Structural Insights into Oxygen-Dependent Signal Transduction within Globin Coupled Sensors

Shannon Rivera, †,§ Paul G. Young, †,§ Eric D. Hoffer, † Gregory E. Vansuch, † Carmen L. Metzler, † Christine M. Dunham, † and Emily E. Weinert †,‡

†Department of Chemistry, Emory University, Atlanta, Georgia 30322, United States
‡Department of Biochemistry, Emory University, Atlanta, Georgia 30322, United States

Supporting Information

ABSTRACT: In order to respond to external stimuli, bacteria have evolved sensor proteins linking external signals to intracellular outputs that can then regulate downstream pathways and phenotypes. Globin coupled sensor proteins (GCSs) serve to link environmental O₂ levels to cellular processes by coupling a heme-containing sensor globin domain to a catalytic output domain. However, the mechanism by which O₂ binding activates these proteins is currently unknown. To provide insights into the signaling mechanism, two distinct dimeric complexes of the isolated globin domain of the GCS from Bordetella pertussis (BpeGlobin) were solved via X-ray crystallography in which differences in ligand-bound states were observed. Both monomers of one dimer contain Fe(II)−O₂ states, while the other dimer consists of the Fe(III)−H₂O and Fe(II)−O₂ states. These data provide the first molecular insights into the heme pocket conformation of the active Fe(II)−O₂ form of these enzymes. In addition, heme distortion modes and heme−protein interactions were found to correlate with the ligation state of the globin, suggesting that these conformational changes play a role in O₂-dependent signaling. Fourier transform infrared spectroscopy (FTIR) of the full-length GCS from B. pertussis (BpeGReg) and the closely related GCS from Pectobacterium carotovorum ssp. carotovorum (PccGCS) confirmed the importance of an ordered water within the heme pocket and two distal residues (Tyr43 and Ser68) as hydrogen-bond donors. Taken together, this work provides mechanistic insights into BpeGReg O₂ sensing and the signaling mechanisms of diguanylate cyclase-containing GCS proteins.

INTRODUCTION

Sensor proteins are used throughout biology to transmit changes in signal concentrations throughout an organism. However, the mechanisms by which sensor proteins transmit the ligand binding event into a downstream signal remain poorly understood. As the presence of molecular oxygen (O₂) is an important determinant of bacterial metabolism as well as reactive oxygen species-mediated host immune responses, it is essential that bacteria be able to respond to changing O₂ levels. The globin coupled sensor (GCS) family of bacterial heme sensor proteins monitors O₂ levels and modulates downstream pathways for adaption, including O₂-dependent biofilm formation, motility, and virulence. Because biofilms are a primary defense against antibiotic treatment, host immune responses, and environmental stress, understanding the mechanisms by which O₂ levels regulate GCS activity has the potential to identify new antibacterial targets.

GCS proteins are widely distributed throughout bacteria, as well as in a number of archaea, suggesting their widespread importance in controlling O₂-dependent bacterial phenotypes. These proteins consist of an N-terminal sensor globin domain linked through a variable middle domain to one of several output domains, such as methyl accepting chemotaxis proteins (MCP), diguanylate cyclases (DGC), phosphodiesterases, sulfate transport and anti-anti- factor domains, and kinases (Figure 1A). Within the GCS family, a number of putative proteins are predicted to contain DGC output domains, which are the enzymes responsible for synthesizing c-di-GMP, a bacterial second messenger that is a major regulator of biofilm formation. To date, ligand binding to the heme within the sensor globin domain of GCSs has been shown to alter activity of all GCS proteins with enzymatic output domains.

While the ligand-dependent enzymatic functions of these proteins have been investigated, the mechanism of activation for GCS proteins remains unknown. This is in part due to subtle differences between closely related GCS proteins. GCS proteins containing a DGC domain must dimerize to exhibit activity; however, GCS proteins from Bordetella pertussis (BpeGReg) and Pectobacterium carotovorum...
Fe(II)–O₂ state for extended periods. Furthermore, the isolated globin domains from BpeGReg and PccGCS (BpeGlobin and PccGlobin) have been shown to be even more stable than their full-length counterparts. While the GCS from oligomeric state alters the heme pocket environment and thus oxygen dissociation rates and their ability to remain in the mechanisms. (HemDGC) is only found as a tetramer. For psychrophila binding. Despite the variety of output domains and native exposure of the globin dimer interface correlates with ligand in which hydrogen/deuterium exchange coupled with mass on the GCS from states of these proteins. This has been corroborated by work dimer interface, and by extension changes in the oligomeric proteins is at least partially mediated by changes at the globin GCS proteins has been performed in either the Fe(II)− as the Fe(II)− O₂ at 4 °C for ∼18 h before showing significant autoxidation, BpeGlobin and PccGlobin remain fully Fe(II)−O₂ after 72 h at 20 °C. Consequently, BpeGlobin and PccGlobin are excellent systems to study the mechanism of O₂ sensing in these important signaling proteins.

To probe the mechanism of O₂-dependent activation and the molecular basis for subtle differences between closely related GCS proteins, X-ray crystallography and FTIR were performed on BpeGReg, PccGCS, and their protein variants. These studies provide novel insights into potential signaling pathways, the effects of ligand binding, the heme pocket environments, and key differences between the two related proteins in the GCS family.

**METHODS**

**Protein Expression.** Codon-optimized genes encoding BpeGReg and PccGCS in pET-20b served as templates for site-directed mutagenesis. Mutagenic primer sequences can be found in Supporting Information Table S1. Isolated globin and full-length distal pocket mutants were produced by standard PCR protocol as previously described. Proteins were expressed in E. coli Tuner (DE3) pLysS cells (Novagen) as previously described. Briefly, cells were transformed with codon-optimized pET20b plasmids encoding each protein via heat shock and grown overnight on LB agar plates containing chloramphenicol (30 μg mL⁻¹) and ampicillin (100 μg mL⁻¹). Transformants were then grown in Luria broth (LB) medium with the appropriate antibiotics at 37 °C with shaking at 225 rpm. These overnight cultures were used to inoculate yeast expression media (RPI) and grown to an OD₆₀₀ of 0.6. At this point, the temperature was dropped to 25 °C, and δ-aminolevulinic acid was added at a final concentration of 50 mM in each expression mixture and left to shake. After 30 min, protein expression was induced by the addition of 100 μM IPTG for 6 h. Cells were harvested by centrifugation (3500 × g, 4 °C, 20 min) and the resulting pellets were collected and stored at −80 °C.

**Protein Purification.** Cell pellets were resuspended in low imidazole buffer (50 mM Tris, 50 mM NaCl, 1 mM DTT, 20 mM imidazole, pH 7.0), lysed using a homogenizer (Avnet), and partially purified by centrifugation (186 000 × g, 4 °C, 1 h). The supernatants were loaded onto an equilibrated HisPur Ni²⁺-column (Fisher Scientific) and washed with low imidazole buffer. Proteins were eluted with high imidazole buffer (50 mM Tris, 50 mM NaCl, 1 mM DTT, 250 mM imidazole, pH 7.0) and further purified and desalted using a S200 gel filtration column (50 mM Tris, 50 mM NaCl, 1 mM DTT, 5% glycerol (v/v), pH 7.0). Proteins were concentrated by ultrafiltration (10 kDa MWCO filter, Millipore), aliquoted, flash frozen in liquid nitrogen, and stored at −80 °C.

**BpeGlobin Crystallization and Structural Determination.** Purified protein was reduced using sodium hydrosulfite salt under an anaerobic atmosphere (Coy Laboratories). BpeGlobin was desalted, oxygenated by mixing with aerobic buffer, and concentrated to 5 mg/mL. Crystals formed in 5-20% polyethylene glycol (PEG) 3350, with 0.2 M calcium acetate and 10 mM Tris, at pH 7.0. BpeGlobin was mixed with the crystallization solution in equal and increasing concentrations (i.e., 1 μL + 1 μL, 2 μL + 2 μL drops, etc.) in both hanging drop and sitting drop fashion. Crystals appeared at 20 °C after 7 days and were cryoprotected (100 mM Tris, pH 7.5, 400 mM CaOAc, 10% PEG 3350, 15% ethylene glycol) and flash frozen in liquid nitrogen. X-ray diffraction data were collected at the Southeastern Regional Collaborative Access Team (SER-CAT) ID22 beamline at the Advanced Photon Source (Chicago, IL). The X-ray diffraction data were indexed, integrated, and scaled in XDS to 2.3 Å. The structure

---

**Figure 1. Structure of BpeGlobin.** (A) Schematic of GCS domain architecture. (B) Orientation of the two different dimeric forms. Each monomer forms associations via salt bridges along helices A, B, and E in the asymmetric unit (solid pink, green, and blue). The ferrous-oxo dimer is shown in green, and the mixed dimer is shown in pink and blue. (C) Close-ups for both Fe(II)–O₂ (green insert, left) and Fe(III)–H₂O (pink insert, right) show heme pockets without significant differences (RMSD = 0.57 Å).

(PccGCS) exist in different mixtures of oligomeric states (i.e., monomer–dimer–tetramer and dimer–tetramer–octamer, respectively), while the GCS from Desulfotalea psychrophila (HemDGC) is only found as a tetramer. For both BpeGReg and PccGCS, the tetrameric assemblies with O₂ bound exhibit the highest activity. Furthermore, differences in dimerization affinity of the isolated globin domains were demonstrated to alter O₂ binding kinetics, suggesting that the oligomeric state alters the heme pocket environment and thus plays an important role in signaling. Activation of GCS proteins is at least partially mediated by changes at the globin dimer interface, and by extension changes in the oligomeric states of these proteins. This has been corroborated by work on the GCS from Anaeromyxobacter sp. Fw109-5 (AfgChHK), in which hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) was used to demonstrate that solvent exposure of the globin dimer interface correlates with ligand binding. Despite the variety of output domains and native oligomeric populations, the similar ligand-dependent responses of these GCSs suggest a common signaling pathway.

In previous work, crystallization of globin domains from GCS proteins has been performed in either the Fe(II)–CO or Fe(III)–CN ligation state to ensure extended heme stability, as the Fe(II)–O₂ state is prone to autoxidation. However, the Fe(II)–O₂ state is typically the high-activity form of GCS proteins, so this ligation state is the most informative for elucidating biologically relevant signaling mechanisms. BpeGReg and PccGCS have greater oxygen binding stability than other GCS proteins, such as EdDosC, HemAT-Bs, and AfgChHK, as demonstrated by their lower oxygen dissociation rates and their ability to remain in the Fe(II)–O₂ state for extended periods. Furthermore, the isolated globin domains from BpeGReg and PccGCS (BpeGlobin and PccGlobin) have been shown to be even more stable than their full-length counterparts. While the GCS from oligomeric state alters the heme pocket environment and thus oxygen dissociation rates and their ability to remain in the
was solved by molecular replacement using an ensemble of four globin structures (globins from two unique proteins; HemAT-Bs and EcDosC; PDB IDs 1OR4, 1OR6, 4ZVA, and 4ZVB) based on predicted structure from the sequence as a search model in PHENIX AutoMR.31 The model was built in Coot32 for residues 13–168 followed by refinement of x/y/z coordinates, occupancies, and B-factors in PHENIX to a final Rmerge/Rfree of 20.9/25.3% (Table 1). All figures of 2048 scans run with a 4 cm−1 resolution. The Varian 660 FTIR spectra were collected on a home-built, dry air purged external box connected to the FTIR spectrometer and equipped with a liquid nitrogen cooled MCT detector.33,34 Sample and reference spectra were collected as an average of 2048 scans run with a 2 cm−1 resolution. For all spectra, the absorbance spectra (\( \log(I_{\text{protein}}/I_{\text{water}}) \)) were baseline corrected with a spline function.

### RESULTS AND DISCUSSION

#### Structure of Ferrous-Oxy and Mixed BpeGlobin Dimers.

The globin domain of BpeGReg crystalized in the space group C121 and contained three monomeric units of BpeGlobin in the asymmetric unit (Figure 1, Table 1). However, the biological unit is a dimer as assessed by Pisa37 and based on previous biochemical studies.25,26 There are two species of biological dimers formed within the crystal lattice. The first dimer contains two monomers with O2 bound to the heme (hereafter referred to as the ferrous-oxy dimer) while the second dimer consists of one O2-bound species and one Fe(III)–H2O monomer (hereafter referred to as the mixed dimer) (Figure 1B and Figure S1). BpeGlobin adopts the classic globin fold of an eight \( \alpha \)-helix bundle with a heme held in place by proximal His99 located on helix F. Interestingly, an ordered water is found within the heme pocket of all monomers regardless of oxidation/ligation status. This water is located ∼4.1 Å away from the bound O2 and is engaged in a direct hydrogen bond with heme distal pocket residues, suggesting a potential role in ligand binding and stabilization (Table 2 and Table S2).

This is the first reported crystal structure of a bacterial sensor globin domain in the Fe(II)–O2 ligation state, offering a novel opportunity to examine a signaling domain in its native active form (Figure 1). Furthermore, both the ferrous-oxy and the mixed dimer are present, allowing for the comparison of the native active heme pocket structure with a low-activity form. Previous work has shown that Fe(II)–O2 BpeGReg is stable for at least 18 h with no significant oxidation, aggregation, or oligomerization changes.3 The exceptionally low autoxidation and O2 dissociation rates of BpeGReg likely enabled the extended stability of the Fe(II)–O2 protein necessary for its structural determination.

#### Comparison to Other Crystallized GCS Proteins.

To investigate the conservation of potentially relevant signaling residues, the structure of BpeGlobin was compared to

---

**Table 1. X-ray Structure Statistics**

| BpeGlobin | wavelength | resolution range | space group | unit cell | total reffns | unique reffns | multiplicity | completeness (%) | mean 1/\( \sigma(I) \) | Wilson B-factor | R-merge | R-meas | R-pim | CC1/2 | CC* | reflns used in refinement | reflns used for R-free | R-work (%) | R-free (%) | CC (work) | CC (free) | no. of non-hydrogen atoms | macromolecules | ligands | solvent | protein residues | RMS (bonds) | RMS (angles) | Ramachandran favored (%) | Ramachandran allowed (%) | Ramachandran outliers (%) | rotamer outliers (%) | clashscore | av B-factor | macromolecules | ligands | solvent |
|-----------|------------|-----------------|-------------|-----------|-------------|--------------|-------------|---------------|----------------|----------------|-------|--------|------|-------|------|--------------------------|----------------------|-----------|---------|-----------|----------|---------------------|-----------------|--------|---------|-------------|-----------|----------|-------------------|-------------------|----------------|----------|----------------|------------|
|          | 1.000      | 35.97–2.3 (2.38–2.30) | C121        | 145.16, 49.64, 68.14, 90, 110.59, 90 | 169919 (18748) | 18911 (2040) | 9.0 (9.2) | 92.14 (99.66) | 26.08 (19.21) | 38.06 | 0.07781 (0.1159) | 0.08293 (0.1227) | 0.02807 (0.03971) | 0.993 (0.994) | 0.998 (0.999) | 18882 (2033) | 1890 (205) | 16.14 (19.75) | 25.27 (29.43) | 0.969 (0.917) | 0.953 (0.817) | 3986 | 3664 | 136 | 186 | 469 | 0.008 | 1.12 | 98.27 | 1.73 | 0 | 1.32 | 46.12 | 45.94 | 46.52 | 49.25 | 39.25 | 38.5 | 37.0 | 14.5 | 12.0 | 22.5 | 21.0 | 20.0 | 19.5 | 20.0 |

*Statistics for the highest-resolution shell are shown in parentheses.*
Table 2. Inter-Residue Distances in Å, Including Nonligated Water Molecules

<table>
<thead>
<tr>
<th>residue</th>
<th>chain A Fe(II)</th>
<th>chain B Fe(II)</th>
<th>chain C Fe(II)</th>
<th>chain D Fe(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe–His99</td>
<td>2.1 ± 0.4</td>
<td>1.9 ± 0.5</td>
<td>2.0 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>Fe–Tyr43</td>
<td>4.5 ± 0.4</td>
<td>4.9 ± 0.5</td>
<td>4.7 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>Fe–Ser68</td>
<td>6.3 ± 0.4</td>
<td>6.2 ± 0.5</td>
<td>6.0 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>Fe–Trp72</td>
<td>8.1 ± 0.4</td>
<td>8.0 ± 0.5</td>
<td>8.3 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>Tyr43–Ser68</td>
<td>4.5 ± 0.4</td>
<td>4.1 ± 0.5</td>
<td>4.3 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>O2–Fe(II)</td>
<td>1.8 ± 0.4</td>
<td>1.9 ± 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2–Tyr43</td>
<td>2.3 ± 0.4</td>
<td>2.4 ± 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2–Ser68</td>
<td>5.1 ± 0.4</td>
<td>5.2 ± 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2O–Tyr43</td>
<td>2.9 ± 0.4</td>
<td>2.5 ± 0.4</td>
<td>2.5 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>H2O–Ser68</td>
<td>2.2 ± 0.4</td>
<td>2.1 ± 0.4</td>
<td>2.3 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>H2O–Fe</td>
<td>4.7 ± 0.4</td>
<td>3.5 ± 0.4</td>
<td>4.2 ± 0.4</td>
<td></td>
</tr>
</tbody>
</table>

Heme Distortion in BpeGlobin Influenced by O2 Binding. Heme deformation is a critical component in the signaling mechanism of several heme-based sensor protein families.40–42 To determine if heme distortion could be involved in signal transduction within GCS proteins, the heme cofactors of different monomers of the BpeGlobin dimers were compared to each other as well as to other GCS hemes (Figure 2 and Tables S3 and S4). The heme group can traditionally adopt different deformation forms including saddling, ruffling, and propellering, all of which can influence the surrounding heme pocket upon ligand binding to transduce the binding signal.40–42 All crystallized heme groups of BpeGlobin have similar saddling (Buu) and propelling (Auu) deformations (Figure 2A,B and Table 3). However, the direction and magnitude of the doming (Auu) of the Fe(II)–O2 hemes are more similar to each other than to the Fe(III)–H2O heme. Additionally, the change in ruffling value (Buu) of the O2-bound hemes from the ferrous-oxy dimer and the mixed dimer suggests that heme deformation for BpeGlobin is ligand-dependent and that heme distortion modes may play a role in GCS signaling. Since DGC activity is dimerization-dependent and the ligated state of the heme within GCS proteins affects oligomeric state, it is hypothesized that the ferrous-oxy dimer represents the active, O2-bound state, while the mixed dimer represents a transition between active and inactive states (Fe(III)–H2O is a low-activity state of the enzyme). Specifically, the difference in heme conformation between Fe(II)–O2 and Fe(III)–H2O hemes suggests that heme deformation propagates the ligand binding signal through interactions with heme pocket residues into the globin fold, thereby influencing the dimer interface, oligomeric state, and activity of the output domain.

Heme distortion within BpeGlobin is likely due to interactions with heme pocket residues; in the proximal pocket, most residues located within 5 Å of the heme are hydrophobic, with the majority being isoleucines and valines (Figure 3). These residues are well-positioned on the F helix to influence the globin dimer interface, changes which may represent a potential signaling mechanism.9,25,27 Such a mechanism is consistent with the findings of Stranava et al., who observed ligation state-dependent structural differences in the heme proximal residues and the dimerization interface of AfGch via HDX-MS.47 In addition to these hydrophobic proximal residues, Trp72 of BpeGlobin sits at the heme edge on the E helix within an appropriate distance for van der Waals contact with the heme. This Trp is conserved in EcDosC and AfGch, while HemAT-Bs contains a histidine residue in the same position (Figure S2).

The location of Trp72 on the E helix of BpeGlobin suggests another potential route for the changes in heme distortion to be propagated through the protein by steric interaction. As Ser68 and Lys64 are also located on the E helix (Figure 3), steric repulsion between the heme cofactor and Trp 72 could result in a shift in conformation of the E helix, which could then lead to global protein rearrangements. Ser68 is involved in stabilizing O2 binding30 while Lys64 forms a hydrogen bond with heme propionate 6 in only the Fe(II)–O2 monomers, suggesting a currently unexplored role for these residues in DGC activation. Recent cross-linking data also has demonstrated that the globin E helix of PccGCS makes a direct interaction with the DGC output domain near the active site (Walker et al., under review). Given the relatively high degree of identity between PccGCS and BpeGlobin (36%), it is inferred previously published structures of the globin domains from AfGch,27 HemAT-Bs,20,38 and EcDosC24 (Figure S2). Alignments of the globin domains of HemAT-Bs, AfGch, and EcDosC with BpeGlobin show a similar helical arrangement as well as a conserved hydrogen-bonding Tyr residue within the distal pocket. The crystal structure of BpeGlobin reveals shorter distances between the distal hydrogen-bonding residues (Tyr43 and Ser68) than EcDosC (Tyr43), HemAT-Bs (Tyr70 and Thr95), and AfGch (Tyr45 and Thr71; Table 2 and Table S2). Additionally, the distances between BpeGlobin Tyr43 and the ordered water, as well as Ser68 and the ordered water, are within range for hydrogen bonding, implying a role for the water in ligand binding and signal transduction (Table 2). The crystal structures of both HemAT-Bs and AfGch also contain water molecules, which are ~4.6 and ~5 Å from the heme ions, respectively (Table S2).20,27,38 In the case of HemAT-Bs, this ordered water has been shown to interact with bound O2 through resonance Raman studies.39 The water molecule for AfGch has not been further examined as of yet. However, no water was found within the distal pocket of EcDosC due to the presence of Leu92, which acts as a hydrophobic gate preventing the water from entering the heme pocket.59 Furthermore, EcDosC does not contain a second distal pocket hydrogen-bond donor, suggesting that the water may contribute to the different catalytic, ligand binding, and oligomerization behaviors of the various GCS proteins.24

DOI: 10.1021/acs.inorgchem.8b02584
that the E helix of BpeGReg likely also interacts with the DGC output domain. Taken together, these findings imply that ligand binding and the heme distortions that follow result in movements of Lys64, Ser68, and Trp72 that can then be propagated through the E helix to regulate the output domain.

**Heme Distortion Comparison to Other GCS Proteins.** It has been noted previously that similar trends in heme distortions may be observed for GCS proteins with the same class of output domain. 40,41,43 Accordingly, the distortions measured in BpeGlobin and EcDosC are more similar to each other than to HemAT-Bs (Table S3). Specifically, the Fe(III)—H₂O heme of BpeGlobin has highly similar 2ₐ₀ values to all units of EcDosC Fe(III)—H₂O (∼0.87 and ∼0.96, respectively). Additionally, heme groups from BpeGlobin and EcDosC both show high 2ₐ₀ and 2₁₀ (ruffling) values (∼0.90/0.85 and ∼0.92/0.70, respectively), whereas the values for HemAT-Bs are lower (∼0.58 and 0.46 for 2ₐ₀ and 2₁₀, respectively) indicating less overall distortion and ruffling. It should be noted that the Fe(III)—H₂O form of BpeGlobin was used in this comparison with other sensor globins in the Fe(III)—H₂O to control for the effects of different ligands and oxidation states on heme distortion. This analysis further supports previous findings that heme distortions are most similar between proteins that have the same function 40,44 and that conserved heme pocket residues assist in signal transmission within DGC-containing GCS proteins. 1,26,45 potentially by modulating heme ruffling (B₁₀) to transmit the ligation status of the heme-Fe into the globin fold. Indeed, a comparison of the pockets of PccGCS, EcDosC, and HemAT-Bs indicates that HemAT-Bs has the lowest sequence identity for these residues (12.5%, as compared to 75% for PccGCS and EcDosC) (Figure S2). In contrast, AfgcHK exhibits 16.5% sequence identity to BpeGReg and has a distinct output domain (histidine kinase). While some of the AfgcHK chains exhibit similar distortion trends to BpeGlobin and other Fe(III) globins, a universal trend within the AfgcHK chains is not apparent. Taken together, these data imply a general model for transduction of the ligand binding signal from the globin to DGC domain in GCS proteins via ligand-dependent distortion of the heme cofactors.

### Table 3. Heme Distortion Modes in BpeGlobin in Å 44

<table>
<thead>
<tr>
<th>Protein</th>
<th>Dₐ₀ (over all)</th>
<th>B₁₀ (saddle)</th>
<th>B₁₀ (ruffle)</th>
<th>A₂₀ (dome)</th>
<th>E₁(x) (wave)</th>
<th>E₂(y) (wave)</th>
<th>A₂₁ (propeller)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ferrous-oxy dimer Fe(II)—O₂</td>
<td>1.21</td>
<td>−0.22</td>
<td>1.13</td>
<td>−0.27</td>
<td>−0.24</td>
<td>−0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>mixed dimer Fe(II)—O₂</td>
<td>1.11</td>
<td>−0.13</td>
<td>−1.03</td>
<td>−0.32</td>
<td>0.09</td>
<td>−0.20</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>mixed dimer Fe(III)—H₂O</td>
<td>0.90</td>
<td>−0.18</td>
<td>0.85</td>
<td>0.15</td>
<td>−0.08</td>
<td>0.14</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

*Data shown is relative to a planar heme where all distortions are 0. Positive (+) numbers indicate distortion above plane; negative (−) indicate distortion below plane. Numerical values represent distortion magnitude.*

Figure 2. Heme deformations in BpeGlobin. (A) O₂-bound heme groups from both dimers exhibit a significant degree of ruffling, while the Fe(III)—H₂O heme exhibits doming in the opposite direction of the O₂-bound hemes. The combination of all six types of deformation yields the overall heme distortion shown. From left to right: Fe(II)—O₂ heme from the ferrous-oxy dimer, Fe(II)—O₂ heme from the mixed dimer, Fe(III)—H₂O heme from the mixed dimer. The differences in heme deformation lead to changes in the orientations of propionate 6, which forms a hydrogen bond with Lys74 in only the O₂-bound units. The 2Fᵣ − Fᵣ electron density is contoured at 1.5σ. (B) Heme cofactors can take on a number of deformation modes, including saddling (B₁₀), ruffling (B₁₀), propellering (A₂₁), doming (A₂₀), or waving (E₁(x), E₂(y)) that may influence ligand-dependent signaling (inspired by work from Shelnutt et al. 40).
Inorganic Chemistry

While previous studies on GCS proteins have not examined the role of heme distortion in signaling, work has been done to address the effect of high-spin/low-spin on activity of GCSs and other heme-based sensors.\textsuperscript{15,45−47} It has been noted that low-spin states of the heme are often correlated with the high-activity states of the enzymes; these findings agree with the data on \textit{Bpe}GReg and \textit{Pcc}GCS, which have identified the Fe(II)−O_{2} state as having the highest enzyme activity. \textit{Ec}DosP, a PAS domain heme sensor, also exhibits high activity in the Fe(II)−O_{2}−CO state, but low activity in the high-spin Fe(II) unligated state.\textsuperscript{45,46} Previous studies on \textit{Af}GcHK also support this trend and suggest that it may be due to the observation that high-spin Fe(II) heme in heme sensor proteins often has a shift of Fe upward out of the heme plane by 0.4−0.5 Å. While ligation pushes the Fe back down toward the plane.\textsuperscript{47−49} However, as \textit{Ec}DosC, the GCS from \textit{E. coli}, was found to exhibit high diguanylate cyclase activity in the high-spin Fe(II) and low-spin Fe(II)−O_{2} and Fe(II)−CO states, the link between spin state and activity is not universal.\textsuperscript{15} Taken together, the previous studies and our results suggest that heme distortion, which is affected by ligand binding, plays an important role in GCS signaling and activity.

**Interactions between Distal Residues and Bound Ligand.** In order to verify putative interactions between bound ligands and protein residues identified from the BpeGlobin crystal structure, FTIR was performed on CO-bound wild-type and BpeGReg/BpeGlobin protein variants, as well as the closely related GCS from \textit{P. carotovorum}, PccGCS/PccGlobin (Figure 4).\textsuperscript{4} PccGCS is often examined in tandem with BpeGReg due to the relatively high degree of sequence identity between the two proteins compared to other GCSs and the information available on enzymatic activity and ligand binding.\textsuperscript{5,25,26}

Furthermore, as \textit{Pcc}GCS has been demonstrated to control O_{2}-dependent virulence of a plant pathogen, additional insight into its mode of signal transduction is important for development of inhibitors.\textsuperscript{9} Although the sensor globins from HemAT-\textit{Bs} and EcDosC have previously been crystallized, they contain lower sequence identity to BpeGReg as compared to \textit{Pcc}GCS (14\%, 34\%, and 36\%, respectively) (Figure S2). \textit{Pcc}GCS exhibits the FTIR spectrum expected for a globin, with peaks at 1922 and 1961 cm\(^{-1}\) corresponding to CO stretches with a hydrogen-bonding interaction and apolar pocket, respectively (Figure 4B and Table S5).\textsuperscript{50,51} Surprisingly, BpeGReg exhibits three CO stretching frequencies, 1925, 1964, and 1972 cm\(^{-1}\) (Figure 4A and Table S5), which likely correspond to a hydrogen bond, apolar interaction, and electrostatic interaction with a negative charge, respectively.\textsuperscript{51} FTIR scanning with isotopically labeled \textsuperscript{13}CO resulted in a downfield shift of all three peaks in the spectrum of BpeGReg (1925, 1964, and 1972 cm\(^{-1}\)) and two peaks in the spectrum of \textit{Pcc}GCS (1922 and 1961 cm\(^{-1}\)) (Figure 4A,B and Table S5), verifying that these stretching frequencies correspond to the CO stretch. In addition, FTIR spectra of full-length BpeGReg and PccGCS were consistent with the corresponding globin spectra (Figure 4C,D and Figure S3), demonstrating that the deletion of the middle and DGC domains does not substantially affect interactions between bound CO and the proteins.

The observation of a 1972/1973 cm\(^{-1}\) stretching frequency for BpeGReg and BpeGlobin, respectively, was intriguing given the lack of charged amino acids within the distal pocket (Figure 1 and Figure S2). Furthermore, the observation of the 1972/1973 cm\(^{-1}\) stretch only in BpeGReg/Globin suggests subtle differences in domain structure and/or conformation, as compared to \textit{Pcc}GCS, supporting previously identified differences in ligand binding kinetics. To probe interactions of heme distal pocket residues with bound CO, protein variants were generated with mutations to the distal hydrogen-bonding tyrosine and serine residues (Tyr43Phe and Ser68Ala for BpeGReg/Globin; Tyr57Phe for PccGCS; Tyr57Phe and Ser82Ala for PccGlobin) (Table S6). For both proteins, only mutation of the distal Tyr produced an appreciable change in the FTIR spectrum; spectra of both BpeGReg/Globin (Tyr43Phe) and PccGlobin (Tyr57Phe) lack the stretches at 1925 and 1922/1921 cm\(^{-1}\), respectively, suggesting that the distal tyrosine provides a hydrogen bond to bound CO (Figure 4D, Figure S3, and Table S7). Furthermore, the apolar stretch near 1960 cm\(^{-1}\) shifted for both BpeGReg/Globin (Tyr43Phe) and PccGlobin (Tyr57Phe), indicating a change in the heme pocket environment upon mutation of Tyr43 (Figure 4D and Figure S3B). This shift may be due to subtle rearrangements that allow the partial positive charge in the center of the phenyl \(\pi\)-ring to interact with the C=O triple bond, increasing the polarity and subtly changing the stretching frequency. Taken together, these data demonstrate that the distal tyrosine residue in these proteins makes a direct interaction with the bound CO ligand, while the distal serine does not. However, this work, along with previous studies,\textsuperscript{26} suggests that the distal serine plays a role in \(O_{2}\) binding, likely through a hydrogen-bond network within the heme pocket.

**Examination of Structural Water.** The observation of a BpeGReg/Globin CO stretching frequency at 1972/1973 cm\(^{-1}\) suggested that the bound water may be involved in modulating

---

**Figure 3.** Heme pocket residues in \textit{Bpe}Globin. (A) The distal hydrogen-bonding residues Ser68 and Tyr43 are shown, along with the proximal His99 and the heme pocket residues Lys64 and Trp72. (B) Hydrophobic residues on the heme proximal side are shown. These residues are thought to propagate ligand-mediated heme distortions throughout the protein.
the distal pocket environment. While the crystal structure identified a hydrogen bond between O2 and the distal pocket water molecule, it was not clear which distal pocket residues the water interacts with to stabilize positioning within the pocket. To explore the role of the water observed in the heme pocket of the BpeGlobin crystal structure, FTIR was performed on wild-type BpeGlobin at pH values ranging from 5.0 to 8.0 (Figure 5). As the buffer pH was decreased, the peak at 1973 cm⁻¹ decreased in intensity, suggesting that the distal pocket ordered water exists in a partially deprotonated state, potentially stabilized through a hydrogen bond to the tyrosine.

To further support the distal water as the moiety responsible for the stretch at 1972 cm⁻¹, FTIR spectra were taken in D₂O and H₂¹⁸O labeled water (Figure 6 and Figure S4A). Within the BpeGlobin FTIR spectrum, the 1973 cm⁻¹ peak shifts and loses intensity when it is in the H₂¹⁸O buffer (Figure 6A). As this peak is tyrosine-dependent and is the only peak significantly affected by the change in pH, it is likely the...
result of the CO interaction with a partially deprotonated distal H₂O, spatially oriented through interactions with the distal Tyr43.

While PccGCS/PccGlobin does not exhibit a peak in the 1970 cm⁻¹ region like BpeGReg/BpeGlobin, the two proteins share high sequence identity (36% overall; 31% for the globin domain), suggesting they adopt a similar tertiary fold. Further, the high sequence identity between BpeGReg and PccGCS suggests that PccGCS/PccGlobin may also be regulated in a similar manner and thus likely also contains a distal pocket water. To interrogate the presence of a structured water in PccGlobin, and thus PccGCS, FTIR was performed on wild-type PccGlobin over a pH range of 7.0–9.0 (Figure 5B). Like BpeGlobin, PccGlobin remained stable at all tested pH values, although it was noted that the stability significantly decreased above pH 9. The spectrum of PccGlobin developed a peak at 1975 cm⁻¹ as pH increased, suggesting the deprotonation of a hydroxyl group in the heme pocket, potentially on either the distal tyrosine or the putative ordered water. FTIR spectra of PccGlobin were also taken in D₂O and H₂¹⁸O labeled water to support the distal water moiety (Figures 5B and 6B, and Figure S4B). Unlike in the H₂¹⁸O spectrum of BpeGlobin, that of PccGlobin does not demonstrate a significant shift in any of its peaks due to the lack of structural water-dependent stretches in the PccGlobin Fe(II)–CO spectrum (Figure 6). This difference in the protonation state of the ordered water between BpeGReg and PccGCS, which must be due to differences in heme pocket environments, likely contributes to the observed differences in O₂ binding kinetics (mono/biexponential dissociation rates) and O₂-dependent cyclase activity of these GCS proteins.⁴

### CONCLUSION

In conclusion, BpeGlobin is the first isolated GCS sensor globin domain to be crystallized in the native active Fe(II)–O₂ ligation state. Differences in ligation state correlate with heme distortion, suggesting a role in transmission of the O₂ binding signal. The differences in ligation state and heme distortion are potentially propagated through interactions with residues on the E helix, allowing for signaling to the DGC output domain. For BpeGReg, a 1973 cm⁻¹ peak is observed by FTIR and demonstrated to be due to the presence of an ordered water that is at least partially deprotonated in the heme pocket; a corresponding water in the PccGCS distal pocket remains protonated. In both proteins, the distal pocket Tyr is required for the stretching frequencies in both the 1970 and 1920 cm⁻¹ regions, indicating that the bound water is stabilized by the Tyr residue and interacts with the bound CO ligand. Taken together, these studies suggest roles for the distal water and heme distortion in transmission of the O₂ binding signal in GCS proteins (Figure 7).

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.inorgchem.8b02584.

Primers list, full table of heme distortion modes and of distance for all known crystallized globins, RMSD values for alignment of the three BpeGlobin chain, along with a sequence alignment figure to highlight key amino acids from all the compared globin domains, and additional FTIR spectra and all peaks (PDF)
Molecular Oxygen Dissociation and Diguanylate Cyclase Activity in Globin-Coupled Sensor Signaling

Fojtikova, V.; Martinkova, M.


ACKNOWLEDGMENTS
This work was supported by NSF CHE1352040 (E.E.W.) and Frasch Foundation Grant 824-H17 (E.E.W.). S.R. was partially supported by funds from the Emory University Laney Graduate School. We thank members of the Weinert and Frasch Foundation Grant 824-H17 (E.E.W.) for assistance and helpful suggestions. Additionally, this research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory, under Contract DE-AC02-06CH11357.

REFERENCES


(46) Tanaka, A.; Shimizu, T. Ligand binding to the Fe (III)-protoporphyrin IX complex of phosphodiesterase from Escherichia coli (Ec DOS) markedly enhances catalysis of cyclic diGMP: Roles of Met95, Arg97, and Phe113 of the putative heme distal side in catalytic regulation and ligand binding. Biochemistry 2008, 47 (50), 13438–13446.


