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)†

HaJeung Park,‡ Christine Suquet,§ James D. Satterlee,*,§ and ChulHee Kang*,‡

School of Molecular Biosciences, Washington State University, Pullman, Washington 99164-4660, and Department of Chemistry, Washington State University, Pullman, Washington 99164-4630

Received November 6, 2003; Revised Manuscript Received January 26, 2004

ABSTRACT: The X-ray crystal structure of the *Escherichia coli* (*Ec*) direct oxygen sensor heme domain (*Ec* DosH) has been solved to 1.8 Å using Fe multiple-wavelength anomalous dispersion (MAD), and the positions of Met95 have been confirmed by selenomethionine (^{Se}Met) MAD. *Ec* DosH is the sensing part of a larger two-domain sensing/signaling protein, in which the signaling domain has phosphodiesterase activity. The asymmetric unit of the crystal lattice contains a dimer comprised of two differently ligated heme domain monomers. Except for the heme ligands, the monomer heme domains are identical. In one monomer, the heme is ligated by molecular oxygen (O_2), while in the other monomer, an endogenous Met95 with S → Fe ligation replaces the exogenous O_2 ligand. In both heme domains, the proximal ligand is His77. Analysis of these structures reveals sizable ligand-dependent conformational changes in the protein chain localized in the FG turn, the G_β-strand, and the HI turn. These changes provide insight to the mechanism of signal propagation within the heme domain following initiation due to O_2 dissociation.

It is now recognized that a collection of heme-b-containing proteins comprise a functionally related group that couple small molecule sensing to biological signaling (*1*−*3*). Prominent among these heme proteins are the FixLs,[1] which are multidomain proteins with separate domains employed for oxygen sensing (heme), signaling (catalytic), and in one case, membrane integration (*4*−*12*). These proteins typically form part of a more complex two-component cellular sensing and signaling system (*4*, *5*, *7*, *11*). FixLs couple heme oxygenation status to expression of the complement of proteins required for nitrogen fixation through a second component, the response regulator FixJ (*4*, *5*, *7*−*9*). Of the heme proteins in this category that have so far been identified, the chemically best characterized are two bacterial FixLs (*Sm* FixL and *Bj* FixL) (*1*−*3*, *10*, *12*−*25*), followed by the *Escherichia coli* Dos (*26*−*33*). Compared to the FixLs, the function of Dos in *E. coli* is more ambiguous but may involve changes between aerobic and anaerobic metabolism.

Because of their size and the complications associated with handling the full-length FixL and Dos proteins, most work has focused initially on the recombinant heme-containing domain of each of these proteins (FixLH, DosH). Heme domains have also been the initial focus because they were the first identified heme-containing PAS[1] proteins and because the signaling function begins in the heme domain, controlled by heme oxygenation status. As currently understood, dissociation of oxygen from the FixL heme domain (sensing domain) initiates the chemical signal that ultimately results in production of the proteins required to reduce atmospheric nitrogen (*1*−*11*). Accordingly, initiation and initial propagation of this signal must occur in the heme domain (*1*−*3*, *12*, *14*). To better understand this process, X-ray crystal structures of *Sinorhyzobium meliloti* (*Sm*) and *Bradyrhizobium japonicum* (*Bj*) FixL heme domains (FixLH) have been published (*12*−*15*, *25*). These structures have resulted in various ideas concerning signal initiation and aspects of transduction within the heme domains (*12*−*16*, *18*, *19*, *22*, *24*, *25*, *33*). In the full-length proteins, FixL and Dos, the signal is transmitted to the protein's catalytic domain. In each FixL, this catalytic region is a kinase (*8*, *19*), while in Dos, it functions as a phosphodiesterase (*26*, *27*).

Structurally, both (*Sm* and *Bj*) of the FixLH domains were shown to be PAS-heme proteins (*12*, *15*), and by sequence analogy, *Ec* DosH was also inferred to be a PAS-heme domain. However, some properties of *Ec* DosH differ from those of the FixLHs (*26*). For example, *Ec* DosH was discovered to be low-spin, indicating a six-coordinate heme, whether or not an exogenous ligand was available for heme

† This work was supported by the National Institutes of Health through Grants RO1 GM47645 (J.D.S.) and RO1 GM66173 (C.H.K.).

* To whom correspondence should be addressed. For J.D.S.: e-mail, hemeteam@wsu.edu; tel, 509-335-8620; fax, 509-335-8867. For C.H.K.: e-mail, chkang@mail.wsu.edu; tel, 509-335-1409; fax, 509-335-9688.

‡ School of Molecular Biosciences.

§ Department of Chemistry.

[1] Abbreviations: *Bj*, *Bradyrhizobium japonicum*; *Sm*, *Sinorhizobium meliloti*, formerly known as *Rhizobium meliloti* (*Rm*); *Ec*, *Escherichia coli*; FixL, full-length sequence including all domains; Dos, direct oxygen sensor full-length sequence including all domains; FixLH, recombinant truncated protein including only the FixL heme domain; DosH, recombinant truncated protein consisting only of the Dos heme domain; *Ec* DosH, the recombinant heme domain containing natural methionines; ^{Se}Met-*Ec* DosH, the recombinant heme domain containing selenomethionines; PAS, abbreviation for the proteins period, ARNT (aryl hydrocarbon nuclear transporter), and simple; ^{Se}Met, selenomethionine; MAD, multiple-wavelength anomalous dispersion; rmsd, root-mean-square deviation.

10.1021/bi035980p CCC: $27.50
Published on Web 02/20/2004

binding (*26*, *27*). In contrast, the FixLHs were found to be five-coordinate in the absence of exogenous ligands (*12*−*15*). *Ec* DosH reacts with exogenous ligands such as O_2, CO (in the ferrous form), and CN^- (in the ferric form). These differently ligated forms can be distinguished by their optical spectra (*26*, *27*), similar to the FixLHs (*10*, *19*). In the absence of an exogenous ligand, it was reasoned that the *Ec* DosH heme must have a sixth ligand originating from the protein itself (an endogenous ligand). Evidence for this was obtained by site-specific mutagenesis experiments that implicated His77 and Met95 as heme ligands in DosH (*27*, *31*).

To further the structural characterization of *Ec* DosH and provide a basis for subsequent structural work on the full-length *Ec* Dos protein, we have solved the crystal structure of *Ec* Dos heme domain (*Ec* DosH) using an Fe MAD method. Here, we describe the structures of *Ec* DosH−O_2, which contains molecular oxygen as a ligand, and *Ec* DosH, in which there is no oxygen ligand. To confirm the existence of two different ligation states in the asymmetric unit, the ^{Se}Met form of *Ec* DosH was prepared and its structure also solved. The crystal structures of both *Ec* DosH and ^{Se}Met-*Ec* DosH are essentially identical when similarly ligated forms are compared and confirm that *Ec* DosH is related in structure to the FixLH proteins. The two structures, with (oxy monomer) and without (deoxy monomer) O_2 ligation, represent the inactive and active *Ec* DosH protein forms, respectively, and provide further insight into the possible mechanics of the signaling mechanism within the heme domain.

EXPERIMENTAL SECTION

Purification, ^{Se}Met Substitution, and Crystallization. Preparation and crystallization of the full-length Dos heme domain was performed as described previously (*33*). A plasmid harboring the *E. coli* Dos heme domain was used to transform *E. coli* 834(DE3) to make the ^{Se}Met-substituted protein. Cells were grown on M9 minimal media with seleno-L-methionine and a nutrient mixture. ^{Se}Met-substituted protein was purified and crystallized in the same manner as for the native protein (*33*).

Data Collection, Structure Determination, and Analysis. The crystals were flash-frozen under a nitrogen stream at 100 K prior to data collection. An Fe MAD dataset of the native protein was collected at three wavelengths (in Å), 1.741 (peak), 1.7382 (edge), and 1.033 (remote), at the SSRL Beamline 9-2, and data for ^{Se}Met-DosH, 0.9796 (peak) and 0.9183 (remote), was collected at the ALS beamline 5.0.2.

Data were processed and merged with the HKL software suite. The recombinant native *Ec* DosH structure was solved by the Fe MAD method (*33*). Briefly, determination of heavy atom positions, initial phasing, and Cα atom tracing were performed using SOLVE and RESOLVE. The model was built into the electron density using the graphics program O. Iterative rounds of model adjustment using O and refinement using CNS were performed to improve the model. The refined model of *Ec* DosH was transformed to ^{Se}Met-DosH with one round of rigid body refinement, and ^{Se}Met-DosH then went through a separate refinement process using CNS. Data and refinement statistics are summarized in Table 1.

Structure graphics, analysis, and viewing were carried out using the programs O and DS Viewer Pro (Accelrys). Protein structures shown in the figures were superimposed by creating "anchors" at identical positions on the heme of each structure and then superimposing the structure so that the corresponding anchor points in each structure occupied identical coordinates. Typical anchor points were the heme α, *β*, and *δ* meso carbon atoms (Figure 1). In this way, any heme structure changes via ruffled, domed, or saddled distortions would be visible. The Fischer numbering system is commonly used throughout the discussion when referring to the heme (Figure 1). The coordinates have been deposited in the Protein Data Bank: recombinant native DosH, 1S66, and ^{Se}Met DosH, 1S67.

Table 1

	Ec DosH	^{Se}Met-*Ec* DosH
	Data	
space group	$P2_12_12_1$	$P2_12_12_1$
cell dimensions (Å)	$a = 47.189$, $b = 68.310$, $c = 82.809$	$a = 47.575$, $b = 67.690$, $c = 82.352$
asymmetric unit	two molecules	two molecules
resolution (Å)	50−1.8	30−1.5
wavelength (Å)	0.9703	0.9800
total reflns	268 182	185 292
unique reflns	23 838	41 747
completeness (%)[a]	88.9 (73.5)	94.4 (66.1)
R_{merge}[a,b]	0.055 (0.204)	0.055 (0.262)
	Refinement	
resolution (Å)	10−1.8	10−1.55
no. of reflns	23 111	34 562
R_{free}[a,c]	22.4 (24.6)	21.4 (23.2)
R_{cryst}[a,d]	18.9 (22.4)	19.7 (22.9)
rmsd[e]		
bond length (Å)	0.023	0.021
bond angle (deg)	1.69	1.56
	Number of Atoms	
protein and ligand	1943	1943
water	205	207

[a] Parentheses refer to statics for the highest 0.05 Å resolution shell. [b] $R_{merge} = (\sum_h \sum_i |\langle F_h \rangle - F_{hi}| / \sum_h F_h$, where $\langle F_h \rangle$ is the mean structure factor magnitude of *i* observations of symmetry-related reflections with Bragg index *h*. [c] R_{free} is calculated with removal of 5% of the data as the test set at the beginning of refinement. [d] $R_{cryst} = \sum ||F_{obsd}| - |F_{calcd}|| / |F_{obsd}|$. [e] rmsd's for mainchain atoms are the root-mean-squared deviations of the bond lengths and bond angles from their respective ideal values as implemented in CNS.

RESULTS AND DISCUSSION

The crystal structure of the heme-containing, oxygen-sensing domain of the recombinant *E. coli* direct oxygen sensor (*Ec* DosH) has been determined using heme Fe as a MAD scattering atom. This structure, especially the exact position of methiones, was confirmed by subsequently crystallizing the ^{Se}Met-substituted *Ec* DosH and determining its structure. Both native and ^{Se}Met-substituted proteins were crystallized in an orthorhombic space group, $P2_12_12_1$, with two molecules in one asymmetric unit. The average rmsd value of backbone Cα atoms between native (solved by Fe MAD) and ^{Se}Met-substituted *Ec* DosH was 0.278 Å, which emphasizes the similarity of the two structures. Unless otherwise indicated in the text, the results and discussion that follow refer specifically to the recombinant *E. coli* direct oxygen sensor (*Ec* DosH) structure and not the ^{Se}Met-substituted *Ec* DosH.

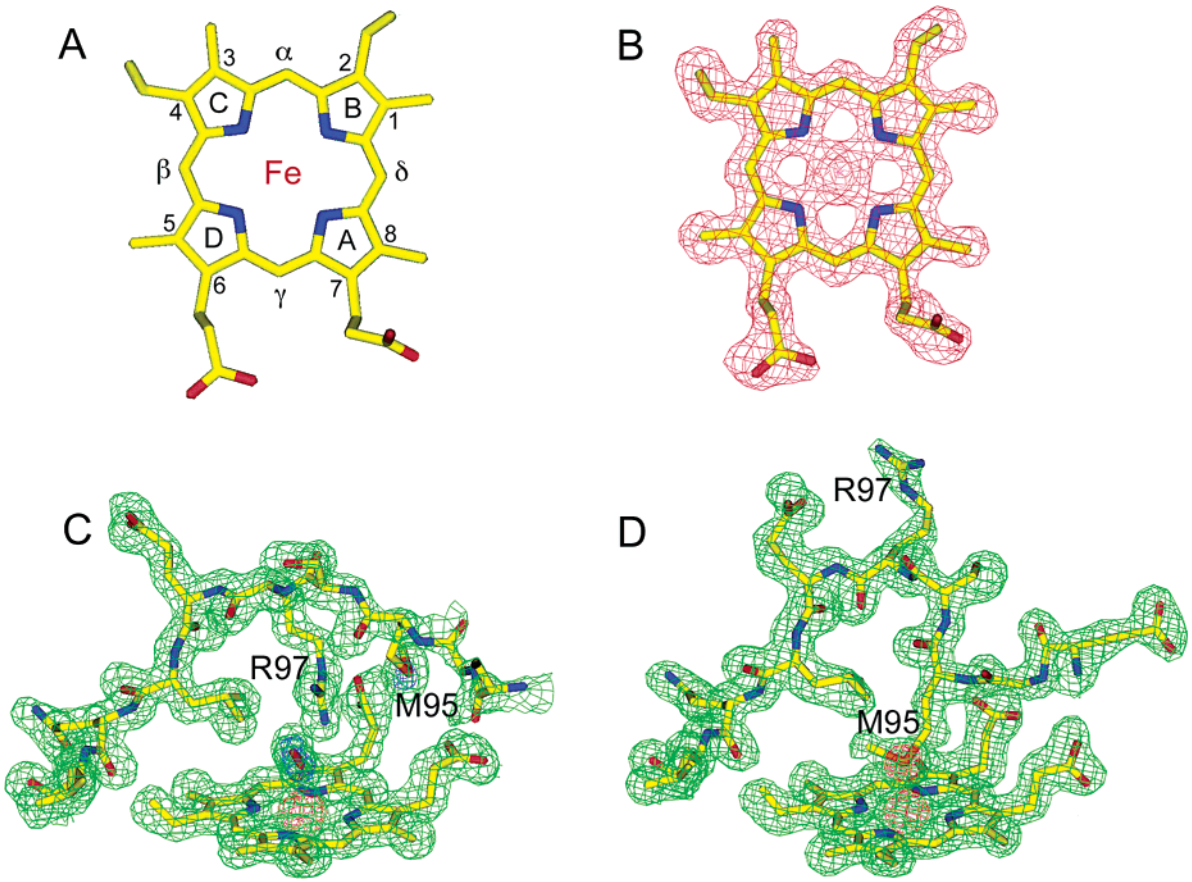

FIGURE 1: Structure (A) of heme-b the DosH prosthetic group. The heme pyrrole rings are labeled using the crystallographic letter notation (A−D) and have the corresponding Fischer numbering A (IV), B (I), C (II), D (III). The meso-carbons and the pyrrole β-carbons are labeled according to the Fischer notation. Panel B shows the electron density map of heme phased using Fe-MAD. Panels C and D are phased using Se-MAD in ^{Se}Met-DosH. Areas around the distal heme pocket of oxy monomer (C) and deoxy monomer (D) are shown contoured at 1σ. The two heme pockets with different ligation state (C, D) are viewed at the same orientation. Red coloring at the heme center in the electron density maps in panels C and D represent heme Fe, which was contoured at 5σ. In panel D, the ligating Se of Met95 is also represented in red and was also contoured at 5σ. In panel C, the blue coloring in the electron density map, representing ligated O_2 and the free Se of Met95, was contoured at 3σ.

The protein backbone was traced, and most of the side chains were fitted using the Fe MAD map. The first and last residue seen in the electron density are, respectively, Asn16 and Ser134 for molecule A (oxy monomer) and Gly20 and Ser134 for molecule B (deoxy monomer). Residues 89−91 of molecule A were also not visible in the electron density. The final *R*-factor is 18.9% (R_{free} = 22.4%) for 23 111 (10−1.8 Å) unique reflections. The root-mean-square deviation from the ideal for bond length is 0.023 Å, and for bond angle, it is 1.69°. There are 205 water molecules included in the final structure. The average *B*-factors of all non-hydrogen atoms (1943 atoms in 231 residues) are 24.5 and 23.9 Å^2 for molecules A and B, respectively. A summary of the crystallographic data is given in Table 1.

Global Structure. The asymmetric units of both crystalline *Ec* DosH samples have two independent molecules. Figure 2 shows the structure of the recombinant heme domain dimer of *E. coli* direct oxygen sensor (*Ec* DosH). However, the axial ligands of the individual heme domains that constitute a dimer are different. While both heme domains in the dimer contain His77 as the proximal heme ligand, what distinguishes them is that one monomer in each dimer contains heme-bound O_2 as the sixth heme ligand, while in the other, the sixth ligand is Met95. Thus, each dimer contains both the O_2-ligated heme domain (oxy monomer),

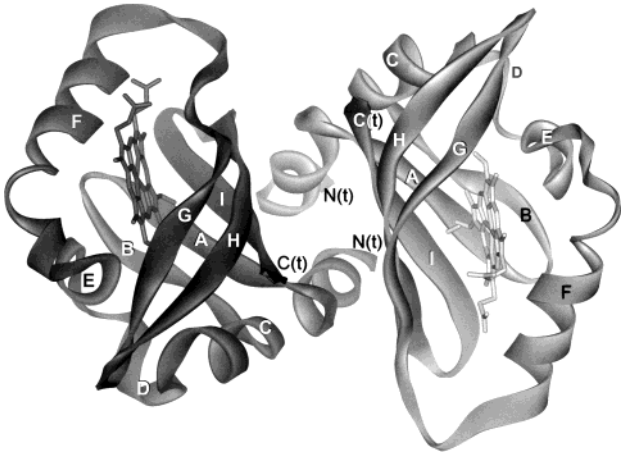

FIGURE 2: Ribbon structure of the recombinant native *Ec* DosH (heme domain) dimer. The lighter structure (right) is the domain with heme-bound O_2, whereas the darker structure (left) is the domain with heme-bound Met95. Secondary structure elements are labeled alphabetically beginning with the amino-terminal β-strand (A).

representing the inactive protein, and the internally ligated Met95 form, representing the active protein. While somewhat unusual, this type of observation has literature precedent (*34*−*37*).

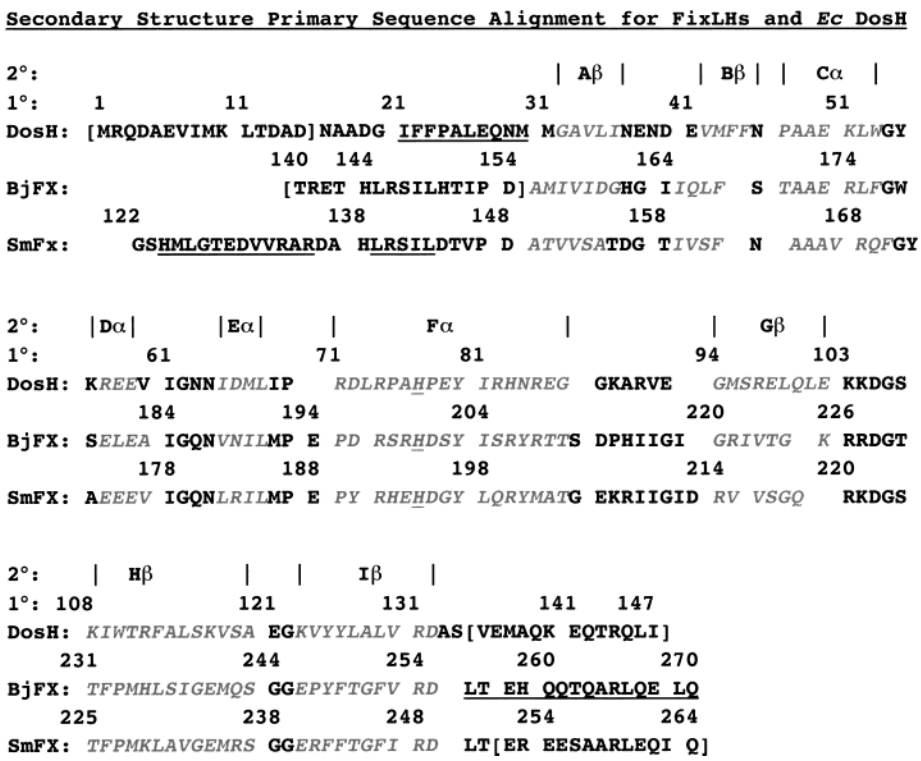

FIGURE 3: Alignment of *Ec* DosH (DosH), *Bj* FixLH (BjFX), and *Sm* FixLH (SmFX) according to secondary structure features defined by DSSP (38). The secondary structure elements are indicated in the top row (2°) of each group, defined by vertical lines, and occur as lighter face type in the individual sequences. Primary sequence (1°) numbers for each protein sequence are shown above each sequence. Brackets surround N-terminal and C-terminal sequence regions of the individual heme domains that were not identifiable in the crystal structures (*12*, *15*, this work). Underlined portions indicate N-terminal and C-terminal helices that are not part of the $\alpha_4\beta_5$ PAS core domain. The heme-coordinating proximal histidine in the middle of each F_α helix is indicated by underline.

Verifying the nature of the different axial heme ligands in the two asymmetric unit proteins was carried out as follows. Sulfur atom positions of all the methionine residues, including Met95 in both ligation states were confirmed by crystallizing the ^{Se}Met-substituted *Ec* DosH and determining its structure (Figure 1D).

From the early stage of refinement, the $F_o - F_c$ map clearly showed electron density corresponding to O_2 at the distal side of the heme in one of the monomers. In this native oxy monomer, the refined *B*-value for the liganded O_2 is very similar to those for the carbonyl oxygen atoms of amino acid residues in the structure. The ligand coordination geometry of the oxy monomer structure also corresponds most closely to the O_2-ligated *Bj* FixLH structure but not to the structures of *Bj* FixLH ligated with other diatomic molecules (CO, CN^-, or NO) (*12*, *13*).

Both heme domains in the dimer are therefore six-coordinate, as predicted from optical spectroscopy measurements (*26*). This fortuitous circumstance allows us to simultaneously analyze the subunit structures for both active and inactive *Ec* DosH.

Figure 2 shows the three-dimensional dimer structure of recombinant native *Ec* DosH with the individual protein chains drawn as ribbons and with secondary structure elements labeled according to a previously proposed notation (*26*). In this labeling system, each distinct secondary structure element is consecutively labeled, A, B, C, ..., beginning with the β-strand situated closest to the N-terminus. The specific structure element type (α-helix, β-strand) is denoted as a subscript. Each *Ec* DosH monomer bears a striking global structural similarity to the previously characterized heme oxygen-sensing proteins, *Bj* FixLH (*3*, *12*−*14*) and *Sm* FixLH (*15*).

Figure 3 summarizes additional comparisons between these three proteins that are the results of structure-based primary sequence alignments. These results come from analysis of the crystal structure of each protein using the program DSSP (*38*) to accurately define and identify specific secondary structure components.

It can be seen from Figures 2 and 3 that each *Ec* DosH monomer contains five distinct α-helical regions; four are labeled C−F, and another, unlabeled, comprises the N-terminus. There are also five separate β-strand regions, labeled A, B, G, H, and I. This $\alpha_5\beta_5$ structure is also typical of *Bj* FixLH and *Sm* FixLH. The comparison shown in Figure 3, and the structure shown in Figure 2 confirm that the *Ec* DosH monomer is a PAS domain protein, similar to the FixLHs. All of these proteins contain the PAS core structure $\beta_2\alpha_4\beta_3$. In addition to the PAS $\beta_2\alpha_4\beta_3$ core both *Ec* DosH and *Sm* FixLH contain one or more N-terminal helices, while *Bj* FixLH contains a C-terminal helix.

The dimer structure presented in Figure 2 shows the heme groups separated by ~30 Å with an interface that includes the most hydrophobic surface of each *Ec* DosH monomer. The interface involves the N-terminal helix of each monomer, which together are juxtaposed, as well as the C-terminus of each subunit. The striking part about this interface is that in these features it is highly similar to the dimer interface described for the *Sm* FixLH dimer structure (*15*, *27*).

The hemes in both monomers within the *Ec* DosH dimer are sandwiched between the longest helical segment, F_α, on the heme proximal side, and two β strands, G_β and H_β, on

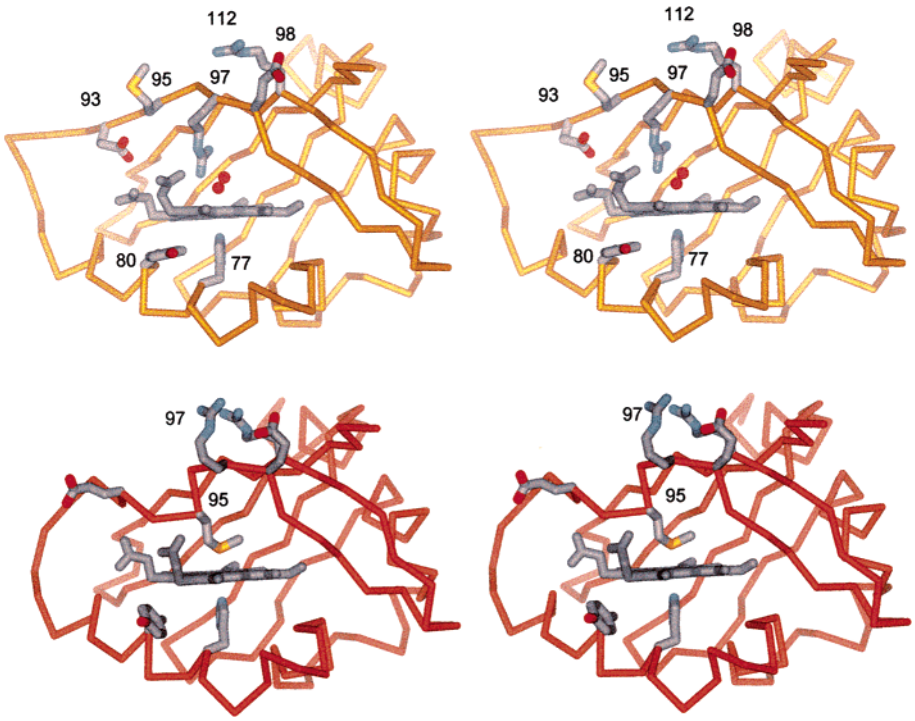

FIGURE 4: The heme environment in the recombinant native *Ec* DosH−O_2 heme domain subunit (yellow, top) and in the recombinant native *Ec* DosH subunit (brown, bottom) presented in stereo. The heme is viewed edge-on with the distal ligand-binding side above the heme plane and the proximal side below the heme plane, showing His77 coordinated to the heme iron ion. Ligated O_2 is shown as a pair of red spheres in the top view. Side chains of several key amino acids closest to the heme, which comprise the heme environment, are shown. Amino acids are numbered and can be identified from the *Ec* DosH sequence shown in Figure 3.

the heme distal (ligand-binding) side. In both monomers, the *Ec* DosH heme is coordinated by His77, whereas distal ligation employs an endogenous Met95 ligand in the absence of an exogenous ligand (Figures 1 and 4). Figure 2 shows that one side of the heme, pyrroles B and C (I and II, respectively, in Fischer notation), is solvent-shielded by additional β strands (A_β, B_β, I_β), whereas the side of the heme with the two propionate substituents, consisting of pyrroles A and D (pyrroles IV and III in Fischer notation) is quite solvent-exposed. Both of the heme propionates form hydrogen bonds directly to water molecules, as well as to other parts of the protein structure (vide infra).

Heme Pocket Environment. Heme pocket environments for each of the differently ligated native *Ec* DosH monomers are shown in Figure 4. The presence or absence of O_2 as a heme ligand causes significant rearrangements of many amino acid side chains that comprise the heme environment, which will be described in detail in the next section. However, the general environment around the hemes in both ligation states is very hydrophobic with nearest neighbors to the heme being isoleucines 36, 65, 69, and 81, leucines 73, 99, and 115, valine 92, asparagine 100, tyrosines 80 and 126, and tryptophan 53 (not shown). Such a hydrophobic heme environment excludes water on the distal or ligand-binding side of the heme.

Structure Differences with and without O_2 Ligation. There are significant and highly localized conformational differences in the global architecture of the two differently ligated monomer heme domains in the DosH structure, as shown in Figures 5 and 6. Visual inspection of the superimposed structures of the differently ligated heme domains, shown in Figure 5A,B, reveals that three regions are affected by the presence or absence of heme-coordinated O_2. These are

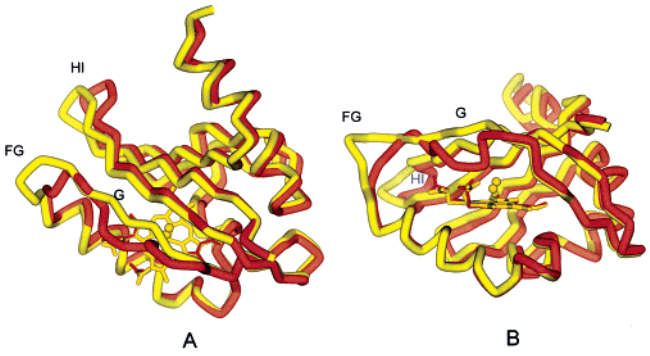

FIGURE 5: Two views of the superimposed global structures of the individual heme domains: *Ec* DosH−O_2 (yellow) and *Ec* DosH (brown). The protein chain of each is shown as a tube. Panel A presents a view looking perpendicular to the heme from the G_β−H_β−I_β strand side (distal side). Panel B presents a view looking toward the heme edge with the F_α helix and the proximal side below the heme and the G_β strand above the distal side of the heme. The O_2 ligand in *Ec* DosH−O_2 is shown as a pair of spheres on the distal side of the heme. No other heme ligands are shown (see Figure 4).

the FG turn (residues 88−93), the early G_β strand (residues 94−97), and the HI turn (residues 121−122). These structural differences are quantified by the pairwise protein chain Cα rmsd comparisons in Figure 6C.

Several of the specific amino acid rearrangements in the *Ec* DosH heme pocket due to the presence or absence of O_2 are shown in Figure 4. The changes are most dramatic on the distal side of the heme where the heme ligand is different. The most significant change involves the heme ligand, which is O_2 in *Ec* DosH−O_2. In the absence of an O_2 ligand, the sixth heme coordination site is occupied by Met95, with S → Fe coordination, Figures 1D and 4.

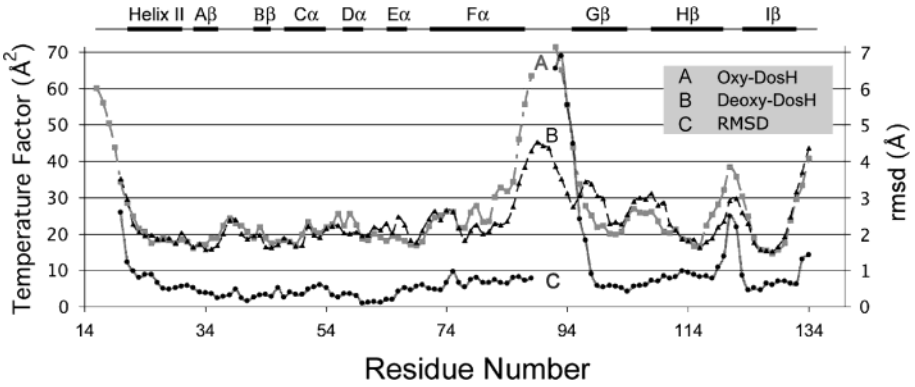

FIGURE 6: Graphs of the backbone Cα atom temperature factors of (A) oxy monomer and (B) deoxy monomer *Ec* DosH (values on left *y*-axis) and (C) the backbone Cα atom rmsd values (right *y*-axis) between the *Ec* DosH oxy monomer and deoxy monomer as a function of primary sequence. The secondary structure regions are shown above the plotted values. Elevated temperature factors and rmsd values were observed in the FG and HI loops of both monomers. Temperature factors are generally lower in the deoxy monomer than in the oxy monomer, except for residues between 97 and 111.

O_2 coordination is accompanied by a nearly 180° rotation of the side chain of Arg97. In the absence of O_2 ligation, the Arg97 side chain is oriented toward the protein's surface and participates in a salt bridge with Arg112 and Glu98 (Figure 4, lower). In the oxy monomer, the Arg97 side chain reorients inward to the heme distal pocket to provide stabilization to heme-coordinated O_2 (Figure 4, upper). Its role is that of a hydrogen bond donor, displaying Arg-N···O- (heme-bound) distances of 2.70 and 2.94 Å. The Arg97 motion represents an 11.67 Å change in position for the guanidinium group.

The side chain of Met95 also adopts radically different orientations in the two ligation states of *Ec* DosH. The difference is also a nearly 180° rotation from heme-coordinated (Figure 4, lower) to pointing toward the protein surface (Figure 4, upper). This motion involves a 10.91 Å change for the Met95 sulfur atom. Whereas both Met95 and Arg97 side chains make similarly large ligation-dependent conformational changes, their effects on the protein main chain is quite different with Met95 distorting the chain significantly. This can be seen by comparing Cα coordinate differences (ΔCα) in the two *Ec* DosH ligation states. For Met95, ΔCα is 5.48 Å, while it is only 1.70 Å for Arg97. As shown in Figures 5 and 6, the effect of the Met95 change is transmitted further along the chain to the area between residues 89 and 96. This also results in a major distortion of the FG turn, including an order−disorder transition in residues 89−91. This structural distortion is demonstrated by the main chain comparison in Figure 5 and the rmsd graph in Figure 6 and is also reflected in the side chain displacements shown in Figure 4 (e.g., Glu93). Amino acid Cα position differences for the two native *Ec* DosH ligation states are 4.98 Å for Lys87 and 4.83 Å for Ala88 (FG turn).

A more subtle difference is that the change in ligation state also affects the structural stability of an area around the two amino acids directly involved in ligand binding (residues 95 and 97). This is indicated by the B-factor plots shown in Figure 6A,B).

Residues between 86 and 95 including the FG loop are stabilized by Met95 ligation to the heme in the deoxy monomer. In contrast, the same area in the O_2-ligated monomer is not stabilized efficiently by Arg97, even though this residue interacts with the ligand.

Because of this, residues 86−95 in the oxy monomer display high-temperature factors and the central residues 89−91 become disordered. This order−disorder transition in residues 89−91 is shown in the *B*-factor plots in Figure 6. On the other hand, the adjacent area of the oxy monomer, between residues 97 and 111, becomes more stable than the comparable region of the deoxy monomer because of indirect anchoring of Arg97.

Although the major ligand-dependent changes just described occur on the heme distal side, there are less significant changes for amino acids lying in the heme plane and on the heme proximal side. As shown in Figure 4, the most significant of these is the ∼115° ring rotation of Tyr80.

The Fe−N_ϵ bond distance between the heme iron and the imidazole group of His77 at the fifth coordination site is 1.98 and 1.95 Å for oxy-DosH and deoxy-DosH, respectively. The distances are slightly shorter than the Fe−N_ϵ bond distance previously reported for the *Bj* FixL heme domain structures (*13*). The orientation of the imidazole group is fixed by a water-mediated hydrogen bond between N_δ of His77 and O_δ of Asp66 in *Ec* DosH−O_2. Two water molecules are involved in a similar hydrogen bond mediation in the deoxy monomer. Water molecules at similar positions are also found in the FixL structures.

Comparison of DosH and Bj FixLH Heme Pockets. Comparisons of the two types of heme pockets, as the locus of signal initiation in the FixLHs and in *Ec* DosH, are instructive and provide a basis for structure/function analysis. Figure 7 shows the combined heme pockets of *Ec* DosH−O_2 (red) and *Bj* FixLH−O_2 (blue) generated by superimposing the heme group of each protein. Both heme-sensor domains contain histidine and molecular oxygen as heme axial ligands. Without O_2 ligation (not shown), *Bj* FixLH is five-coordinate and high-spin, lacking a ligand at the sixth coordination position, but *Ec* DosH is six-coordinate with Met95 ligation (Figure 1D and 4).

Similarities in the *Bj* FixLH and *Ec* DosH ligand binding site include the presence of arginine stabilizing heme-coordinated O_2. This is shown by the positions of Arg97 (red, *Ec* DosH) and Arg220 (blue, *Bj* FixLH) in Figure 7 and is presumed to occur also in *Sm* FixLH, which has the corresponding Arg214, although no structure of *Sm* FixLH−O_2 has yet been published.

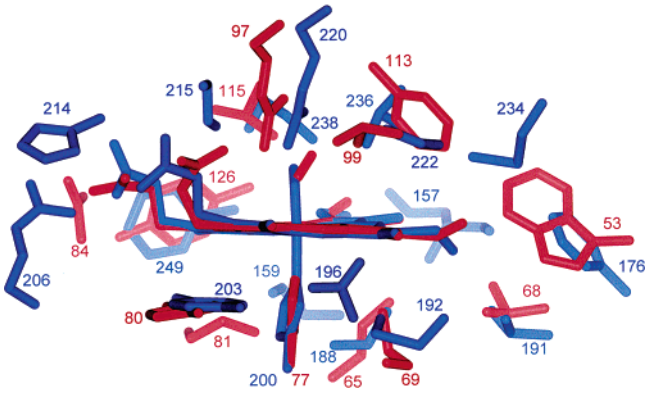

FIGURE 7: Combined view of the heme environments for *Ec* DosH−O_2 (red) and *Bj* FixLH−O_2 (blue), created by superimposing the heme groups of each protein. Amino acids are numbered so that they can be identified from the primary sequences shown in Figure 3.

The nonpolar nature of the heme pocket in FixLHs is due to a "hydrophobic triad" of similarly situated distal amino acid side chains: Ile215(*Bj*)/Ile209(*Sm*), Leu236(*Bj*)/Leu230(*Sm*), and Ile236(*Bj*)/Val232(*Sm*). In *Ec* DosH, there is no amino acid that spatially corresponds to Ile215(*Bj*)/Ile209(*Sm*), but Phe113 and Leu115 correspond, respectively, to the other two members of the "hydrophobic triad" of the FixLHs. Leu99 in *Ec* DosH serves as a third hydrophobic heme contact in DosH, but it is situated more closely to the position occupied by Val222 in *Bj* FixLH than to the Ile215/Ile209 position (Figure 7). Other corresponding hydrophobic side chains can be seen in Figure 7, including Ile65(*Ec* DosH)/Val88(*Bj*), Ile69(*Ec* DosH)/Met192(*Bj*), Tyr80(*Ec* DosH)/Tyr203(*Bj*), Tyr126(*Ec* DosH)/Phe249(*Bj*), Leu68(*Ec* DosH)/Leu191(*Bj*), and Trp53(*Ec* DosH)/Phe176(*Bj*).

Comparison of Ec DosH and Bj FixLH Ligation-Dependent Structure Changes. Comparing the structural changes in native *Ec* DosH and *Bj* FixLH that depend on the presence or absence of an O_2 ligand can delineate the conformational changes that accompany signal initiation. For *Ec* DosH, significant O_2-dependent structural changes appear to be localized in the FG turn, early G_β strand, and HI turn, as shown in Figure 5. However, closer inspection of the coordinates reveals that the distortion of the FG turn is essentially continuous into the G_β strand, as illustrated by the rmsd differences shown in Figure 6C. This distortion can also be demonstrated by the corresponding amino acid differences between the O_2-ligated and Met95-ligated structures, represented by ΔCα in the FG/G_β region. At the end of the F_α helix, ΔCα is 0.73 Å for Arg82 but jumps to 6.70 Å for Val92 in the middle of the FG turn. In the G_β strand, the distortion continues, being 5.48 Å for Met95, 2.58 Å for Ser96 and 1.70 Å for Arg97.

By comparison, the O_2-induced structural changes in *Bj* FixLH are less extensive and of lower magnitude than those in *Ec* DosH. Figure 8 shows superimposed structures of O_2-ligated *Bj* FixLH (yellow, pdb; 1DP6) and ferrous unligated *Bj* FixLH (blue, pdb; 1LSW), analogous to the views of *Ec* DosH presented in Figure 5. It is immediately obvious that any structural differences in the FG and HI turn regions and the G_β strand of FixLH are not as extensive as those in *Ec* DosH. For example, in the FG turn, the maximum O_2-induced ΔCα differences for *Bj* FixLH are 1.22 Å (Ser211) and 1.30 Å (Pro213). Thereafter in the sequence, chain

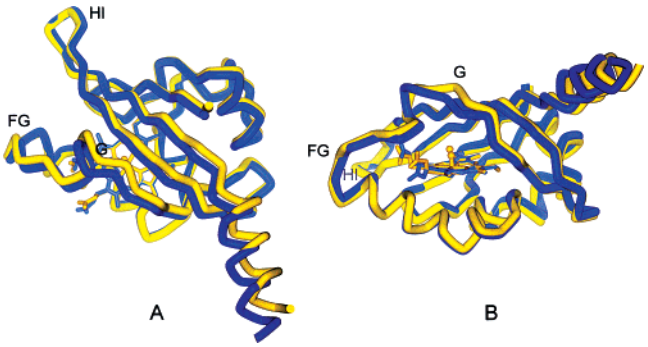

FIGURE 8: Two views of the superimposed global structures of the individual heme domains, *Bj* FixLH−O_2 (yellow) and deoxy *Bj* FixLH (blue). The protein chain of each is shown as a tube, and these views are analogous to those shown for *Ec* DosH in Figure 5. Panel A presents a view looking perpendicular to the heme from the G_β−H_β−I_β strand side (distal side). Panel B presents a view looking toward the heme edge with the F_α helix and the proximal side below the heme and the G_β strand above the distal side of the heme. The O_2 ligand in *Bj* FixLH−O_2 is shown as a pair of spheres on the distal side of the heme. No other heme ligands are shown.

overlap in the two forms improves until the early G_β strand where the coordinates diverge with ΔCα being 2.06 Å (Gly217), 2.31 Å (Ile218), 2.35 Å (Gly219), and 1.99 Å (Arg220). Except for Arg220, these are all much smaller changes than those found for their structurally corresponding *Ec* DosH amino acids. The Arg220 side-chain reorientation upon O_2-ligation is also much smaller than that found for Arg97 in *Ec* DosH. The terminal guanidinium group in Arg220 is required to move only 5.40 Å to stabilize heme-bound O_2 in *Bj* FixLH−O_2.

A proposed key element in signal initiation in *Bj* FixLH was an observed change in heme planarity accompanied by an alteration in the hydrogen bonding experienced by the heme propionates resulting from the ligation change (e.g., dissociation of heme-bound O_2). In *Bj* FixLH−O_2, the heme 7-propionate is hydrogen-bonded to a water molecule and His214, while the heme 6-propionate is hydrogen-bonded to Arg206 and His214 (see Figure 5 in ref *14*). After O_2 dissociates, thereby initiating signal transduction, the 7-propionate ends up hydrogen-bonded to both His214 and Arg220, while the 6-propionate hydrogen bonds to His214 and a water molecule. Judging from the two *Bj* FixLH crystal structures (pdb; 1DP6 and 1LSW), the results of this shift in hydrogen bonding are changes in each propionate's orientation. This is slight for the 6-propionate, but for the 7-propionate, there is a definite change from an "out-of-heme-plane" conformation to a more extended "in-heme-plane" orientation, as shown in Figure 9.

In *Ec* DosH, there are no amino acids structurally analogous to Arg206 and His214 in *Bj* FixLH. In *Bj* FixLH, these two amino acids also help to shield the propionate edge of the heme from solvent exposure. As already pointed out, this means that the *Ec* DosH propionate heme edge is much more solvent-exposed. As shown in Figure 9, both heme propionates in *Ec* DosH adopt "out-of-plane" conformations in both ligation states. In *Ec* DosH−O_2, only one water molecule lies within hydrogen-bonding distance of the 7-propionate, which is also bent away from the 6-propionate (Figure 9). This conformation allows the 7-propionate carboxyl terminus, O2A, to come within 3.04 Å of the Arg97-N_ϵ. Taken together, this combination of orientation

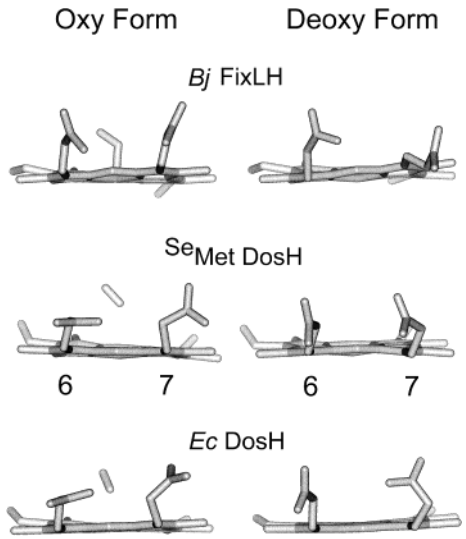

FIGURE 9: Identical edge-on views of the heme groups of the three indicated structures. The heme edge with the 6-propionate (D or III pyrrole) and 7-propionate (A or IV pyrrole) positions (labeled on the middle structure) face the viewer. Hemes derived from the structures of proteins with O_2-ligation (oxy) and without O_2-ligation (deoxy) are shown.

and separation distance indicate a moderately strong hydrogen bond between the 7-propionate and Arg97 (*39*), which could help stabilize Arg97 within the hydrophobic heme pocket, perhaps through charge neutralization. The 6-propionate hydrogen-bonds to Asn84 and a water molecule in *Ec* DosH−O_2.

When the O_2 ligand is not present, the 7-propionate conformation changes so that its carboxylate group points toward the 6-propionate (Figure 9), and it hydrogen bonds to the Met95 peptide N−H as a consequence of the large polypeptide chain deformation accompanying Met95 heme ligation. The 7-propionate carboxylate O1A is separated from the Met95-N_p by only 2.98 Å. Two water molecules are within hydrogen bond distance as well. The 6-propionate is within hydrogen-bonding distance of Asn84 (N_δ), the peptide NH of Gly94, and a water molecule. There are no resolved water molecules in the *Ec* DosH heme pocket, in either ligation state, which is another difference with the FixLHs.

Conformational Changes and Mechanism. As previously described, in *Ec* DosH large O_2-dependent conformational changes are found in the FG turn and the G_β strand. These changes alone are large enough to serve as a conformational mechanism for initial signal transduction, but we also find significant changes in the HI turn (Figure 5). The two primary amino acids in this turn, Glu121 and Gly122, display substantial $\Delta C\alpha$ of 4.05 and 3.49 Å, respectively, between the two ligation states. In contrast, for the two *Bj* FixLH ligation states (Figure 8), only small differences in Cα coordinates are found for the HI turn. $\Delta C\alpha$ is only 0.47 and 0.55 Å for Gly244 and Gly245, the two corresponding primary amino acids in the HI turn.

If the initial signal is conveyed away from the heme-O_2 site by conformational changes, then there are three localized but adjacent regions in *Ec* DosH that suggest the transduction pathway: FG and HI turns and G_β strand. The adjacent spatial localization of these three regions can be seen in Figure 5A,B. Also obvious in Figure 5A is that the FG and HI regions lie farther apart from each other when molecular oxygen is not the ligand (deoxy). Thus, O_2 dissociation from the heme, which is the signal initiating event, results in opening a cleft between the FG and HI turns and extending the separation further between the G_β and H_β strands. For this area to represent the signal transduction pathway requires that the phosphodiester catalytic domain of full-length *Ec* Dos must interact through these regions with the heme domain. Consequently, for *Ec* Dos, the signal is likely to be propagated from the ligand-binding site by initial dissociation of O_2 through Met95/Arg97 rearrangements to immediate neighbors in the heme distal pocket, which displace the FG and HI turns.

This proposed mechanism is somewhat different from that envisioned for FixL, similarly formulated from the FixLH structures, where distal heme pocket participation was minimized in favor of a proposal that combined heme flattening and propionate hydrogen bond rearrangements (vide supra; *12*−*14*). Whereas the *Ec* DosH heme propionates experience hydrogen bonding differences between the two ligation states similar to those identified in *Bj* FixLH, to us these seem to be incidental to the dramatically larger heme distal pocket changes. Moreover, as shown in Figure 8 and described above, comparing the *Bj* FixLH oxy and deoxy structures reveals that there are only very slight O_2-dependent differences in the FG and HI turns. As in *Ec* DosH, the early G_β strand is where the most significant structural change occurs, but in *Bj* FixLH, this is the only area in the PAS domain core exhibiting other than minor structural changes.

It may be that signal propagation is different in FixLs and *Ec* Dos, perhaps because of the different catalytic domains. It may also be that crystal packing forces in *Bj* FixLH inhibit the true extent of O_2-dependent conformational rearrangement when the deoxy monomer is generated in situ from the *Bj* FixLH−O_2 crystal. Further work on these questions is in progress.

ACKNOWLEDGMENT

J.D.S. acknowledges Harold M. Goldberg, MD, Michael Kwasman, MD, Lawrence Hammond, MD, and Stacey Dean, MD, without whose skill and compassion this work would not have been possible. Support of the Murdock Charitable Trust for the BXC Center is also acknowledged. We gratefully acknowledge warm conversations with Professor C. S. Raman and Dr. Dong-Sun Lee at the University of Texas Medical School at Houston who independently and simultaneously solved several *Ec* DosH structures, even though our work may precede theirs in print.

REFERENCES

1. Rodgers, K. R. (1999) Heme-based sensors in biological systems, *Curr. Opin. Chem. Biol. 3*, 158−167.
2. Pellequer, J.-L., Brudler, R., and Getzoff, E. D. (1999) Biological sensors: More than one way to sense oxygen, *Curr. Biol. 9*, R496−R498.
3. Chan, M. K. (2001) Recent advances in heme-protein sensors, *Curr. Opin. Chem. Biol. 5*, 216−222.
4. David, M., Daveran M.-L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., and Kahn, D. (1988) Cascade regulation of nif gene expression in *Rhizobium meliloti*, *Cell 54*, 671−683.
5. de Philip, P., Batut, J., and Boistard, P. (1990) *Rhizobium meliloti* FixL is an oxygen sensor and regulates *Rhizobium meliloti* nifA and fixK genes differently in *Escherichia coli*, *J. Bacteriol. 172*, 4255−4262.

6. Gilles-Gonzalez, M.-A., Ditta, G. S., and Helinski, D. R.(1991) A hemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*, *Nature 350*, 170−172.
7. Anthamatten, D., and Hennecke, H. (1991) The regulatory status of the FixL-like and FixJ-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*, *Mol. Gen. Genet. 225*, 38−48.
8. Monson, E. K., Weinstein, M., Ditta, G. S., and Helinski, D. R. (1992) The FixL protein of *Rhizobium meliloti* can be separated into a heme binding oxygen sensing domain and a functional C-terminal kinase domain, *Proc. Natl. Acad. Sci. U.S.A. 89*, 4280−4294.
9. Lois, A. F., Ditta, G. S., and Helinski, D. R. (1993) The oxygen sensor FixL of *Rhizobium meliloti* is a membrane protein containing four possible transmembrane segments, *J. Bacteriol. 175*, 1103−1109.
10. Gilles-Gonzalez, M.-A., Gonzalez, G., Perutz, M. F., Kiger, L., Marden, M. C., and Poyart, C. (1994) Heme based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation, *Biochemistry 33*, 8067−8073.
11. Monson, E. K., Ditta, G. S., and Helinski, D. R. (1995) The oxygen sensor protein FixL of *Rhizobium meliloti*—role of histidine residues in heme binding, phosphorylation and signal transduction, *J. Biol. Chem. 270*, 5243−5250.
12. 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.
13. 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.
14. 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.
15. Miyatake, H., Mukai, M., Park, S.-Y., Adachi, S.-I., 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.
16. Rodgers, K. R., Lukat-Rodgers, G. S., and Barron, J. A. (1996) Structural basis for ligand discrimination and response initiation in the heme-based oxygen sensor FixL, *Biochemistry 35*, 9539−9548.
17. Miyatake, H., Mukai, M., Adachi, S., Nakamura, H., Tamura, K., Iizuka, T., and Shiro, Y. J. (1999) Iron coordination structures of oxygen sensor FixL characterized by FeK-edge extended X-ray absorption fine structure and resonance Raman spectroscopy, *J. Biol. Chem. 274*, 23176−23184.
18. Tuckerman, J. R., Gonzalez, G., Dioum, E. M., and Gilles-Gonzalez, M.-A. (2002) Ligand and oxidation-state specific regulation of the heme-based oxygen sensor fixL from *Sinorhizobium meliloti*, *Biochemistry 41*, 6170−6177.
19. Gilles-Gonzalez, M.-A., Gonzalez, G., and Perutz, M. F. (1995) Kinase activity of oxygen sensor FixL depends on the spin state of its heme iron, *Biochemistry 34*, 232−236.
20. Mansy, S. S., Olson, J. S., Gonzalez, G., and Gilles-Gonzalez, M.-A. (1998) Imidazole is a sensitive probe of steric hindrance in the distal pockets of oxygen-binding heme proteins, *Biochemistry 37*, 12452−12457.
21. Bertolucci, C., Ming, L.-J., Gonzalez, G., and Gilles-Gonzalez, M.-A. (1996) Assignment of the hyperfine-shifted ^{1}H NMR signals of the heme in the oxygen sensor FixL from *Rhizobium meliloti*, *Chem. Biol. 3*, 561−566.
22. Akimoto, S., Tanaka, A., Nakamura, K., Shiro, Y., and Nakamura, H. (2003) O_2-specific regulation of the ferrous heme-based sensor kinase FixL from *Sinorhizobium meliloti* and its aberrant inactivation in the ferric form, *Biochem. Biophys. Res. Commun. 304*, 136−142.
23. Tuckerman, J. R., Gonzalez, G., and Gilles-Gonzalez, M.-A. (2001) Complexation precedes phosphorylation for two-component regulatory system fixL/fixJ of *Sinorhizobium meliloti*, *J. Mol. Biol. 308*, 449−455.
24. Satterlee, J. D., Suquet, C. M., Savenkova, M. I., and Lian, C. (2003) Proton NMR characterization of recombinant ferric heme domains of the oxygen sensors FixL and Dos: Evidence for protein heterogeneity, in *Paramagnetic Resonance of Metallobiomolecules* (Telser, J., Ed.) ACS Symposium Series 858, pp 244−257, American Chemical Society, Washington, DC.
25. Dunham, C. M., Dioum, E. M., Tuckerman, J. R., Gonzalez, G., Scott, W. G.l, and Gilles-Gonzalez, M.-A. (2003) A distal arginine in oxygen-sensing Heme-PAS domains is essential to ligand binding, signal transduction, and structure, *Biochemistry 42*, 7701−7708.
26. Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M.-A. (2000) Dos, a heme-binding PAS protein from *Escherichia coli*, is a direct oxygen sensor, *Biochemistry 39*, 2685−2691.
27. Sasakura, Y., Hirata, S., Sugiyama, S., Suzuki, S., Taguchi, S., Watanabe, M., Matsui, T., Sagami, I., and Shimizu, T. (2002) Characterization of a direct oxygen sensor heme protein from *Escherichia coli* − Effects of the heme redox states and mutations at the heme-binding site on catalysis and structure, *J. Biol. Chem. 277*, 23821−23827.
28. Watanabe, M., Matsui, T., Sasakura, Y., Sagami, I., and Shimizu, T. (2002) Unusual cyanide bindings to a heme-regulated phosphodiesterase from *Escherichia coli*: effect of Met95 mutations, *Biochem. Biophys. Res. Commun. 299*, 169−172.
29. 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.
30. Sato, A., Sasakura, Y., Sugiyama, S., Sagami, I., Shimizu, T., Mizutani, Y., and Kitagawa, T. (2002) Stationary and time-resolved resonance Raman spectra of His(77) and Met(95) mutants of the isolated heme domain of a direct oxygen sensor from *Escherichia coli*, *J. Biol. Chem. 277*, 32650−32658.
31. Gonzalez, G., Dioum, E. M., Bertolucci, C., Tomita, T., Ikeda-Saito, M., Cheesman, M. R., Watmough, N. J., and Gilles-Gonzalez, M.-A. (2002) Nature of the displaceable heme-axial residue in the EcDos protein, a heme-based sensor from *Escherichia coli*, *Biochemistry 41*, 8414−8421.
32. Gilles-Gonzalez, M.-A., and Gonzalez, G. (1993) Regulation of the kinase activity of heme protein FixL from the two-component system FixL/FixJ of *Rhizobium meliloti*, *J. Biol. Chem. 268*, 16293−16297.
33. Park, H., Suquet, C. M., Savenkova, M., Satterlee, J. D., and Kang, C. H. (2002) Cloning, purification, crystallization and preliminary X-ray analysis of DOS heme domain, a new heme oxygen sensor in *Escherichia coli*, *Acta Crystallogr. D58*, 1504−1506.
34. Zhang, W., and Phillips, G. N., Jr. (2003) Structure of the oxygen sensor in *Bacillus subtilis*: signal transduction of chemotaxis by control of symmetry, *Structure 11*, 1097−1110.
35. Tan, S., and Richmond, T. J. (1998) Crystal structure of the yeast MAT alpha 2/MCM1/DNA ternary complex, *Nature 391*, 660−666.
36. Iwata S., Kamata K., Yoshida S., Minowa T., and Ohta T. (1994) T and R states in the crystals of bacterial L-lactate dehydrogenase reveal the mechanism for allosteric control, *Nat. Struct. Biol. 1*, 176−185.
37. Carrell, R. W., Stein, P. E., Fermi, G., and Wardell, M. R. (1994) Biological implications of a 3-angstrom structure of dimeric antithrombin, *Structure 2*, 257−270.
38. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical structures, *Biopolymers 22*, 2577−2637; implemented online at http://bioweb.pasteur.fr/seqanal/interfaces/dssp-simple.html.
39. Jeffrey, G. A. (1997) *An Introduction to Hydrogen Bonding*, Oxford University Press, London.

BI035980P