2.3 Å X-ray Structure of the Heme-Bound GAF Domain of Sensory Histidine Kinase DosT of *Mycobacterium tuberculosis*[†,‡]

Larissa M. Podust, Alexandra Ioanoviciu, and Paul R. Ortiz de Montellano*

Department of Pharmaceutical Chemistry, University of California, 600 16th Street, San Francisco, California 94158-2517

Received June 30, 2008; Revised Manuscript Received August 25, 2008

ABSTRACT: *Mycobacterium tuberculosis* responds to changes in environmental conditions through a two-component signaling system that detects reduced O_2 tension and NO and CO exposures via the heme-binding GAF domains of two sensory histidine kinases, DosT and DevS, and the transcriptional regulator DosR. We report the first X-ray structure of the DosT heme-bound GAF domain (GAF_{DosT}) in both oxy and deoxy forms determined to a resolution of 2.3 Å. In GAF_{DosT}, heme binds in an orientation orthogonal to that in the PAS domains via a highly conserved motif, including invariant H147 as a proximal heme axial ligand. On the distal side, invariant Y169 forms stacking interactions with the heme with its long axis parallel and the plane of the ring orthogonal to the heme plane. In one of the two protein monomers in an asymmetric unit, O_2 binds as a second axial ligand to the heme iron and is stabilized via a H-bond to the OH group of Y169. The structure reveals two small tunnel-connected cavities and a pore on the protein surface that suggest a potential route for the access of O_2 to the sensing pocket. The limited conformational differences observed between differently heme iron-ligated GAF_{DosT} monomers in the asymmetric unit may result from crystal lattice limitations since atmospheric oxygen binding likely occurs in the crystal as a result of X-ray-induced Fe^{3+} photoreduction during diffraction data collection. Determination of the GAF_{DosT} structure sets up a framework in which to address ligand recognition, discrimination, and signal propagation schemes in the heme-based GAF domains of biological sensors.

Mycobacterium tuberculosis can survive for decades in a dormant state in hypoxic granulomas in the lung, emerging periodically to cause recurrent infections (*1*). *M. tuberculosis* enters latency in response to hypoxia or exposure to physiological concentrations of NO and/or CO via induction of the dormancy regulon Dos (*2*−*6*) [originally designated Dev (*7*)], comprised of the signaling histidine kinases DevS (also called DosS) and DosT, and the cognate response regulator DosR (also called DevR) (*8*−*10*). DevS and DosT each contain two tandem GAF[1] domains whose N-terminal heme-binding sensory elements (*11*−*14*) directly bind O_2, NO, and CO (*12*−*14*). NO or CO binding to or O_2 dissociation from the N-terminal GAF domain initiates autophosphorylation by the C-terminal protein-histidine kinase domain, with subsequent phosphoryl transfer to the aspartate residue of the transcriptional regulator DosR (*9, 10*).

The oxidation and ligation states of the heme iron modulate DevS and DosT autokinase activity. Ferrous DevS and deoxy and CO- or NO-bound DosT show significant increases in autokinase activity compared to ferric DevS and oxy DosT, respectively, which led to the conclusion that DevS functions as a redox sensor whereas DosT functions as a hypoxia sensor (*14*). However, DosT binds NO 5000 times more tightly than O_2 and 190 times more tightly than CO (*13*), suggesting that one specific function for this sensor may be to detect low concentrations of NO against a high background concentration of O_2. NO is generated by activated macrophages as an immune response factor which exhibits antimycobacterial properties and can kill tubercle bacilli in vitro (*15*). On the other hand, NO acts as a potent reversible inhibitor of aerobic respiration in mitochondria and in bacteria (*16*) and is a versatile signaling agent in eukaryotic systems (*17*). It has been demonstrated that NO and CO produced by the activated host immune system are "hijacked" by *M. tuberculosis* to perform a host-to-pathogen signaling function (*4*−*6*), thus disclosing the level of immune activation at the early stages of infection. O_2 and low concentrations of NO and/or CO mediate opposing effects on gene expression, as NO and CO induce and O_2 represses Dos regulon expression (*4*). Given its prominently high affinity for NO, DosT would be capable of mediating an early *M. tuberculosis* response on this host immune factor, even though the oxygen tension may still be normal or close to normal.

GAF and PAS domains are ubiquitous cytoplasmic signaling units in prokaryotic and eukaryotic organisms (*18, 19*). They have a ligand binding pocket that accommodates a variety of small-molecule ligands, including flavin, adenine, and guanine, as well as linear tetrapyrrole chromophores

[†] This work was supported by NIH RO1 Grants AI74824 (P.R.O.d.M.) and GM078553 (L.M.P.).

[‡] Coordinates have been deposited in the Protein Data Bank as entry 2VZW.

* To whom correspondence should be addressed: Department of Pharmaceutical Chemistry, University of California, 600 16th St., N572D, San Francisco, CA 94158. Telephone: (415) 476-2903. Fax: (415) 502-4728. E-mail: ortiz@cgl.ucsf.edu.

[1] Abbreviations: GAF_{DosT}, N-terminal GAF sensory domain of DosT; GAF domain, sensory domain named for its association with cGMP-regulated cyclic nucleotide phosphodiesterases, adenylate cyclases, and the bacterial transcriptional regulator FhlA; PAS domain, sensory domain named for its association with the Per, ARNT, and Sim proteins.

10.1021/bi8012356 CCC: $40.75
Published on Web 11/04/2008

(bilins) (*20*) and heme (*11*−*13*). Although GAF and PAS domains are unrelated in amino acid sequence, the similarities of their structural topologies have previously been recognized (*21*) through determination of a number of X-ray structures for both sensory domains from different organisms. Among them, crystal structures of the heme-bound PAS domains have been reported (*22*−*26*). Despite the finding that GAF domains can bind heme (*11*−*13*), and the importance of heme-bound GAF sensors for activation of the DosR-dependent hypoxia response in *M. tuberculosis*, no structural information is available for any heme-bound member of the GAF domain family. To improve our understanding of oxygen deprivation and NO and CO signaling in this obligate aerobic pathogen, we have determined the X-ray structure of the sensing GAF domain of *M. tuberculosis* DosT (GAF_{DosT}) in the oxygen-free and oxygen-bound states.

EXPERIMENTAL PROCEDURES

GAF_{DosT} Cloning and Expression. The sequence encoding amino acids 61−208 was amplified from the full-length DosT using Pfu Turbo DNA polymerase (Stratagene) and upstream CCGCCGCCATATG**CATCATCATCATCATCAC**GAGAACTTATATTTTCAAGGAAAGCTCGACGCCACCCTGCGCGCC and downstream GCGTCGGATCCCTATTACCGTGATTCCTCGAAGAGAC primers. The boldface indicates the His_6 tag. The underlining denotes an NdeI restriction cloning site in the upstream primer and the BamHI restriction cloning site in the downstream primer. The purified PCR product was digested with NdeI and BamHI and ligated into a pET23a+ vector (Invitrogen) cleaved with the same restriction enzymes and treated with alkaline phosphatase. The identity of the resulting expression vector pET23a+ GAF_{DosT} was confirmed by DNA sequencing.

GAF_{DosT} was heterologously expressed according to the protocol previously described for DevS (*12*). Briefly, BL21gold DE3 cells were cotransformed with pET23a+ GAF_{DosT} and pT-GroE. The cells were grown on Luria-Bertani agar plates containing both ampicillin (50 μg/mL) and chloramphenicol (34 μg/mL). Overnight cultures were grown at 37 °C and then used to inoculate flasks containing 1.5 L of Luria-Bertani medium supplemented with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). The cells were grown at 37 °C and 230 rpm until the OD_{600} reached 0.8−1.0. At that time, isopropyl 1-thio-β-D-galactopyranoside (1 mM) and hemin (30 mg/L) were added. The cells were incubated for an additional 20 h at 18 °C and then harvested by centrifugation at 5000 rpm for 25 min.

GAF_{DosT} Purification. The cells were lysed in 50 mM sodium phosphate (pH 7.6) containing 10% glycerol, 200 mM NaCl, 1% Triton X-100, 0.5 mg/mL lysozyme, 5 mM $MgCl_2$, 5 mM ATP, and the protease inhibitors antipain (1 μg/mL), leupeptin (1 μM), pepstatin (1 μM), and PMSF (0.1 mM). The cell suspension was incubated with shaking at 37 °C for 30 min. The cell membranes were disrupted by sonication using a Branson sonicator (VWR Scientific) while cooling on ice. Cell debris was removed by centrifugation at 35000 rpm for 1 h at 4 °C. The soluble extract was applied to a 5 mL HisTrap HP column (GE Healthcare) at a rate of 1 mL/min. The column was then washed with 50 mL each of 20 and 50 mM imidazole in 50 mM sodium phosphate (pH 7.6), 10% glycerol, and 500 mM NaCl at a rate of 5 mL/min. The recombinant protein was eluted with 50 mL of 250 mM imidazole in the same buffer. The protein was then mixed with AcTEV (Invitrogen) protease (10 μg/μL) and DTT (1 mM) and dialyzed for 4 days at 4 °C against 20 mM HEPES (pH 7.6) containing 150 mM NaCl, 1 mM DTT, and 10% glycerol. The cleaved protein was separated from the His_6-tagged counterpart using a 5 mL HisTrap HP column, then concentrated, and oxidized with potassium ferricyanide (10 mg/mL). Potassium ferricyanide was removed using a PD10 column (GE Healthcare) equilibrated with 20 mM HEPES (pH 7.6) containing 150 mM NaCl, and virtually homogeneous ferric GAF_{DosT} was concentrated to ~0.5 mM, as estimated on the basis of the Soret absorbance at 407 nm using the extinction coefficient of 166076 M^{-1} cm^{-1} previously calculated for DevS on the basis of the correlation between amino acid analysis and heme content (*12*).

Crystallization and Data Collection. Crystallization conditions were identified by the hanging drop vapor diffusion method using an Index high-throughput screen matrix (Hampton Research) and a nanoliter drop setter Mosquito (TTP LabTech). Diffraction-quality crystals were generated in the 24-well crystallization plates by mixing 1−2 μL of 0.5 mM fully oxidized GAF_{DosT} with 2 μL of well solution containing 50% Tacsimate (pH 6.0), a crystallization reagent composed of a mixture of seven organic acid salts. Prior to data collection, the crystals were plunged in the well solution supplemented with 20% glycerol and then flash-frozen in liquid nitrogen. Native and two-wavelength anomalous dispersion X-ray diffraction data were collected at 100−110 K on beamline 8.3.1 (Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA). Diffraction data reduction was performed using HKL2000 (*27*) or MOSFLM (*28*). Anomalous diffraction data were collected from one crystal at two wavelengths, one corresponding to the median between the Fe peak and the inflection point and the other at a point 375 eV higher (Table 1). The crystal diffracted in the $P4_3$ space group and had two molecules per asymmetric unit.

Structure Determination and Refinement. The ELVES automated software system (*29*) was used to determine the GAF_{DosT} crystal structure from the anomalous dispersion of the heme iron. ELVES identified and refined positions of the two possible Fe sites. Initial phases with an overall figure of merit of 0.20 were improved by solvent flattening (mean figure of merit after solvent flattening of 0.67) to provide an interpretable electron density map. Automated model building with ARP/wARP (*30*) and RESOLVE (*31*) placed a polyalanine backbone for 76% of the residues in the asymmetric unit. The remaining model was traced manually with COOT (*32*), alternated with TLS and positional refinement against the 2.3 Å native data set using REFMAC5 (*33, 34*). The composite omit map was calculated using CNS (*35*).

The final refined model includes 146 (60−205) and 149 (60−208) amino acid residues in chains A and B in an asymmetric unit, respectively, each bound to a heme prosthetic group with the distal pocket occupied with a water molecule (chain A) or dioxygen (chain B) as the second axial ligand (Table 1). The monitoring of quality of the final structure with PROCHECK indicated 90.2% of the residues are in the most favored, 5.7% of the residues are in additional allowed, 1.6% of the residues are in generously allowed, and 2.5% of the

Table 1: Diffraction Data and Phasing and Refinement Statistics

	native	Fe anomalous dispersion peak/inflection median	Fe anomalous dispersion remote
Data Collection			
space group	$P4_3$		$P4_3$
cell dimensions			
a, *b*, *c* (Å)	87.9, 87.9, 66.8	88.3, 88.3, 66.4	88.3, 88.3, 66.4
α, *β*, *γ* (deg)	90, 90, 90	90, 90, 90	90, 90, 90
wavelength	1.11587	1.73989	1.65312
resolution (Å)	2.3	3.1	2.95
no. of molecules in the asymmetric unit	2	2	2
R_{sym} (in last shell)[a]	6.7 (48.8)	8.2 (25.1)	8.1 (33.8)
$I/\sigma I$[a]	43.0 (4.6)	31.0 (6.4)	30.5 (4.5)
completeness (%)[a]	100.0 (99.8)	100.0 (99.8)	100.0 (100.0)
redundancy[a]	7.9 (6.4)	7.6 (6.9)	7.6 (6.9)
Refinement			
resolution (Å)	2.3		
no. of reflections	21731		
R_{work}/R_{free}	23.6/28.9		
no. of atoms			
protein	2262		
heme	86		
O_2 ligand	2		
water	115		
B-factor			
protein	50.7		
heme	37.1		
O_2 ligand	43.5		
water	42.5		
root-mean-square deviation (rmsd)			
bond lengths (Å)	0.022		
bond angles (deg)	2.28		

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

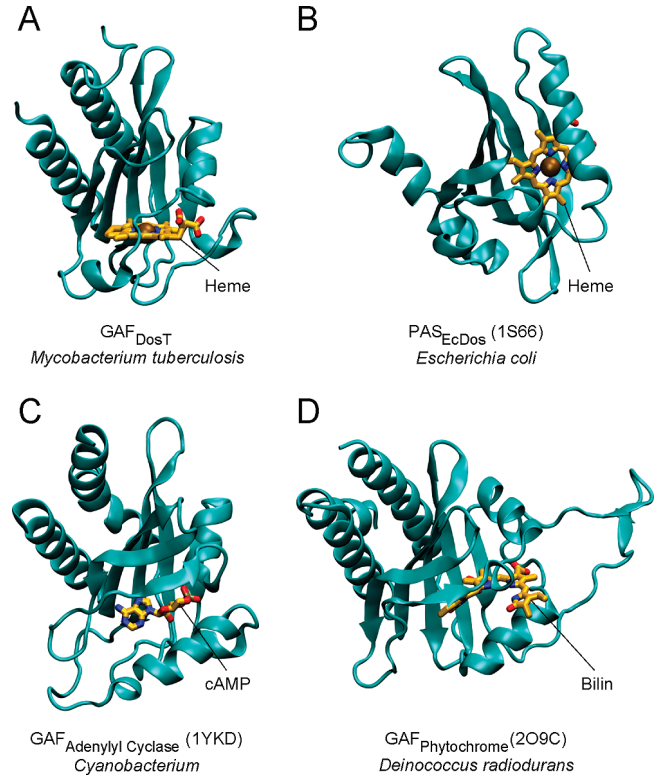

FIGURE 1: GAF and PAS sensory domains. Ribbon diagrams of GAF and PAS domains from different organisms are shown in the same orientation. Bound ligands are colored yellow. PDB entry codes are given in parentheses.

residues (three) are in disallowed regions of the Ramachandran plot for chain A and 84.7% of the residues are in the most favored, 11.3% of the residues are in additional allowed, 2.4% of the residues are in generously allowed, and 1.6% of the residues (two) are in disallowed regions for chain B.

RESULTS

GAF_{DosT} Photoreduction. During the first few seconds of diffraction data collection, the color of the irradiated part of the crystal changed to a more saturated hue of red (Figure S1 of the Supporting Information) and remained that way under prolonged X-ray exposure. This phenomenon can be readily explained by X-ray-induced photoreduction of metal centers (*36*−*38*), which has been studied particularly for hemoproteins (*39, 40*). Thus, the crystal structure reported in this work is most likely for a reduced state of the protein, even though initially its fully oxidized form was used to generate the crystals.

Overall Structure of GAF_{DosT}. The GAF_{DosT} structure was determined to a resolution of 2.3 Å by using heme Fe as an anomalously scattering atom (Table 1). The asymmetric unit contains two differently heme iron-ligated GAF_{DosT} monomers adopting similar conformations (rmsd of 0.5 Å for Cα atoms). Although, as expected, the overall GAF_{DosT} topology resembles that of previously defined GAF and PAS domains (Figure 1), GAF_{DosT} bears little structural similarity to previously characterized heme oxygen-sensing PAS domains (*22*−*26*) (Figure 1A,B). First, in GAF_{DosT}, heme binds in a virtually orthogonal orientation with respect to the twisted surface of the *β*-sheet (Figure 1A), in contrast to the more coplanar orientation of heme in the PAS domain (Figure 1B). Second, the histidine residue serving as a proximal axial ligand originates from the different structural elements: an equivalent of helix α3 in the PAS domain and a bend of the long *β*3−*β*4 loop in GAF_{DosT} (see Figure 2A for a reference). Collectively, the heme binding mode in GAF_{DosT} is similar to the cyclic nucleotide binding in GAF, rather than heme binding in a PAS domain, even though GAF_{DosT} contains a five-stranded antiparallel *β*-sheet instead of the six-stranded one found in GAF motifs that bind cyclic nucleotides (*41*) (Figure 1C) or linear tetrapyrrole chromophores (*42*) (Figure 1D). The *β*-sheet of GAF_{DosT} is packed directly against two long antiparallel α-helices (α1 and α4) on one side of the sheet and short helix α3 on the opposite side (Figure 2), whereas in the PAS domain, a pocket between helix α3 and the *β*-sheet accommodates the heme.

Heme Binding Site. The heme in GAF_{DosT} is solvent-shielded by short helix α2 and the α2−α3 connecting loop, both of which encircle the heme along the macrocycle edge, making contacts at residues Tyr98, Arg106, Ile109, Gly110, Ser111, Leu112, Glu114, Gly115, Arg116, Gly117, and Val118, which are within 4 Å of the heme. In addition, the long *β*3−*β*4 loop generates a proximal (with respect to the heme) surface of GAF_{DosT}, contacting heme at residues Ile134, Ala139, Ser140, Val141, Gly142, Phe143, Pro144, His147, Pro148, Pro149, Met150, Phe153, and Thr171 (marked by blue triangles in Figure 3), which are within 6 Å of the heme macrocycle. This latter arrangement constitutes a proximal heme-binding motif with a proximal iron axial ligand, H147 (highlighted in cyan).

Ligand Binding Site. A small ligand-binding pocket with an accessible volume of 10 $Å^3$, as determined using a 1.4 Å

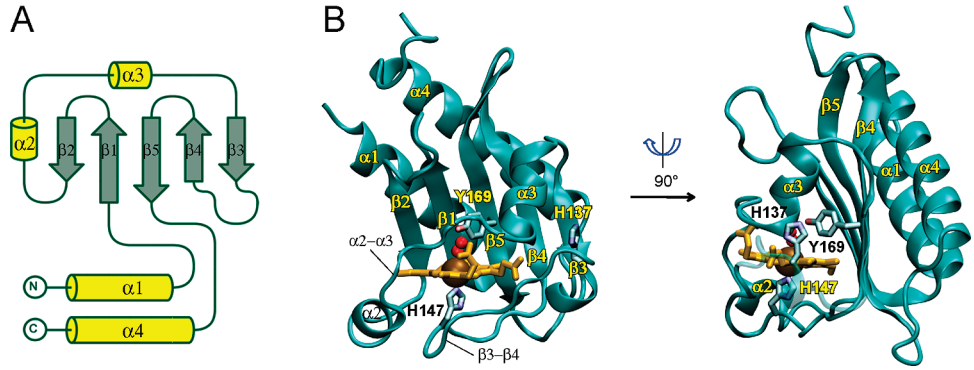

FIGURE 2: Overall structure of GAF_{DosT}. (A) Topology diagram of GAF_{DosT}. (B) Ribbon representation of the two orthogonal views of GAF_{DosT}. The O_2 ligand is represented by red spheres, and the heme (orange) and H137, H147, and Y169 are shown as sticks.

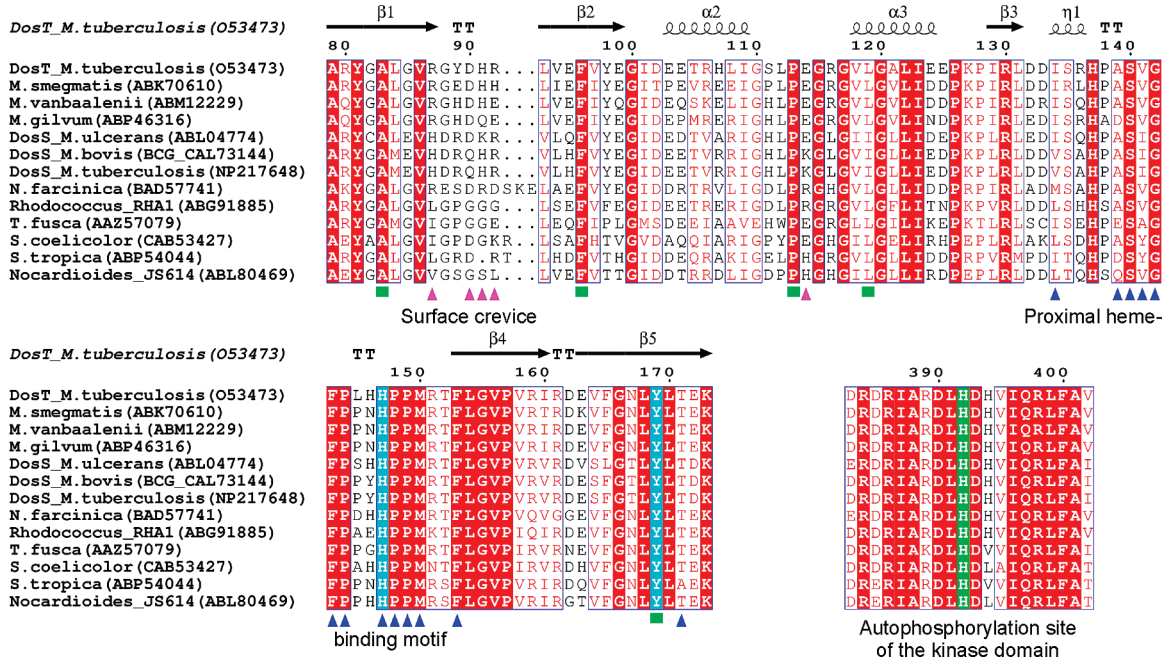

FIGURE 3: Heme-binding and ligand-binding motifs. Two fragments of multiple-sequence alignments between putative two-component sensory histidine kinases from different organisms are shown. Accession numbers of the proteins in Swiss-Prot/TrEMBL (http://us.expasy.org/sprot) or NCBI (http://www.ncbi.nlm.nih.gov) databases are given next to the name of the host organism (*M. smegmatis*, *Mycobacterium smegmatis*; *M. vanbaalenii*, *Mycobacterium vanbaalenii*; *M. gilvum*, *Mycobacterium gilvum*; *M. ulcerans*, *Mycobacterium ulcerans*; *M. bovis*, *Mycobacterium bovis*; *N. farcinica*, *Nocardia farcinica*; *T. fusca*, *Thermobifida fusca*; *S. coelicolor*, *Streptomyces coelicolor*; *S. tropica*, *Salinispora tropica*). Secondary structure annotation and residue numbering at the top correspond to GAF_{DosT} of *M. tuberculosis*. The α-helices are represented by spirals and *β*-strands by arrows. Alignments were performed for the full-length sequences using the MAP algorithm as implemented in the BCM Search Launcher (*59*). Entire alignments are shown in Figure S2 of the Supporting Information. Residues constituting the proximal heme-binding motif as deduced from the crystal structure are marked with blue triangles. Residues constituting the distal ligand-binding pocket are marked with green squares. Residues making up the surface crevice are marked with pink triangles. Invariant H147 and Y169 are highlighted in cyan. Highly conserved H392 of the C-terminal kinase domain implicated as the site of autophosphorylation (*9, 10*) is highlighted in green.

probe, is located on the distal side of the heme (Figure 4A). Five side chains within 4 Å of the heme (marked by green squares in Figure 3), spread along the primary sequence, constitute a ligand-binding GAF motif (A83-F96-P113-L119-Y169), which may convey a signal simultaneously to the different regions of GAF_{DosT}. The hydroxyphenyl ring of Y169 protrudes from strand *β*5 with the long axis parallel to the heme plane and the OH group projecting onto the nitrogen atom of pyrrole ring I (Figure 4A). The plane of the Y169 ring is orthogonal to the heme plane with the ring edge only 3.44 Å from the heme *δ*-meso carbon atom, suggesting $\pi-\pi$ interactions between these aromatic systems. Y169 separates the ligand-binding pocket from a larger connected cavity (accessible volume of 45.0 $Å^3$) comprised of an amino acid motif (R87-L119-I123-N167-Y169). The cavity accommodates two water molecules separated from each other by 2.5 Å in the O_2-bound monomer and one water molecule in the O_2-free monomer. These water molecules H-bond to the guanidinium group of R87 and the carboxamide group of N167 (Figure 4A).

The tunnel-connected cavities can be considered a pathway leading from the ligand-binding site to a small pore on the molecular surface adjacent to a surface crevice formed by the set of residues R87, D90, H91, R92, and E114 (marked by pink triangles in Figure 3) holding an acetate group bound in both monomers of the crystal structure (Figure 4B). The surface pore may serve as a port of ligand entry. The 275 D molecular dipole moment of GAF_{DosT} points straight through that pore (Figure 4C).

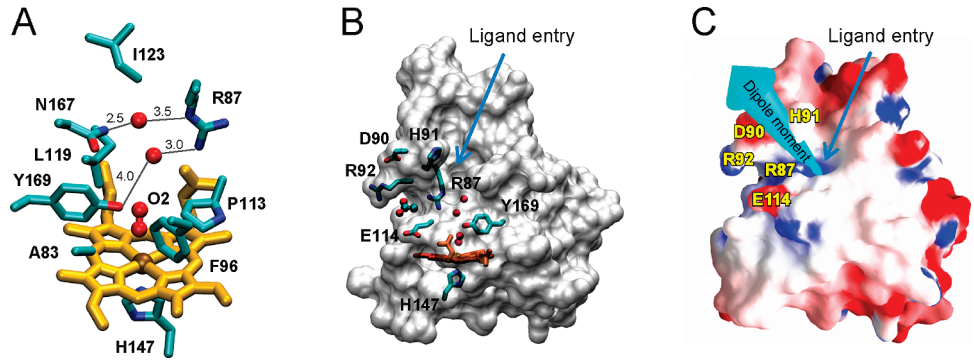

FIGURE 4: O_2 binding in GAF_{DosT}. (A) GAF_{DosT} tunnel-connected ligang-binding pocket and the adjacent cavity. The O_2 ligand and the water molecules that may indicate a ligand migration route through the protein matrix are represented by the red spheres. (B) Semitransparent surface of GAF_{DosT}. The putative port of the ligand entry is denoted with a blue arrow. (C) Solvent-accessible molecular surface of the deoxy form color-coded for electrostatic potential as generated with GRASP. The deepest shades of red and blue correspond to potentials of −5.5 and 7.3 kcal, respectively. The dipole moment of 275 D is denoted with a cyan arrow.

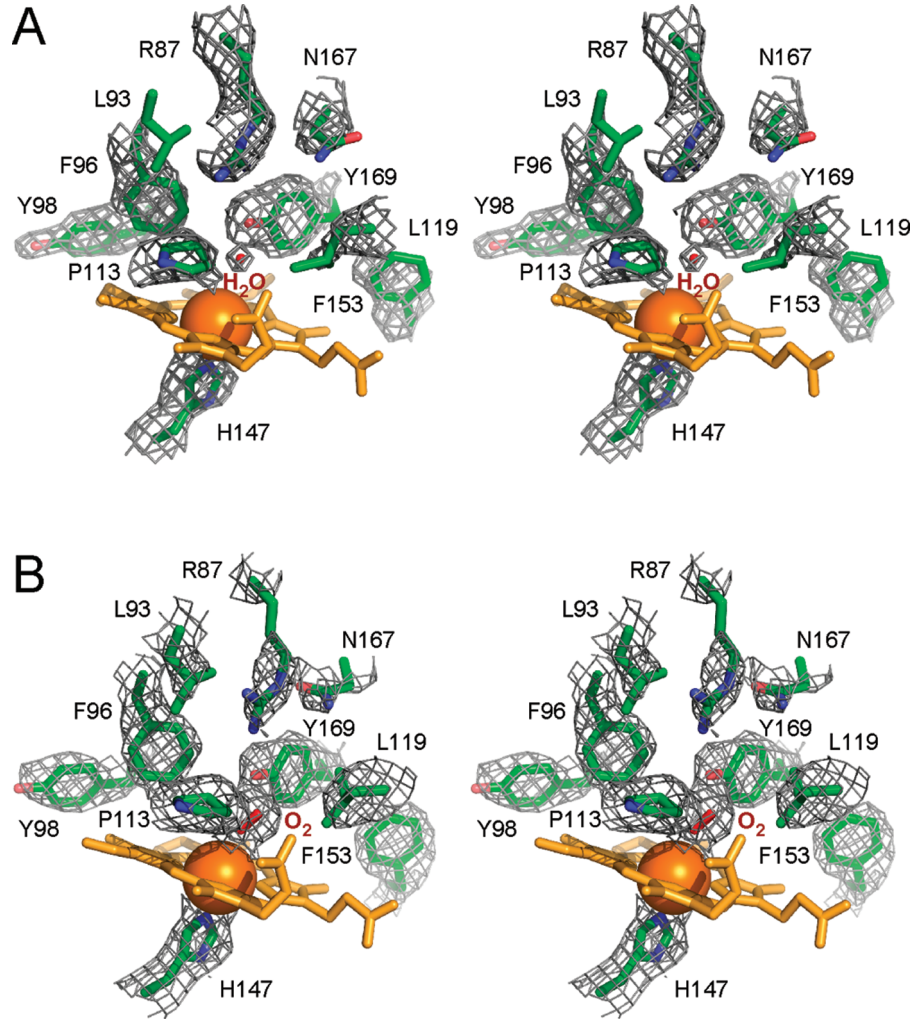

FIGURE 5: Electron density map. Fragments of the composite omit electron density map generated around the ligand-binding site contoured at 1.1σ are shown with a gray mesh in stereo mode for deoxy (A) and oxy (B) GAF_{DosT}. To avoid excessive cluttering, heme (orange) was excluded from the map calculation in panels A and B, and Leu93 was excluded in panel A only. Amino acid side chains are colored green, oxygen atoms red, and nitrogens blue.

Binding of a Second Axial Iron Ligand. In one of the two monomers in the asymmetric unit, weak electron density for a water molecule centered within 3.7 Å of the heme iron and at a H-bond distance of 2.46 Å from the Y169 phenolic O atom is evident in the electron density composite omit map (Figure 5A). The water molecule tends to maintain a perfect H-bonding geometry with respect to Y169 at the expense of the coordination bond to the heme iron. Consistent with reduction of the Fe center during X-ray data collection, this water molecule might have been released from the iron coordination sphere but nevertheless trapped in the vicinity by a H-bond to Y169.

In contrast, prominent electron density was observed next to the heme iron in the second monomer in the asymmetric unit (Figure 5B), where an axial water ligand was also placed initially. Later in the refinement, the residual $F_o - F_c$ map

revealed a positive peak adjacent to the oxygen atom, consistent with the binding of a diatomic molecule as a sixth iron ligand. Three diatomic gases, O_2, CO, and NO, have been shown to bind and modulate the autokinase activity of DevS and DosT (*13, 14*). We reasoned that O_2, abundant under the aerobic conditions used for crystallization and data collection, has the highest probability to be a second axial ligand in the irradiated crystals. Therefore, O_2 was fitted into the electron density (Figure 5B) with full occupancy within 2.60 Å of the heme iron for the proximal oxygen atom (with respect to the heme iron) and 2.50 Å of the Y169 OH group for the distal oxygen atom and 2.94 Å for the proximal one, suggesting a strong H-bonding role for Y169. The Fe−O bond in GAF_{DosT} (2.60 Å) is somewhat longer than that reported in the oxy complex of the heme-bound PAS domain of *Escherichia coli* EcDos [2.37 Å (PDB entry 1VB6) (*26*) and 2.31 Å (PDB entry 1S66) (*25*)]. In contrast, the Fe−O bond reported for the *Bradyrhizobium japonicum* FixL oxy complex (PDB entry 1DP6) (*43*) is notably shorter (1.81 Å), perhaps due to the polar environment of the O_2-binding pocket caused by the proximity of the side chain of the PAS domain invariant arginine. This trend in Fe−O bond distances is consistent with the Fe−O stretching frequencies observed in the resonance Raman spectra. Indeed, the Fe−O_2 stretching mode is observed near 560 cm^{-1} in DevS (65% identical to DosT, including the H-bonding distal tyrosine residue) (*44*) and EcDos (*45*), but at 571 cm^{-1} in FixL (*46*).

The oxygen molecule is tilted at an ~123° angle to the heme plane and leans toward the porphyrin γ-meso carbon atom. The bond angle of O_2 in GAF_{DosT} is close to those, 118° (*26*) and 130° (*25*), previously observed in the crystal structures reported for *E. coli* EcDOS. As in GAF_{DosT}, in both EcDos X-ray structures, only one of the two protein monomers in an asymmetric unit is in the oxy form. For *B. japonicum* FixL, the reported bond angles for O_2 and NO are 142° (*43*) and 122° (*43*), respectively. At the same time, both CO and CN^- bind almost orthogonally to the heme plane (*24, 47*), thus excluding a possibility of any of them being bound in the GAF_{DosT} sensing pocket. Given the demonstrated ease of X-ray-induced heme iron photoreduction (*39, 40*) and the abundance of oxygen (but not NO) in the air, coordination of the oxygen molecule in GAF_{DosT} almost certainly occurred to the ferrous iron in the crystal during diffraction data collection.

Inspection of the crystal structure reported here indicates that the distal Y169 cannot itself coordinate to the heme Fe atom without a significant conformational rearrangement of the distal heme pocket.

Conformational Differences between GAF_{DosT} Monomers. Small conformational differences observed between differently ligated GAF_{DosT} monomers largely converge to two protein regions, the N-terminus of helix α2 and the C-terminus of helix α3, where pairs of consecutive glutamate residues, E103-E104 and E124-E125, respectively, are located (Figure 6). The occurrence of pairs of negatively charged residues conserved between DosT and DevS (Figure 3) in both conformationally variable spots is intriguing and may suggest that the surface electrostatic potential is modulated during signaling, possibly affecting downstream interactions with other protein domains. However, as far as the conformational changes are concerned, GAF_{DosT} does not operate as an isolated domain; hence, the possibility remains that it may behave differently as a part of the full-length enzyme.

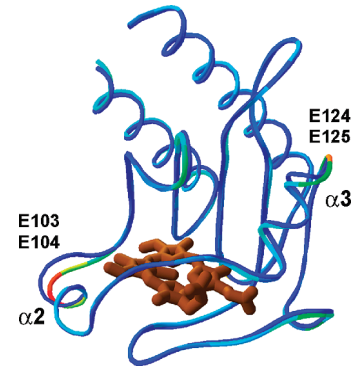

FIGURE 6: Conformational differences between GAF_{DosT} monomers. Superimposition of the two GAF_{DosT} monomers in the asymmetric unit is shown as a ribbon diagram. C-Termini are trimmed up to A201 to emphasize the structural differences between monomers by coloring the oxy form according to the rms deviations (highest colored red and lowest blue). The deoxy form is shown as a blue ribbon.

Heme-Binding and Ligand-Binding Motifs. A BLAST search of the NCBI protein database against the full-length *M. tuberculosis* DosT sequence returns 13 prokaryotic proteins with sequence homology extending nearly the entire length of the protein, including an N-terminal GAF domain and a C-terminal kinase region (Figure S2 of the Supporting Information). Two fragments of Figure S2 representing the structurally defined portion of the N-terminal GAF domain and the autophosphorylation site of the C-terminal kinase domain are shown in Figure 3. This group of clearly related proteins in addition to DosT includes three DevS proteins from *Mycobacterium* species and nine proteins of unknown function from aerobic bacteria. Worth mentioning is the fact that His392 (highlighted in green in Figure 3), implicated as the site of autophosphorylation in DosT (*10*) and DevS (*9*), is invariant in all 13 proteins. Also, the iron proximal ligand H147 (highlighted in cyan) is one of two invariant histidine residues in the heme-binding GAF motif, with the other, H137, situated on the domain surface (Figure 2B), suggesting a role for it in protein−protein interactions. Y169, which provides the H-bond to the bound O_2 ligand, is also invariant.

The alignments show conservation of the proximal heme-binding and ligand-binding motifs as they were deduced from the GAF_{DosT} crystal structure, suggesting that these bacterial proteins may all be utilized by their host organisms as heme-based GAF sensors for adaptive responses to fluctuating O_2, CO, or NO levels. Thus, nine of the total of 13 residues (marked by blue triangles in Figure 3) that constitute the proximal heme-binding motif [Ile134-Ala139-**Ser**140-Val141-**Gly**142-**Phe**143-**Pro**144-**His**147-**Pro**148-**Pro**149-**Met**150-**Phe**153-Thr/Ala171 (boldface indicates invariant residues)] are invariant in all the putative family members, suggesting a common heme binding mode. Four of the total of five residues (marked by green squares) of the ligand-binding motif [**Ala**83-**Phe**96-**Pro**113-Leu(Ile)119-**Tyr**169] are invariant, whereas leucine at position 119 is substituted with isoleucine in DevS sequences. At the same time, the residues contacting the heme edge are more variable. In contrast, residues R87, D90, H91, R92, and E114 (marked by pink

triangles) constituting the surface crevice adjacent to the putative port of ligand entry (Figure 4B,C) are, remarkably, least conserved among all the residues of the N-terminal GAF domain.

DISCUSSION

Understanding oxygen deprivation and NO/CO signaling in the obligate aerobic pathogen *M. tuberculosis* is an important pharmacological objective since in the dormant state achieved through the expression of the Dos regulon genes *M. tuberculosis* is least susceptible for treatment with the available antimycobacterial drugs. In this work, we report the X-ray structure of heme-bound GAF_{DosT}, which may serve as a prototype for the heme-based GAF sensory domains of histidine kinases.

Despite the recognized similarities of structural topologies between GAF and PAS domains (Figure 1), the heme binding mode of GAF_{DosT} differs from that of the PAS domains, in both heme orientation and specific heme−protein and ligand−protein contacts. GAF_{DosT} provides the first insight into the ligand-binding pocket architecture of heme-based GAF sensors. It predicts a key role for Y169 in sensor−ligand interactions, as Y169 is the only residue in the GAF_{DosT} ligand-binding site strategically positioned to provide a H-bond to the bound oxygen. The immediate presence of the Y169 OH group in GAF_{DosT} contrasts with the hydrophobic environment of the heme-containing PAS domain of FixL sensors that govern nitrogen fixation in *Rhizobia*, in which the side chain of a highly conserved arginine enters the O_2-binding pocket to directly interact with the bound oxygen only in response to its binding (*22, 24*).

Precedents for a tyrosine in the distal pocket of heme proteins that bind diatomic gaseous ligands are available in the literature. Tyrosine is the primary factor responsible for high oxygen affinity in heme enzymes acting as NO dioxygenases, such as truncated-N hemoglobins from *Ascaris* (*48*) or *M. tuberculosis* (*49*), that catalyze NO detoxification to innocuous nitrate under hypoxic conditions. Replacement of tyrosine with phenylalanine in truncated-N hemoglobin increases the O_2 rate dissociation constant (k_{off}) by 2 orders of magnitude (*50*). However, in NO dioxygenases, O_2 is stabilized by two H-bond donor residues, tyrosine and glutamine (*51*), whereas in GAF_{DosT}, Y169 is the only residue in the ligand-binding site capable of providing a H-bond to finely tune the affinity of the sensor for a precise response to a given oxygen concentration.

As the heme in GAF_{DosT} is not readily accessible from the environment, the ligand molecule has to migrate to the active site across the protein matrix, probably via tunnel-connected cavities on the distal side of the heme. The limited number of water molecules associated via H-bonds with tunnel amino acid residues (Figure 4A,B) may delineate a putative ligand migration route. Y169, which deeply protrudes into the sensing pocket, blocks the continuity of the tunnel and is a potential barrier for ligand diffusion. Thus, protein motions would be expected to affect the dynamics of ligand migration along the tunnel. In this regard, small cavities (13−45 $Å^3$) have been observed in sperm whale myoglobin that are recognized to act as transient binding sites for gaseous ligands O_2, NO, and CO (*52*−*56*).

Although a conformational switch following binding of a signaling ligand has been postulated for oxygen sensors (*57*), the experimentally observed magnitudes of such changes vary between deoxy and oxy states among the two heme-containing sensory PAS domains for which crystal structures are available. Sizable and highly localized conformational differences between differently ligated heme domains were described for the *E. coli* O_2 sensor EcDos (*25, 26*), where a methionine side chain axially coordinates to the heme iron in the absence of the O_2 ligand. Less extensive changes accompany O_2 binding in FixL sensors (*22*−*24, 43*). Even smaller conformational changes between differently ligated GAF_{DosT} monomers observed in this study may result from crystal lattice limitations, since binding of atmospheric oxygen in the crystal likely occurred as a result of X-ray-induced Fe^{3+} photoreduction. Alternatively, since the heme binding mode in GAF_{DosT} is very different compared to that in more extensively studied heme-based PAS sensors, the modest conformational variations observed between the two GAF_{DosT} structures may be indicative of a new mode of signal propagation in heme-based sensors.

In summary, the determination of the first X-ray crystal structure of the oxygen-sensing GAF domain, reported here, is an important step toward understanding the mechanism of signal recognition in the *M. tuberculosis* sensory histidine kinases DosT and DevS. The crystal structure reveals the key functional residues that are highly conserved in both proteins and sets up a framework in which to test the hypothesis. Thus, the H-bonding role of the distal tyrosine in discriminating the gaseous ligands has been confirmed in parallel studies (*58*) via site-directed mutagenesis of the full-length histidine kinase DevS, combined with analysis of the mutants in autophosphorylation assays and by resonance Raman spectroscopy.

ACKNOWLEDGMENT

We thank Potter Wickware for critical reading of the manuscript, Clinton Nishida and Chris Waddling for assistance with software and instrumentation, and the staff members of beamline 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory for assistance. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

SUPPORTING INFORMATION AVAILABLE

Spectral characterization of GAF_{DosT} (Figure S1), GAF_{DosT} X-ray-induced photoreduction (Figure S2), and full-length multiple-sequence alignments between putative two-component sensory protein histidine kinases from different organisms (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wayne, L. G., and Sohaskey, C. D. (2001) Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol. 55*, 139–163.
2. Ohno, H., Zhu, G., Mohan, V. P., Chu, D., Kohno, S., Jacobs, W. R., Jr., and Chan, J. (2003) The effects of reactive nitrogen intermediates on gene expression in *Mycobacterium tuberculosis*. *Cell. Microbiol. 5*, 637–648.
3. Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., Schoolnik, G. K., and Sherman, D. R. (2003) Rv3133c/

dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol. 48*, 833–843.
4. Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., and Schoolnik, G. K. (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J. Exp. Med. 198*, 705–713.
5. Shiloh, M. U., Manzanillo, P., and Cox, J. S. (2008) *Mycobacterium tuberculosis* senses host-derived carbon monoxide during macrophage infection. *Cell Host Microbe 3*, 323–330.
6. Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Agarwal, A., and Steyn, A. J. (2008) Heme oxygenase-1-derived carbon monoxide induces the *Mycobacterium tuberculosis* dormancy regulon. *J. Biol. Chem. 283*, 18032–18039.
7. Kinger, A. K., and Tyagi, J. S. (1993) Identification and cloning of genes differentially expressed in the virulent strain of *Mycobacterium tuberculosis*. *Gene 131*, 113–117.
8. Roberts, D. M., Liao, R. P., Wisedchaisri, G., Hol, W. G., and Sherman, D. R. (2004) Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem. 279*, 23082–23087.
9. Saini, D. K., Malhotra, V., Dey, D., Pant, N., Das, T. K., and Tyagi, J. S. (2004) DevR-DevS is a bona fide two-component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology 150*, 865–875.
10. Saini, D. K., Malhotra, V., and Tyagi, J. S. (2004) Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. *FEBS Lett. 565*, 75–80.
11. Sardiwal, S., Kendall, S. L., Movahedzadeh, F., Rison, S. C., Stoker, N. G., and Djordjevic, S. (2005) A GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem. *J. Mol. Biol. 353*, 929–936.
12. Ioanoviciu, A., Yukl, E. T., Moenne-Loccoz, P., and Ortiz de Montellano, P. R. (2007) DevS, a heme-containing two-component oxygen sensor of *Mycobacterium tuberculosis*. *Biochemistry 46*, 4250–4260.
13. Sousa, E. H., Tuckerman, J. R., Gonzalez, G., and Gilles-Gonzalez, M. A. (2007) DosT and DevS are oxygen-switched kinases in *Mycobacterium tuberculosis*. *Protein Sci. 16*, 1708–1719.
14. Kumar, A., Toledo, J. C., Patel, R. P., Lancaster, J. R., Jr., and Steyn, A. J. (2007) *Mycobacterium tuberculosis* DosS is a redox sensor and DosT is a hypoxia sensor. *Proc. Natl. Acad. Sci. U.S.A. 104*, 11568–11573.
15. Nathan, C., and Ehrt, S. (2003) Nitric oxide in tuberculosis. In *Tuberculosis* (Rom, W., and Garay, S., Eds.) pp 215−235, Lippincott, Williams and Walkins, New York.
16. Brown, G. C. (2001) Regulation of mitochondrial respiration by nitric oxide inhibition of cytochrome c oxidase. *Biochim. Biophys. Acta 1504*, 46–57.
17. Martin, E., Davis, K., Bian, K., Lee, Y. C., and Murad, F. (2000) Cellular signaling with nitric oxide and cyclic guanosine monophosphate. *Semin. Perinatol. 24*, 2–6.
18. Martinez, S. E., Beavo, J. A., and Hol, W. G. (2002) GAF domains: Two-billion-year-old molecular switches that bind cyclic nucleotides. *Mol. Interventions 2*, 317–323.
19. Galperin, M. Y. (2004) Bacterial signal transduction network in a genomic perspective. *Environ. Microbiol. 6*, 552–567.
20. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature 438*, 325–331.
21. Ho, Y. S., Burden, L. M., and Hurley, J. H. (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *EMBO J. 19*, 5288–5299.
22. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998) Structure of a biological oxygen sensor: A new mechanism for heme-driven signal transduction. *Proc. Natl. Acad. Sci. U.S.A. 95*, 15177–15182.
23. Miyatake, H., Mukai, M., Park, S. Y., Adachi, S., Tamura, K., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T., and Shiro, Y. (2000) Sensory mechanism of oxygen sensor FixL from *Rhizobium meliloti*: Crystallographic, mutagenesis and resonance Raman spectroscopic studies. *J. Mol. Biol. 301*, 415–431.
24. Hao, B., Isaza, C., Arndt, J., Soltis, M., and Chan, M. K. (2002) Structure-based mechanism of O_2 sensing and ligand discrimination by the FixL heme domain of *Bradyrhizobium japonicum*. *Biochemistry 41*, 12952–12958.
25. Park, H., Suquet, C., Satterlee, J. D., and Kang, C. (2004) Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of *Escherichia coli* Dos heme domain (Ec DosH). *Biochemistry 43*, 2738–2746.
26. Kurokawa, H., Lee, D. S., Watanabe, M., Sagami, I., Mikami, B., Raman, C. S., and Shimizu, T. (2004) A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor. *J. Biol. Chem. 279*, 20186–20193.
27. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
28. Leslie, A. G. W. (1992) *Joint CCP4 ESF-EAMCB Newsletter on Protein Crystallography,* 26.
29. Holton, J., and Alber, T. (2004) Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. U.S.A. 101*, 1537–1542.
30. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol. 6*, 458–463.
31. Terwilliger, T. (2004) SOLVE and RESOLVE: Automated structure solution, density modification and model building. *J. Synchrotron Radiat. 11*, 49–52.
32. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
33. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240–255.
34. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D57*, 122–133.
35. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
36. Teng, T. Y., and Moffat, K. (2000) Primary radiation damage of protein crystals by an intense synchrotron X-ray beam. *J. Synchrotron Radiat. 7*, 313–317.
37. Ravelli, R. B., and McSweeney, S. M. (2000) The 'fingerprint' that X-rays can leave on structures. *Structure 8*, 315–328.
38. Burmeister, W. P. (2000) Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D56*, 328–341.
39. Beitlich, T., Kuhnel, K., Schulze-Briese, C., Shoeman, R. L., and Schlichting, I. (2007) Cryoradiolytic reduction of crystalline heme proteins: analysis by UV-Vis spectroscopy and X-ray crystallography. *J. Synchrotron Radiat. 14*, 11–23.
40. Arcovito, A., Moschetti, T., D'Angelo, P., Mancini, G., Vallone, B., Brunori, M., and Della Longa, S. (2008) An x-ray diffraction and X-ray absorption spectroscopy joint study of neuroglobin. *Arch. Biochem. Biophys., 475*, 7–13.
41. Martinez, S. E., Bruder, S., Schultz, A., Zheng, N., Schultz, J. E., Beavo, J. A., and Linder, J. U. (2005) Crystal structure of the tandem GAF domains from a cyanobacterial adenylyl cyclase: Modes of ligand binding and dimerization. *Proc. Natl. Acad. Sci. U.S.A. 102*, 3082–3087.
42. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of *Deinococcus* bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem. 282*, 12298–12309.
43. Gong, W., Hao, B., and Chan, M. K. (2000) New mechanistic insights from structural studies of the oxygen-sensing domain of *Bradyrhizobium japonicum* FixL. *Biochemistry 39*, 3955–3962.
44. Yukl, E. T., Ioanoviciu, A., de Montellano, P. R., and Moenne-Loccoz, P. (2007) Interdomain interactions within the two-component heme-based sensor DevS from *Mycobacterium tuberculosis*. *Biochemistry 46*, 9728–9736.
45. Tomita, T., Gonzalez, G., Chang, A. L., Ikeda-Saito, M., and Gilles-Gonzalez, M. A. (2002) A comparative resonance Raman analysis of heme-binding PAS domains: Heme iron coordination structures of the BjFixL, AxPDEA1, EcDos, and MtDos proteins. *Biochemistry 41*, 4819–4826.
46. Tamura, K., Nakamura, H., Tanaka, Y., Oue, S., Tsukamoto, K., Nomura, M., Tsuchiya, T., Adachi, S., Takahashi, S., Iizuka, T., and Shiro, Y. (1996) Nature of endogenous ligand binding to heme iron in oxygen sensor FixL. *J. Am. Chem. Soc. 118*, 9434–9435.
47. Key, J., and Moffat, K. (2005) Crystal structures of deoxy and CO-bound bjFixLH reveal details of ligand recognition and signaling. *Biochemistry 44*, 4627–4635.

48. Minning, D. M., Gow, A. J., Bonaventura, J., Braun, R., Dewhirst, M., Goldberg, D. E., and Stamler, J. S. (1999) *Ascaris* haemoglobin is a nitric oxide-activated 'deoxygenase'. *Nature 401*, 497–502.
49. Milani, M., Pesce, A., Nardini, M., Ouellet, H., Ouellet, Y., Dewilde, S., Bocedi, A., Ascenzi, P., Guertin, M., Moens, L., Friedman, J. M., Wittenberg, J. B., and Bolognesi, M. (2005) Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins. *J. Inorg. Biochem. 99*, 97–109.
50. Couture, M., Yeh, S. R., Wittenberg, B. A., Wittenberg, J. B., Ouellet, Y., Rousseau, D. L., and Guertin, M. (1999) A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A. 96*, 11223–11228.
51. Ouellet, Y., Milani, M., Couture, M., Bolognesi, M., and Guertin, M. (2006) Ligand interactions in the distal heme pocket of *Mycobacterium tuberculosis* truncated hemoglobin N: Roles of TyrB10 and GlnE11 residues. *Biochemistry 45*, 8770–8781.
52. Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) Cavities in proteins: Structure of a metmyoglobin-xenon complex solved to 1.9 Å. *Biochemistry 23*, 2849–2857.
53. Brunori, M., and Gibson, Q. H. (2001) Cavities and packing defects in the structural dynamics of myoglobin. *EMBO Rep. 2*, 674–679.
54. Scott, E. E., Gibson, Q. H., and Olson, J. S. (2001) Mapping the pathways for O_2 entry into and exit from myoglobin. *J. Biol. Chem. 276*, 5177–5188.
55. Srajer, V., Ren, Z., Teng, T. Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M., and Moffat, K. (2001) Protein conformational relaxation and ligand migration in myoglobin: A nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. *Biochemistry 40*, 13802–13815.
56. Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Jr., Wulff, M., and Anfinrud, P. A. (2003) Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. *Science 300*, 1944–1947.
57. Gilles-Gonzalez, M. A., and Gonzalez, G. (2005) Heme-based sensors: Defining characteristics, recent developments, and regulatory hypotheses. *J. Inorg. Biochem. 99*, 1–22.
58. Yukl, E. T., Ioanoviciu, A., Nakano, M. M., Ortiz de Montellano, P. R., and Moenne-Loccoz, P. (2008) A distal tyrosine residue is required for exogenous ligand discrimination in DevS from *Mycobacterium tuberculosis*. *Biochemistry 47*, 12532–12539.
59. Smith, R. F., Wiese, B. A., Wojzynski, M. K., Davison, D. B., and Worley, K. C. (1996) BCM Search Launcher: An integrated interface to molecular biology data base search and analysis services available on the World Wide Web. *Genome Res. 6*, 454–462.

BI8012356