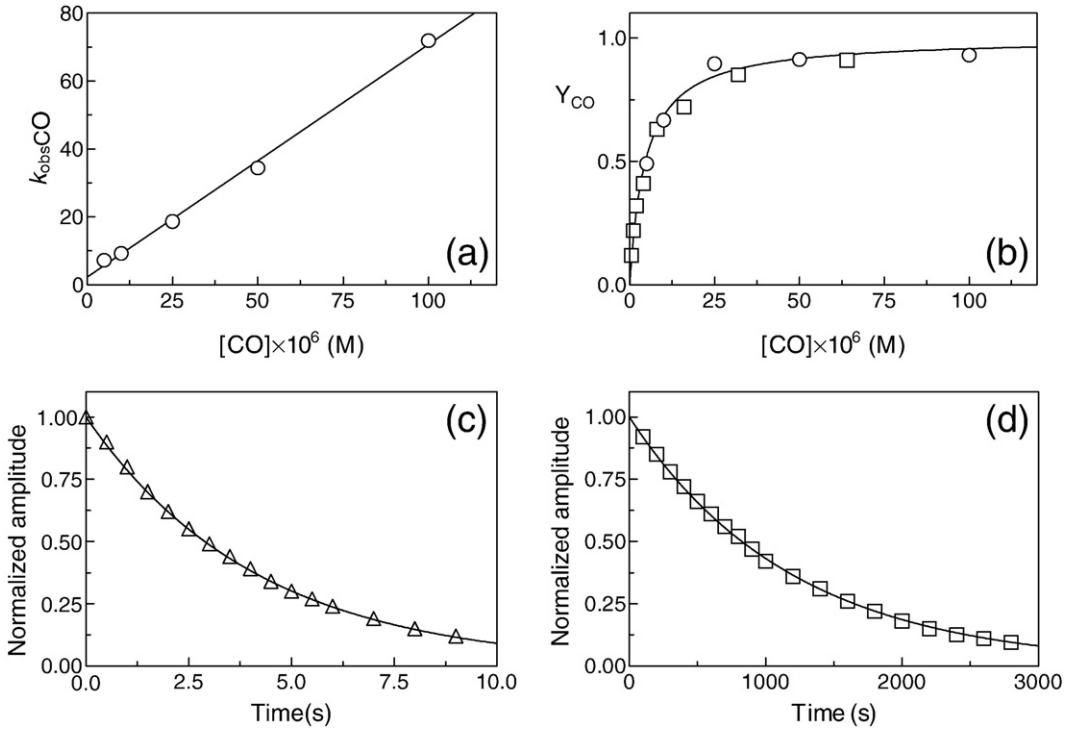

Fig. 6. Kinetics and thermodynamics of *Gs*GCS^{162} (de)carbonylation, deoxygenation, and denitrosylation. a, Dependence of Y_{CO} on [CO] for *Gs*GCS^{162} carbonylation. Values of Y_{CO} were obtained from equilibrium (squares) and kinetic (circles) experiments. Data analysis according to Eq. (1) (see Supplementary Data) allowed the determination of K_{CO}=4.6±0.6 μM. b, Dependence of the pseudo first-order constant for *Gs*GCS^{162} carbonylation (k_{obsCO}) on the CO concentration (i.e., [CO]). Data analysis according to Eq. (2) (see Supplementary Data) allowed the determination of k_{onCO}=6.8±0.8×10^5 $M^{-1}s^{-1}$ and k_{offCO}=2.3±1.0 s^{-1}. c, Normalized averaged time-courses for O_2 dissociation from *Gs*GCS^{162}-O_2. The time-course analysis according to Eq. (4) (see Supplementary Data) allowed the determination of k_{offO2}=2.4×10^1 s^{-1}. Spectra were collected every 5 min. The O_2 and dithionite concentration was 5.0×10^4 M and 1.0×10^2 M respectively. d, Normalized averaged time-courses for NO dissociation from *Gs*GCS^{162}-NO. The time-course analysis according to Eq. (5) (see Supplementary Data) allowed determination of k_{offNO}=8.4×10^4 s^{-1}. Spectra were collected every 5 min. The CO and dithionite concentration was 5.0×10^4 M and 1.0×10^2 M, respectively.

determination of the following kinetic parameters at pH 7.0 and 20 °C (Fig. 6B):

$$k_{onCO} = 6.8 \times 10^5 \pm 0.8 \times 10^5\, M^{-1}s^{-1}$$

$$k_{offCO} = 2.3 \pm 1.0\, s^{-1}$$

The value of k_{offCO} has been confirmed by direct measurements of CO dissociation kinetics by NO replacement (data not shown; see Materials and Methods). According to expectations (Supplementary Data, Experimental Procedures, Scheme 1), the value of K_{CO} calculated from values of k_{onCO} and k_{offCO} (3.4±1.6 μM) is in good agreement, within the uncertainty range, with the value determined experimentally (4.6±0.6 μM). As such, the data obtained from kinetic and equilibrium experiments are in excellent agreement. Further, observations at longer time intervals indicate that there is only a single bimolecular CO binding process and no slower event was observed.

*Gs*GCS^{162} deoxygenation

The O_2 dissociation from *Gs*GCS^{162}-O_2 conforms to a single exponential for more than 95% of its time-course. The wavelength-independent value of the first-order rate constant for *Gs*GCS^{162}-O_2 deoxygenation at pH 7.0 and 20 °C (Fig. 6c) is:

$$k_{offO2} = 2.4 \times 10^{-1} \pm 0.3 \times 10^{-1} s^{-1}$$

*Gs*GCS^{162}-NO denitrosylation

The NO dissociation from *Gs*GCS^{162}-NO conforms to a single-exponential process for more than 90% of its time-course (Fig. 6d). The value of the first-order rate constant for NO dissociation from *Gs*GCS^{162}-NO at pH 8.2 and 20 °C:

$$k_{offNO} = 8.2 \times 10^{-4} \pm 0.7 \times 10^{-4}\, s^{-1}$$

is wavelength- and [CO]-independent in the presence of dithionite excess. In agreement with the spectroscopic properties (Supplementary Data Figs. S3 and S4), the value of k_{offNO} for denistrosylation of *Gs*GCS^{162}-NO is reminiscent of that of penta-coordinated α-chains of ferrous nitrosylated human Hb in the T-state.[34]

Discussion

We have shown that the *Gs*GCS^{162} fold is a variant of the canonical "globin fold", characterized by the presence of an additional N-terminal Z-helix and the total absence of the D-helix. This fold is reminiscent of that of *Bs*HemAT[4] and *Ma*Pgb[7] in terms of tertiary and quaternary structures, with differences at specific sites that make the *Gs*GCS^{162} structure unique among GCS (Figs. 1 and 2). From the quaternary structure viewpoint, *Gs*GCS^{162} has highly symmetrical homodimeric architecture, based on a four-helix central bundle built by the G- and H-helices of both subunits, coupled to interface contacts provided by the Z-helix and the BC inter-helical hinge (Fig. 3a). The involvement of the GCS-specific Z-helix at the dimerization interface and the common dimerization mode found in *Gs*GCS^{162}, *Bs*HemATs,[4] and *Ma*Pgb[7] (Fig. 3), suggests this globin fold variant as a stable scaffold for the formation of GCSs. It is unclear, however, if this dimerization of *Gs*GCS^{162} can be related to signal transduction in response to stimuli as shown for *Bs*HemATs. From the tertiary structure viewpoint, the ferric *Gs*GCS^{162} structure reveals a novel heme-Fe bis-histidyl coordination mode, based on the proximal His(93)F8 and the distal His(66)E11 axial residues (Fig. 4). The unprecedented heme-Fe E11-hexa-coordination results in a significant rotation of the E-helix (relative to the orientation adopted in swMb) that reshapes part of the globin distal site cavity (Fig. 2a). Superposition of the two subunits of the *Gs*GCS^{162} dimer reveals their conformational variability at the CE and EF regions, particularly marked at the carboxy-end of the C-helix (Fig. 1).

The structural variability of the C-helix in the two *Gs*GCS^{162} subunits propagates to side-chain multiple conformations at the distal site level. Indeed, in both subunits, the His(66)E11 imidazole ring displays an eclipsed orientation relative to the NB-ND pyrrole nitrogen atoms of the heme group, but in the B chain an alternative coexisting staggered orientation is present, in turn correlated with alternate side-chain conformations at the B9 and E7 topological sites, and with the presence of two water molecules located in the distal site cavity (one of these hydrogen bonded to the distal His66 (E11)). Both water molecules are absent when His (66)E11 is in the eclipsed conformation (Fig. 4). Direct access of water (or diatomic ligand) molecules from the solvent region to the heme distal site may be provided through a narrow entrance localized between the B- and E-helices, or through a region located between the heme propionates and His(66)E11.

The hexa-coordination found in the met-*Gs*GCS^{162} crystal structure is confirmed by UV-visible and RR experiments, with the hexa-coordinated form being dominant in ferric *Gs*GCS^{162}, whereas a mixture of both hexa-coordinated and penta-coordinated forms is characteristic of the ferrous deoxygenated derivative. The ligand-binding kinetics of ferrous *Gs*GCS^{162} suggests that the equilibrium between the hexa-coordinated non-reactive species and the penta-coordinated reactive form is faster than the pseudo first-order CO ligand binding process. Furthermore, the RR spectra on *Gs*GCS^{162}-CO reveal an open conformation of the heme pocket, indicating a weak or no interaction of distal His(66)E11 with the heme-bound CO. Taking this in to account, it is not likely that the hexacoordination will have a major role in the signal transduction; however, we cannot exclude the possibility. A similar behaviour has been reported for *Bs*HemATs and the globin-coupled

sensor of *Azotobacter vinelandii* (*Av*GReg).[35,36] The affinity of CO for *Gs*GCS^{162} is approximately equal to that of the sensor domain of *Av*GReg,[37] but lower than that of swMb (26) (Supplementary Data Table S1). Since values of k_{onCO} for *Gs*GCS^{162}, *Av*GReg and swMb carbonylation are very similar, the different affinity of CO for *Gs*GCS^{162} and swMb arises from very different values of k_{offCO} (differing by three orders of magnitude). This is confirmed by our data on *Gs*GCS^{162}, where k_{offCO} ($2.3 \pm 1.0\ s^{-1}$) is more than 100-fold higher than that for swMb decarbonylation. The structural bases for such difference and for the faster k_{offCO} in *Gs*GCS^{162} are not yet explored. However, it is known that the cleavage (or the severe weakening) of the heme-Fe proximal coordination bond brings about a dramatic enhancement (by six orders of magnitude) of the CO dissociation rate constant.[38] Therefore, an increase of two orders of magnitude, as observed for *Gs*GCS^{162}-CO with respect to swMb-CO, might be related simply to strain exerted on the proximal coordination bond, in keeping with azimuthal heterogeneity at the His(93)F8 residue observed in *Gs*GCS^{162} A chain (Fig. 4a).

The O_2 dissociation rate constant of GsGCS^{162} is about 100-fold lower than that of mammalian Mbs. In most bacterial globins such slowing of the O_2 dissociation process is achieved through the formation of a H bond, often between TyrB10 and the bound O_2 molecule. In the ferric *Gs*GCS^{162} structure the unique length of the C-helix (Supplementary Data Fig. S2) forces Tyr(40)B10 side-chains to point away from the heme distal site (Fig. 4). However, a cavity is available in the *Gs*GCS^{162} distal site that could help the Tyr(40)B10 side chain relocate near the O_2 binding site upon breaking of the bis-histidyl hexa-coordination. Alternatively, ligand stabilization may be provided by His(66) E11 (which might form a strong H bond with the bound O_2), and by Trp(69)E14, which might cooperate in raising the activation barrier for oxygen dissociation, with a consequent slowing of the dissociation process.

G. sulfurreducens has important potential for applications in bioremediation;[8] it can oxidize organic substrates by using the metal ion Fe(III), which is very abundant at the oxic–anoxic interface, as the electron acceptor. *G. sulfurreducens*, in the past seen as an obligate anaerobic, has been reported to survive after exposure to atmospheric O_2 and to grow in the presence of low concentrations of O_2.[10] The ability to sense O_2 and, consequently, to adapt to changes of O_2 levels may allow *G. sulfurreducens* to thrive in these environments. In this context, the role of *G. sulfurreducens* GCS appears physiologically relevant.

In addition, the differential protein expression of *G. sulfurreducens* under various anaerobic growth conditions revealed that *G. sulfurreducens* GCS is expressed more highly during growth under Fe(III)-reducing conditions than with the alternative electron acceptor fumarate. This suggests that the GCS may be required for optimal Fe(III) reduction.[39]

Materials and Methods

Cloning, expression and purification of recombinant *Gs*GCS^{162}

G. sulfurreducens strain was purchased from the American Type Culture Collection (ATCC-51573). Genomic DNA was isolated with a GNOME DNA isolation kit (QBiogene, Morgan Irvine, CA). The gene fragment encoding the N-terminal sensor domain of *Gs*GCS (codons 1–162; *Gs*GCS^{162}) was amplified by PCR, and cloned into the PCR4Blunt-TOPO vector, and then subcloned into pET3a. Expression of *Gs*GCS^{162} in *Escherichia coli* BL21(DE3)pLysS and its purification from inclusion bodies were as described (see also Supplementary Data).[40]

Crystallization, structure determination and refinement

Crystallization of ferric *Gs*GCS^{162} was achieved using the sitting-drop, vapour-diffusion method, following the results of a wide screen (566 in-house-designed conditions) set up with a robotic apparatus (Genesis RSP100 - Tecan). The protein solution, at 33 mg/ml concentration, was equilibrated against a precipitant solution (10% (w/v) PEG 4000, 0.1 M Mes (pH 6.5), 0.2 M $MgCl_2$) at 277 K. Large single crystals grew within one month and were transferred to 30% (w/v) PEG 4000, 0.1 M Mes (pH 6.5), 0.2 M $MgCl_2$, 15% (v/v) glycerol immediately before cryo-cooling and data collection. The crystals diffracted up to 1.5 Å resolution, using synchrotron radiation (beamline ID23-2, ESRF, Grenoble, France), and belong to the orthorhombic space group $P2_12_12_1$, with unit cell parameters $a = 56.6$ Å, $b = 73.8$ Å, $c = 89.9$ Å, $\alpha = \beta = \gamma = 90°$ (two protein molecules in the asymmetric unit). All data were reduced and scaled using MOSFLM[13] and SCALA[14] programs from the CCP4 suite (Table 1).[41]

The *Gs*GCS^{162} structure was solved by MAD methods, based on the heme-Fe absorption edge. Initial phases were calculated at 2.5 Å resolution using the SOLVE program[42] and further improved by density modification using DM.[43] ARP/wARP was used to automatically trace about 94% of the model residues.[44] Further model building was performed using Coot[45] and REFMAC5 was used for restrained crystallographic refinement.[46] At the end of the refinement stages (including anisotropic *B*-factor refinement), 378 water molecules, and three glycerol molecules were located through inspection of difference Fourier maps. The final *R*-factor was 16.4%, and R_{free} was 20.7% (Table 1). The programs Procheck and Surfnet[47,48] were used to assess stereochemical quality and to explore protein matrix cavities. The program PISA was used to identify quaternary assemblies within the crystal unit cell.[48]

UV-vis spectroscopy

UV-visible absorption spectra (200–800 nm) of ferric and ferrous *Gs*GCS^{162} derivatives were recorded on a Varian Cary-5 UV-VIS-NIR spectrophotometer. Typical protein concentrations varied between 45 μM and 78 μM (pH 8.5 at 20 °C).

Resonance Raman (RR) spectroscopy

RR measurements were done with a Dilor XY-800 Raman scattering spectrometer consisting of a triple 800 mm

spectrograph operating in low-dispersion mode and using a liquid nitrogen-cooled CCD detector. The excitation source was a mixed gas Kr/Ar ion laser (Spectra-Physics BeamLok 2060) operating at 413.1 nm and the spectra were recorded at room temperature. The protein solution was stirred at 500 rpm to avoid local heating and photochemical decomposition in the laser beam. From five to ten spectra (120–240 s recording time each) were acquired to allow the removal of cosmic ray spikes. This was done by eliminating the lowest and highest data points for each frequency value, and averaging the remaining values. Laser power was in the range 0.5–100 mW. Samples with protein concentrations between 45 μM and 78 μM (pH 8.5, at 20 °C) were employed.

Electron paramagnetic resonance spectroscopy

The X-band continuous-wave electron paramagnetic resonance (CW-EPR) experiment was performed on a Bruker ESP380E spectrometer (microwave frequency 9.73 GHz) equipped with a liquid helium cryostat from Oxford Instruments. The spectrum was recorded at 17 K with a microwave power of 0.72 mW, a modulation amplitude of 0.2 mT and a modulation frequency of 100 kHz. The simulation was performed with the EasySpin program.[50]

Kinetics of *Gs*GCS^{162}(-CO) (de)carbonylation

The kinetics of CO binding to *Gs*GCS^{162} was determined by rapidly mixing the ferrous deoxygenated heme protein solution (final concentration 1–2 μM) with the CO-containing solution (final concentration in the range 5 μM–100 μM) in the presence of 10 mg/mL sodium dithionite, at pH 7.0 (0.1 M phosphate buffer) and 20 °C.[23] The kinetics of CO dissociation from *Gs*GCS^{162} was determined by mixing the CO-saturated *Gs*GCS^{162} (final concentration 1–2 μM) with sodium nitrite (final concentration 5 mM) in the presence of 10 mg/mL sodium dithionite, at pH 7.0 (0.1 M phosphate buffer) and 20 °C.[51] The kinetic progress curves were monitored between 390 nm and 500 nm (wavelength interval 3 nm) using a rapid-mixing SX.18MV stopped-flow apparatus equipped with a PDA.1 photodiode array accessory (Applied Photophysics, Salisbury, UK). Absorption spectra were recorded between 10^{-3} s and 1000 s on both a linear and a logarithmic time base.

Equilibrium measurements of *Gs*GCS^{162} carbonylation

Thermodynamics of CO binding to ferrous *Gs*GCS^{162} was obtained at pH 7.0 (0.1 M phosphate buffer) and 20 °C, by adding aliquots of gaseous CO with a precision syringe into the tonometer sealed to a 1 cm spectrophotometric cell (volume 258 mL) containing 3 mL of the heme protein solution (final concentration 1.8 μM), in the presence of 10 mg/mL sodium dithionite, kept under N_2.[51] The absorbance changes accompanying *Gs*GCS^{162} carbonylation were monitored between 360 nm and 460 nm. No longer than 10 min was needed to reach equilibration.

Kinetics of *Gs*GCS^{162}-O_2 deoxygenation

The kinetics of O_2 dissociation from *Gs*GCS^{162}-O_2 was obtained at pH 7.0 (0.1 M phosphate buffer) and 20 °C, by rapidly mixing the deoxygenated heme protein solution (final concentration 1–2 μM), in the presence of 10 mg/mL sodium dithionite, with oxygenated phosphate buffer solution (final concentration 0.1 M).[52] The kinetic progress curves were monitored between 390 nm and 500 nm (wavelength interval 3 nm) using the rapid-mixing SX.18MV stopped-flow apparatus. Absorbance spectra were recorded between 3×10^{-3} s and 50 s on both a linear and a logarithmic time base.

Kinetics of *Gs*GCS^{162}-NO denitrosylation

The kinetics of NO dissociation from *Gs*GCS^{162}-NO was obtained at pH 8.2 (0.04 M tricine buffer) and 20 °C, by mixing the *Gs*GCS^{162}-NO solution (final concentration 3.1 μM) with the CO-dithionite solution (final concentration 100–500 μM and 10 mg/mL, respectively) under anaerobic conditions. No gaseous phase was present. The kinetic progress curves were monitored spectrophotometrically between 360 nm and 460 nm.[53]

Protein Data Bank accession codes:

Atomic coordinates and structure factors have been deposited with the Protein Data Bank[49] with entry codes **2w31** and **r2w31sf**, respectively.

Acknowledgements

The authors thank Kirsten Mees (University of Antwerp, Belgium) for help with the expression/purification of the recombinant proteins, and Dr Chiara Ciaccio (University of Roma Tor Vergata, Roma, Italy) for data analysis. This work was supported by grants from the Ministry for University and Scientific Research of Italy (FIRB Project 'Biologia Strutturale' to M.B., 'CLAR 2007' to P.A., and '60%' to M.C.), from the Ministry for Health of Italy ('Ricerca Corrente 2006' to P.A.), and from the Fund for Scientific Research-Flanders (FWO; to L.M.). S.D. is a post doctoral fellow of the Fund of Scientific Research Flanders (FWO). F.D. acknowledges the University of Antwerp (BOF funds) for PhD funding.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.12.023

References

1. Gilles-Gonzalez, M. A. & Gonzalez, G. (2005). Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. *J. Inorg. Biochem.* **99**, 1–22.
2. Freitas, T. A., Hou, S. & Alam, M. (2003). The diversity of globin-coupled sensors. *FEBS Lett.* **552**, 99–104.
3. Hou, S., Freitas, T., Larsen, R. W., Piatibratov, M., Sivozhelezov, V., Yamamoto, A. *et al.* (2001). Globin-coupled sensors: a class of heme-containing sensors in Archaea and Bacteria. *Proc. Natl Acad. Sci. USA*, **98**, 9353–9358.

4. Zhang, W. & Phillips, G. N., Jr (2003). Structure of the oxygen sensor in Bacillus subtilis: signal transduction of chemotaxis by control of symmetry. *Structure (Camb.)*, **11**, 1097–1110.
5. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Mizuguchi, K., Dewilde, S., Moens, L. & Vanfleteren, J. R. (2007). A model of globin evolution. *Gene*, **398**, 132–142.
6. Freitas, T. A., Hou, S., Dioum, E. M., Saito, J. A., Newhouse, J., Gonzalez, G. *et al.* (2004). Ancestral hemoglobins in Archaea. *Proc. Natl Acad. Sci. USA*, **101**, 6675–6680.
7. Nardini, M., Pesce, A., Thijs, L., Saito, J. A., Dewilde, S., Alam, M. *et al.* (2008). Archaeal protoglobin structure indicates new ligand diffusion paths and modulation of haem-reactivity. *EMBO Rep.* **9**, 157–163.
8. Lovley, D. R. (2003). Cleaning up with genomics: applying molecular biology to bioremediation. *Nature Rev. Microbiol.* **1**, 35–44.
9. Methe, B. A., Nelson, K. E., Eisen, J. A., Paulsen, I. T., Nelson, W., Heidelberg, J. F. *et al.* (2003). Genome of Geobacter sulfurreducens: metal reduction in subsurface environments. *Science*, **302**, 1967–1969.
10. Lin, W. C., Coppi, M. V. & Lovley, D. R. (2004). Geobacter sulfurreducens can grow with oxygen as a terminal electron acceptor. *Appl. Environ. Microbiol.* **70**, 2525–2528.
11. Bolognesi, M., Bordo, D., Rizzi, M., Tarricone, C. & Ascenzi, P. (1997). Nonvertebrate hemoglobins: structural bases for reactivity. *Progr. Biophys. Mol. Biol.* **68**, 29–68.
12. Holm, L. & Sander, C. (1993). Structural alignment of globins, phycocyanins and colicin A. *FEBS Lett.* **315**, 301–306.
13. de Sanctis, D., Pesce, A., Nardini, M., Bolognesi, M., Bocedi, A. & Ascenzi, P. (2004). Structure-function relationships in the growing hexa-coordinate hemoglobin sub-family. *IUBMB Life*, **56**, 643–651.
14. Vuletich, D. A. & Lecomte, J. T. (2006). A phylogenetic and structural analysis of truncated hemoglobins. *J. Mol. Evol.* **62**, 196–210.
15. Nardini, M., Pesce, A., Milani, M. & Bolognesi, M. (2007). Protein fold and structure in the truncated (2/2) globin family. *Gene*, **398**, 2–11.
16. Scott, N. L., Falzone, C. J., Vuletich, D. A., Zhao, J., Bryant, D. A. & Lecomte, J. T. (2002). Truncated hemoglobin from the cyanobacterium Synechococcus sp. PCC 7002: evidence for hexacoordination and covalent adduct formation in the ferric recombinant protein. *Biochemistry*, **41**, 6902–6910.
17. Hoy, J. A., Kundu, S., Trent, J. T., III, Ramaswamy, S. & Hargrove, M. S. (2004). The crystal structure of Synechocystis hemoglobin with a covalent heme linkage. *J. Biol. Chem.* **279**, 16535–16542.
18. Trent, J. T., III, Kundu, S., Hoy, J. A. & Hargrove, M. S. (2004). Crystallographic analysis of synechocystis cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. *J. Mol. Biol.* **341**, 1097–1108.
19. Dewilde, S., Kiger, L., Burmester, T., Hankeln, T., Baudin-Creuza, V., Aerts, T. *et al.* (2001). Biochemical characterization and ligand binding properties of neuroglobin, a novel member of the globin family. *J. Biol. Chem.* **276**, 38949–38955.
20. Couture, M., Das, T. K., Lee, H. C., Peisach, J., Rousseau, D. L., Wittenberg, B. A. *et al.* (1999). Chlamydomonas chloroplast ferrous hemoglobin. Heme pocket structure and reactions with ligands. *J. Biol. Chem.* **274**, 6898–6910.
21. Aono, S., Kato, T., Matsuki, M., Nakajima, H., Ohta, T., Uchida, T. & Kitagawa, T. (2002). Resonance Raman and ligand binding studies of the oxygen-sensing signal transducer protein HemAT from Bacillus subtilis. *J. Biol. Chem.* **277**, 13528–13538.
22. Das, T. K., Weber, R. E., Dewilde, S., Wittenberg, J. B., Wittenberg, B. A., Yamauchi, K. *et al.* (2000). Ligand binding in the ferric and ferrous states of Paramecium hemoglobin. *Biochemistry*, **39**, 14330–14340.
23. Antonini, E. & Brunori, M. (1971). *Hemoglobin and Myoglobin in their Reaction with Ligands*, North Holland Publishing, Amsterdam.
24. Coletta, M., Boffi, A., Ascenzi, P., Brunori, M. & Chiancone, E. (1990). A novel mechanism of heme-heme interaction in the homodimeric hemoglobin from Scapharca inaequivalvis as manifested upon cleavage of the proximal Fe-N epsilon bond at low pH. *J. Biol. Chem.* **265**, 4828–4830.
25. Perutz, M. F. (1979). Regulation of oxygen affinity of hemoglobin: influence of structure of the globin on the heme iron. *Annu. Rev. Biochem.* **48**, 327–386.
26. Hu, S. Z., Smith, K. & Spiro, T. G. (1996). Assignment of protoheme resonance Raman spectrum by heme labeling in myoglobin. *J. Am. Chem. Soc.* **118**, 12638–12646.
27. Lou, B. S., Snyder, J. K., Marshall, P., Wang, J. S., Wu, G., Kulmacz, R. J. *et al.* (2000). Resonance Raman studies indicate a unique heme active site in prostaglandin H synthase. *Biochemistry*, **39**, 12424–12434.
28. Cerda-Colon, J. F., 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.
29. Uchida, T., Sato, E., Sato, A., Sagami, I., Shimizu, T. & Kitagawa, T. (2005). CO-dependent activity-controlling mechanism of heme-containing CO-sensor protein, neuronal PAS domain protein 2. *J. Biol. Chem.* **280**, 21358–21368.
30. Smulevich, G., Hu, S. Z., Rodgers, K. R., Goodin, D. B., Smith, K. M. & Spiro, T. G. (1996). Heme-protein interactions in cytochrome c peroxidase revealed by site-directed mutagenesis and resonance Raman spectra of isotopically labeled hemes. *Biospectroscopy*, **2**, 365–376.
31. Jin, Y., Nagai, M., Nagai, Y., Nagatomo, S. & Kitagawa, T. (2004). Heme structures of five variants of hemoglobin M probed by resonance Raman spectroscopy. *Biochemistry*, **43**, 8517–8527.
32. Dewilde, S., Ebner, B., Vinck, E., Gilany, K., Hankeln, T., Burmester, T. *et al.* (2006). The nerve hemoglobin of the bivalve mollusc Spisula solidissima: molecular cloning, ligand binding studies, and phylogenetic analysis. *J. Biol. Chem.* **281**, 5364–5372.
33. Spiro, T. G. & Wasbotten, I. H. (2005). CO as a vibrational probe of heme protein active sites. *J. Inorg. Biochem.* **99**, 34–44.
34. Herold, S. & Rock, G. (2005). Mechanistic studies of the oxygen-mediated oxidation of nitrosylhemoglobin. *Biochemistry*, **44**, 6223–6231.
35. Thijs, L., Vinck, E., Bolli, A., Trandafir, F., Wan, X., Hoogewijs, D. *et al.* (2007). Characterization of a globin-coupled oxygen sensor with a gene-regulating function. *J. Biol. Chem.* **282**, 37325–37340.
36. Zhang, W., Olson, J. S. & Phillips, G. N., Jr (2005). Biophysical and kinetic characterization of HemAT, an aerotaxis receptor from Bacillus subtilis. *Biophys. J.* **88**, 2801–2814.
37. Ascenzi, P., Bocedi, A., Leoni, L., Visca, P., Zennaro, E., Milani, M. & Bolognesi, M. (2004). CO sniffing through heme-based sensor proteins. *IUBMB Life*, **56**, 309–315.

38. White, D. K., Cannon, J. B. & Traylor, T. G. (1979). A kinetic model for R- and T-state haemoglobin. *J. Am. Chem. Soc.* **101**, 2443–2454.
39. Ding, Y. H., Hixson, K. K., Giometti, C. S., Stanley, A., Esteve-Nunez, A., Khare, T. *et al.* (2006). The proteome of dissimilatory metal-reducing microorganism Geobacter sulfurreducens under various growth conditions. *Biochim. Biophys. Acta*, **1764**, 1198–1206.
40. Dewilde, S., Mees, K., Kiger, L., Lechauve, C., Marden, M. C., Pesce, A. *et al.* (2008). Expression, purification, and crystallization of neuro- and cytoglobin. *Methods Enzymol.* **436**, 341–357.
41. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
42. Terwilliger, T. C. & Berendzen, J. (1999). Discrimination of solvent from protein regions in native Fouriers as a means of evaluating heavy-atom solutions in the MIR and MAD methods. *Acta Crystallogr. D*, **55**, 501–505.
43. Cowtan, K. (1994). Joint CCP4 and ESF-EACBM newsletter on protein crystallography. 31, 34–38.
44. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.
45. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.
47. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
48. Krissinel, E. & Henrick, K. (2005). *CompLife 2005, LNBI 3695*, pp. 163–174, Springer-Verlag Berlin, Heidelberg.
49. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
50. Stoll, S. & Schweiger, A. (2006). EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42–55.
51. Amiconi, G., Antonini, E., Brunori, M., Formaneck, H. & Huber, R. (1972). Functional properties of native and reconstituted hemoglobins from Chironomus thummi thummi. *Eur J. Biochem.* **31**, 52–58.
52. Gibson, Q. H., Olson, J. S., McKinnie, R. E. & Rohlfs, R. J. (1986). A kinetic description of ligand binding to sperm whale myoglobin. *J. Biol. Chem.* **261**, 10228–10239.
53. Moore, E. G. & Gibson, Q. H. (1976). Cooperativity in the dissociation of nitric oxide from hemoglobin. *J. Biol. Chem.* **251**, 2788–2794.