Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Shannon Rivera Date

Elucidating the Various Roles of the Globin Domain from Globin Coupled Sensors
By

Shannon Rivera
Doctor of Philosophy

Chemistry

_____
Dr. Emily Weinert
Advisor

_____
Dr. Brian Dyer
Committee Member

_____
Dr. Stefan Lutz
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Elucidating the Various Roles of the Globin Domain from Globin Coupled Sensors

By

Shannon Rivera
Chemistry B.S., University of Georgia, 2014

Advisor: Dr. Emily Weinert, Ph.D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Uni of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry.
2019

Abstract

Elucidating the Various Roles of the Globin Domain from Globin Coupled Sensors

By Shannon Rivera

Globin coupled sensors (GCS) are sensory proteins used by bacteria to determine the surrounding gaseous environment. The function of a GCS is determined by the output domain of the GCS, which include phosphodiesterases, kinases, and diguanylate cyclases (DGC). Diguanylate cyclase domains produce cyclic dimeric guanosine monophosphate (c-di-GMP) from guanosine triphosphate (GTP). C-di-GMP is a bacterial secondary messenger and a major regulator of biofilm formation. *Pectobacterium carotovorum ssp. carotovorum* and *Bordetella pertussis* both contain GCS proteins (*Pcc*GCS and *Bpe*GReg, respectively) with DGC output domains. Previous works has shown that oxygen binding in the globin domain regulates the output domain, but the signaling mechanism and structure of GCSs are not well characterized.

Here we present our work on elucidating these previously uncharacterized properties. In doing so, the isolated globin domains from *Pcc*GCS (*Pcc*Globin) and *Bpe*GReg (*Bpe*Globin) have been characterized. The oligomeric state of *Pcc*Globin is dimeric while *Bpe*Globin is monomeric, indicating potential oligomer binding sites in the globin domain. As full length PccGCS and *Bpe*GReg exist as different oligomeric states (dimer-tetramer-octamer and monomer-dimer-tetramer, respectively), the globin domain appears to be a major determinant of oligomerization. Additionally, the globin truncations revealed altered oxygen dissociation kinetics, as compared to *Pcc*GCS and *Bpe*GReg.

Furthermore, crystallization of *Bpe*Globin allowed for additional identification of key residues in the heme pocket. Site-directed mutagenesis has been used to interrogate the relative roles of each of these residues in stabilizing bound O_2 and contributing to each dissociation rate in both the full-length and isolated globin proteins. Further investigation into the heme pocket and the dimerization of the globin domain has been accomplished using domain swapping, cross-linking, Fourier Transform Infrared (FTIR) spectroscopy, electrochemistry and more. This work elucidates the global effects of protein oligomerization on conformation of the globin domain, as well as identifies key requirement for signal transduction within the globin coupled sensor family.

Elucidating the Various Roles of the Globin Domain from Globin Coupled Sensors

By

Shannon Rivera
Chemistry B.S., University of Georgia, 2014

Advisor: Dr. Emily Weinert, Ph. D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2019.

Contents

Chapter 4: Determining Heme Residue Effects on Kinetic Turnover in GCSs

Chapter 5: Conclusion - Signaling in Globin Coupled Sensors Proteins

List of Figures and Tables

Chapter 1:

Chapter 2:

Chapter 3: Structural Insights into Globin Coupled Sensors

Chapter 4:

Chapter 1: Introduction to Globin Coupled Sensor Signaling

Adapted from: Walker, J.A.; Rivera, S.; Weinert, E.E. (2017) Mechanism and Role of Globin Coupled Sensor Signaling. *Adv. Microb. Physiol. 71*, 133-169.
DOI: 10.1016/bs.ampbs.2017.05.003, with permission form Elsevier.

Chapter 1:
Introduction to Globin Coupled Sensor Signaling

Chapter Overview:

1. Introduction

Heme sensor proteins are ubiquitous from bacteria to mammals and are used to sense gaseous ligands, such as nitric oxide (NO), carbon monoxide (CO), and oxygen (O_2), to control diverse physiological responses.[6-8] The diverse families of heme sensor proteins include Heme Nitric oxide/OXygen (H-NOX) binding domains, heme-PAS domains, CooA proteins, and sensor globin domains (for a recent review of heme sensor proteins, please see). [6] Sensor globin domains are found as the heme domain within globin coupled sensors (GCS), a family of putative O_2 sensors consisting of an N-terminal sensor globin domain linked to a C-terminal output domain (Fig.1). [13] Putative GCS proteins are predicted to contain many types of output

domains, such as methyl accepting chemotaxis protein (MCP), kinase, diguanylate cyclase (DGC), adenylate cyclase, gene regulatory function, and domains of unknown function.

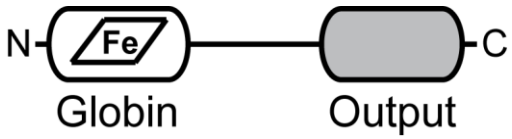

Figure 1. GCS proteins consist of a N-terminal globin domain containing a heme linked to a C-terminal output domain.

GCS proteins were first discovered as heme sensors that controlled aerotaxis in *Bacillus subtilis* and *Halobacterium salinarium* [13-15]. Further studies have identified roles for GCS proteins in controlling O_2-dependent second messenger production [4-5, 9, 11-12, 16-21], phosphorylation cascades [10, 22-24], and biofilm formation [5, 20, 25-26]. Given that putative GCS proteins are found within the genome sequences of numerous Gram-negative and Gram-positive bacteria, archaea, and even some protozoa [28], GCS proteins likely play important roles in controlling O_2-dependent physiology and phenotypes.

To date, characterized GCS proteins include MCP-containing proteins from *B. subtilis* [14], *Bacillus halodurans* C-125 [15], and *H. salinarium* [14]; DGC-containing proteins from *Escherichia coli* [16], *Shewanella putrefaciens* [20], *Bordetella pertussis* [5], *Pectobacterium carotovorum* [11], *Desulfotalea psychrophila* [21], and *Azotobacter vinelandii* [29]; kinase-containing GCS protein from *Anaeromyxobacter* sp. Fw109-5 [23]; adenylate cyclase-containing GCS from *Leishmania major* [19]; putative gene-regulating GCS from *Vibrio brasiliensis* [10], and GCS of unknown function from *Geobacter sulfureducens* [30]. Through a compilation of biochemical, spectroscopic, and physiological studies, the importance of these proteins in sensing the gaseous environment and regulating key organismal phenotypes is being elucidated.

2. General Structural and Ligand Binding Characteristics

The sensor globin domains within GCS proteins share sequence and structural homology with other globin domains, including mammalian myoglobin, and consist of an α-helical fold (Fig. 2 and 3) [1, 12, 30-31]. Compared to myoglobin, sensor globins lack the D helix found in myoglobin (Mb) and hemoglobin (Hb) and typically have a shortened E helix. The heme is bound within a pocket primarily consisting of helices C, E, F, and G, with the histidine that serves as the proximal ligand to the heme residing on the F helix. In sensor globins characterized to date, hydrogen bonding residues are found in the distal pocket and typically consist of a tyrosine located on helix B, and a threonine or serine residue also involved in hydrogen bonding found on helix E within some GCS proteins. All of the GCS proteins characterized so far have been demonstrated to form the expected Fe^{II}-NO, Fe^{II}-CO, and Fe^{II}-O_2 complexes with spectral characteristics (Fig. 4) that are very similar to myoglobin and other histidyl-ligated heme proteins [2, 5, 11, 14, 21, 23].

Crystal structures have been solved of the sensor globin domain from multiple GCS proteins, including the MCP-containing GCS from *Bacillus subtilis* (HemAT-*Bs*) [1] and the diguanylate cyclase-containing GCS from *Escherichia coli* (*Ec*DosC) [12]. Within these structures, the globin domains were found to crystallize as dimers, with a large buried dimer interface (1300-1800 $Å^2$). Helices G and H, as well as part of the Z helix in the case of HemAT-*Bs* globin, on each globin monomer form the interface with an anti-parallel four-helix bundle. The heme pocket opening is found away from the dimer interface, although residues from the G helix also line the back of the

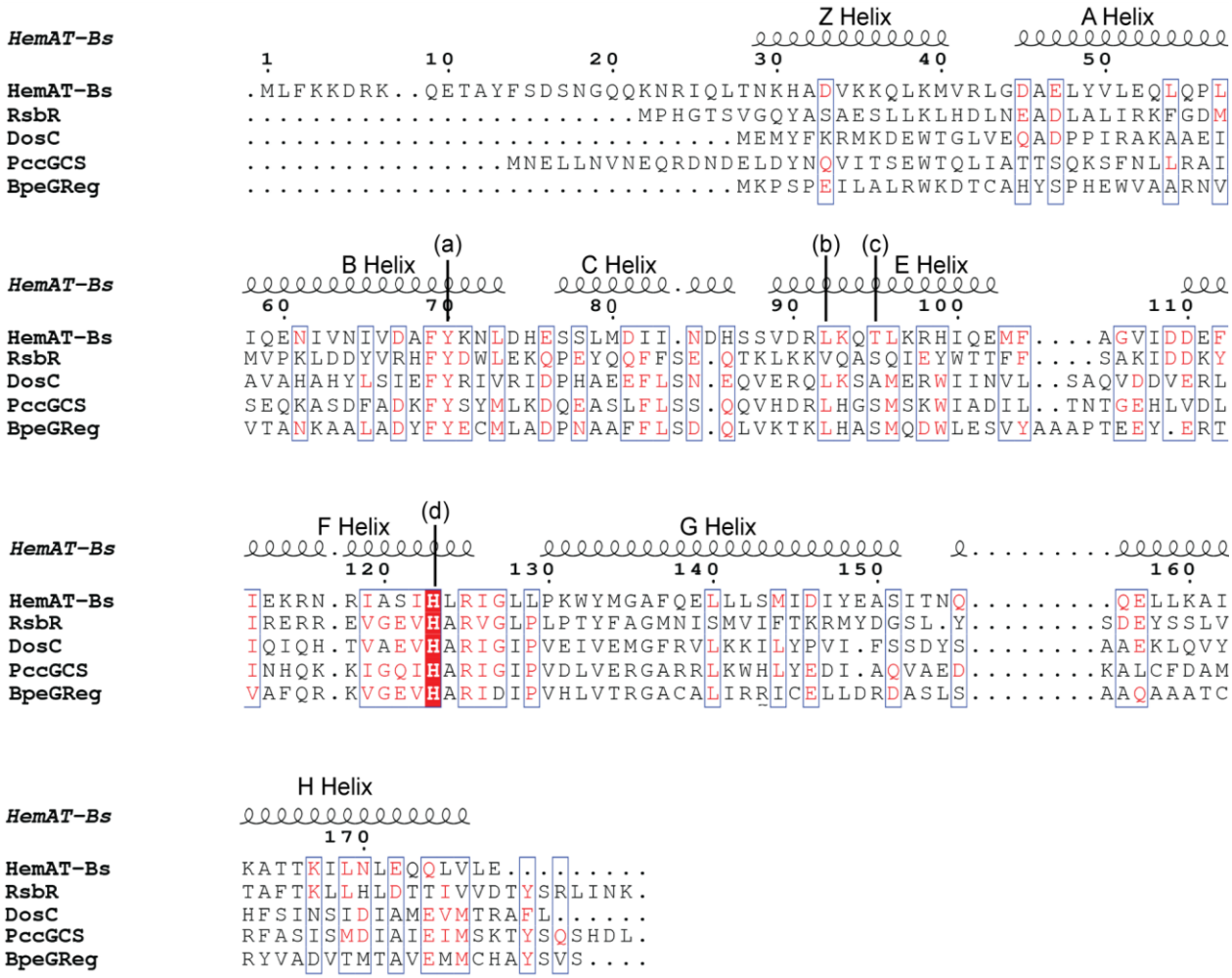

Figure 2. Globin domain sequence of GCS with key residues highlighted: (a) Distal pocket Tyr. (b) Distal pocket Leu. (c) Distal pocket Thr/Ser. (d) Proximal His. The helix labels can be seen above sequence and is based on HemAT-*Bs* (PDB: 1OR4).

heme pocket. In addition, the heme is buried deeper than in Mb and Hb, with decreased solvent exposed area and a large heme cavity. While structures of individual domains from GCS proteins and homologous domains have been crystallized, a full-length structure so far has remained elusive.

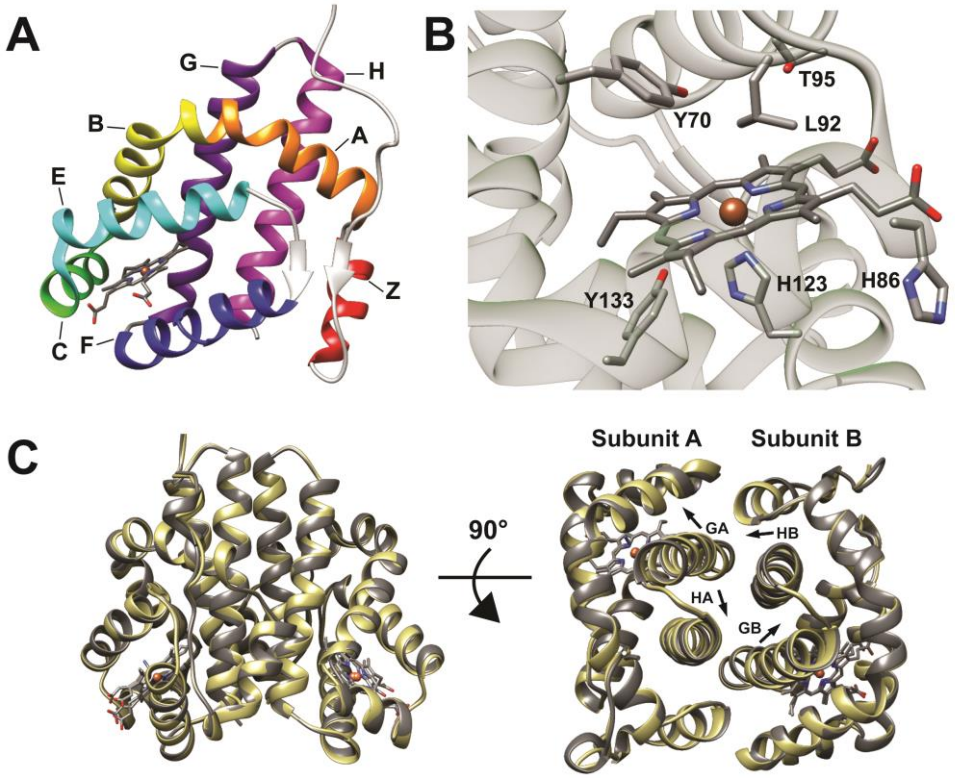

Figure 3. HemAT-*Bs* globin domain structure (PDB ID: 1OR4, 1OR6) [1]. (a) Labeling of globin domain helices. (b) Residues of the heme pocket (c) Globin dimer structure with overlay of two ligation states: cyano-liganded form (gray) - 1OR4, unliganded form (gold) - 1OR6. The labels GA and HA, and GB and HB represent G- and H-helices of subunit A and G- and H-helices of subunit B, respectively. The arrows indicate the movement of the helices when the unliganded dimer binds cyanide.

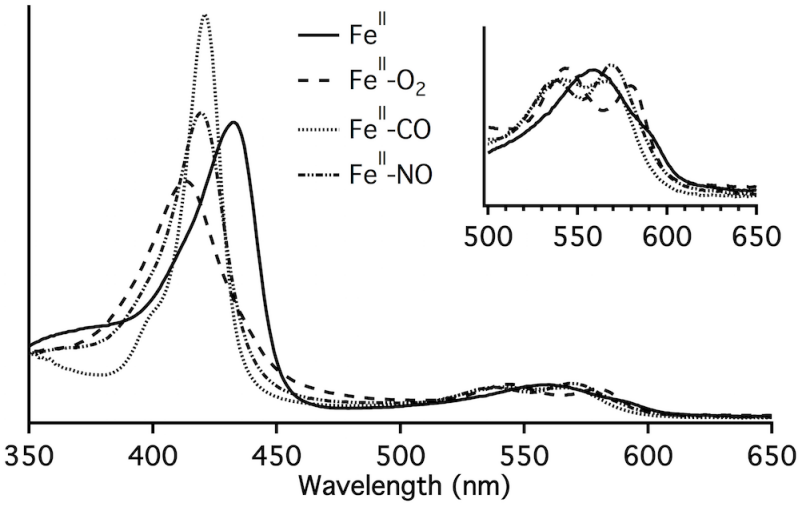

Figure 4. Representative UV-visible spectra of a GCS protein (*Pcc*GCS) [11]. Fe(II) exhibits a soret band at 430 nm with α/*β* around 550 nm; Fe(II)-O_2 exhibits a soret band at 416 nm with α/*β* at 540 and 580 nm; Fe(II)-CO presents a sharp soret band at 422 nm with α/*β* at 540 and 560 nm; Fe(II)-NO exhibits a soret band of 421 nm with α/*β* 540 and 570 nm.

3 HemAT-*Bs*: The First GCS Characterized

HemAT-*Bs* is a 432 amino acid protein and was one of the first two GCS proteins to be identified and characterized. It affects chemotaxis of *B. subtilis* to in response to O_2 levels, resulting in an aerophilic response. However, HemAT-*Bs* is not essential for aerotaxis, as all other putative MCP-like proteins in *B. subtilis* must be deleted to generate a robust HemAT-*Bs*-dependent aerophilic response [14]. In addition, HemAT-*Bs* does not appear to be methylated by methyltransferase CheR during the O_2 response, which is in contrast to many other bacteria in which methylation of an MCP results in the activation of histidine kinase CheA, part of the two component CheA/CheY signal transduction pathway in bacterial aerotaxis [14, 32]. Despite the lack of observable methylation, HemAT-*Bs* interacts with CheA, which presumably results in the observed O_2-dependent effects on taxis [33]. Furthermore, HemAT-*Bs* has been demonstrated to cluster with transmembrane chemoreceptors at the poles of *B. subtilis*, possibly forming complexes with these chemoreceptors [34], and is involved in pellicle formation of cultures at the air-liquid interface [35].

3.1 Ligand Binding Characteristics

Biochemical and spectroscopic characterization of purified HemAT-*Bs* and the isolated HemAT-*Bs* globin domain have provided insights into protein structure and ligand binding of GCS proteins [1-2, 15, 36-42]. Full-length HemAT-*Bs* forms a dimer in solution with an elongated, asymmetric shape, while the globin domain forms a monomer in solution but has been crystallized as a dimer [1-2].

The protein binds O_2 reversibly and exhibits biphasic binding kinetics, as measured by both kinetic measurements (stopped flow and laser flash photolysis) and equilibrium binding measurements (Table 1). [2, 13] Through resonance Raman and mutagenic studies, biphasic ligand binding kinetics have been demonstrated to originate from interactions with distal pocket residues (Y70 and T95; Fig. 3B) and different open/closed forms of the distal pocket (Fig. 5) [2, 37-38]. In the full-length, WT protein, three Fe-O stretching frequencies are observed, which, through mutagenesis and isotopic labeling, can be assigned to interactions of the bound O_2 with a water molecule hydrogen bonded to the distal pocket threonine (T95) within a closed distal pocket, interactions with T95 in a hydrogen bonding network with the distal tyrosine, and an open form of the pocket without direct hydrogen bonding interactions with O_2. From these studies, it has been proposed that T95 is required for O_2 sensing, while Y70 is crucial for signal transduction in HemAT-*Bs* [37]. In addition, the differences in hydrogen bonding and pocket conformation alter O_2 affinity, allowing HemAT-*Bs* to respond to O_2 over a concentration range of ~300 μM [2, 39].

The distal pocket residues of HemAT-*Bs* (Y70 and T95) adopt distinct conformations depending on the bound ligand, assisting with ligand recognition and discrimination, resulting in low CO affinity and monophasic binding [2, 38]. In addition to the distal hydrogen bonding residues, a heme pocket leucine (L92; Fig. 3B) serves as a conformational gate that both helps to maintain protein conformations following ligand binding and serves to direct ligand migration. Within the proximal pocket, a cavity near Y133 serves to bind and/or store gas molecules that also can be accessed through the L92 conformational gate once ligands have dissociated from the heme iron [38]. While considerable attention has been given to the ligand binding characteristics of HemAT

Table 1. Oxygen binding kinetics.

Protein	k_{on} ($uM^{-1}s^{-1}$)	$k_{off,1}$ (s^{-1})	$k_{off,2}$ (s^{-1})	K_d (uM)	Reference
HemAT-*Bs*	19	1900	87	0.010 0.22	2
*Ec*DosC (YddV)	1.4	21	N.R.	15	3
*Bpe*GReg	7[a]	0.82[b]	6.30b	0.117[c] 0.900[c]	a5 b11 c27
*Bpe*GReg Tetramer	N.R.	1.33	6.16	N.R.	11
*Bpe*GReg Dimer	N.R.	1.23	7.50	N.R.	11
*Pcc*GCS	7.2[a]	0.56[b]	3.87[b]	0.078[a] 0.538[a]	a27 b11
*Pcc*GCS Dimer	N.R.	0.668	4.68	N.R.	11
*Pcc*GCS Tetramer	N.R.	0.641	3.98	N.R.	11
*Vb*RsbR	5.2	7.0	67.4	1.4 13.0	10

N.R. – not reported

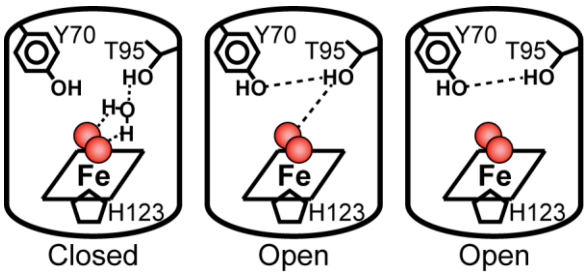

Figure 5. Schematic diagram of the conformations of O_2-bound HemAT-*Bs* based on resonance Raman data from Ohta et al., 2004. The left panel of the closed form shows the hydrogen bonding of T95 to water and water to O_2. The middle and right panels of the open forms show the hydrogen bonding of T95 with Y70 and O_2, and T95 with Y70 only, respectively.

proteins, *in vitro* enzyme kinetic assays to probe the effects of O_2 have proved elusive, as the MCP domain does not generate a measurable product and reconstituting methyl transfer pathways *in vitro* can be quite difficult [8, 32].

3.2 Signal Transduction in HemAT-Bs

Based on the aerotaxis data, binding of O_2 to HemAT-*Bs* should result in a conformational change that can transmit the O_2 binding event to the MCP domain and then to other chemotaxis-related proteins. While there is no full-length structural information on HemAT-*Bs*, or any other GCS protein, crystal structures of the globin domain in different ligation states and a variety of resonance Raman experiments on the full-length protein have been used to probe the global effects of ligand binding to generate models of signal transduction [1-2, 36-42].

HemAT-*Bs* globin has been crystallized as a dimer in both the Fe^{III}-CN and Fe^{II} states, providing a comparison of GCS globin structures with and without ligand bound to the heme (Fig. 3C). From these structures, a higher degree of symmetry within the dimer is observed in the ligand bound state, as compared with the unligated structure. In addition, the two subunits displayed differential structural changes when comparing Fe^{III}-CN and Fe^{II} structures, with Y70 rotating ~100 degrees away from the heme in the Fe^{II} structure, suggesting that heterogeneity or negative cooperativity in HemAT-*Bs* may be a cause of the biphasic O_2 binding kinetics and large sensing concentration range. Comparison of the ligand bound and unbound structures also demonstrates changes throughout the globin, with the G helix at the dimer interface being displaced relative to the H helix on each subunit, and the G and H helices undergoing small rotational movements, in addition to the globin monomers rotating relative to each other [1].

Further differences observed upon ligand binding in the full-length protein include changes around the heme propionate (Fig. 3B), which are propagated through hydrogen bonding between a heme-peripheral histidine (H86) and a heme propionate, resulting in conformational changes of the CE loop and E helix (which contains T95) [39]. Furthermore, mutation of H86 or T95 to alanine, preventing key hydrogen bonding interactions, disrupts the conformational changes of the B helix (which contains Y70) and the G helix, disrupting signal transduction from the globin domain to the middle domain [40]. Overall, ligand binding has been demonstrated to result in small rearrangements throughout the globin domain that alter symmetry and are propagated through the middle domain to control interactions of the MCP output domain.

4. Diguanylate Cyclase-Containing GCS Proteins

To date, GCS proteins with diguanylate cyclase (DGC) output domains have been characterized from *E. coli* (*Ec*DosC), *Shewanella putrefaciens* (DosD), *Bordetella pertussis* (*Bpe*GReg), *Pectobacterium carotovorum* (*Pcc*GCS), *Desulfotalia psychrophila* (HemDGC), and *Azotobacter vinelandii* (*Av*GReg) [3-5, 9, 11-12, 16, 20-21, 25-26, 43]. All of these DGC-containing GCS proteins have a linking middle domain between the N-terminal sensor globin and C-terminal DGC domain. However, the middle domains vary greatly in length, from ~40 amino acids in HemDGC to ~140 amino acids for *Ec*DosC, *Bpe*GReg, DosD, *Pcc*GCS, and *Av*GReg. In addition, the proteins exhibit modest sequence similarity (30-60%), supporting the subtle differences observed experimentally.

The C-terminal DGC domain in this subclass of GCS proteins synthesizes cyclic dimeric 3',5'-guanosine monophosphate (c-di-GMP) from two molecules of guanosine triphosphate (GTP). The DGC domain exerts requirements on the overall protein structure, as DGC domains are catalytically active as dimers, with each monomer binding one molecule of GTP and the cyclization occurring across the DGC dimer interface. Catalysis occurs within the GGDEF active site of the conserved domain, with the GTP monomers being arranged in an anti-parallel alignment. DGC domains also often contain a product-binding inhibitory site motif (RxxD) that binds c-di-GMP and inhibits enzyme activity. Therefore, within DGC-containing GCS proteins, cyclase activity potentially can be controlled by both ligand binding to the globin domain and c-di-GMP binding to the cyclase domain [44].

In bacteria, c-di-GMP has been demonstrated to serve as a second messenger that controls numerous cellular pathways and processes, including biofilm formation [9, 16, 44-45]. Given the importance of biofilm formation in bacterial growth and survival, including being of great import in infections by various bacterial pathogens, the roles of DGC-containing GCS proteins in controlling O_2-dependent c-di-GMP production have garnered considerable interest [20, 25-26]. In addition, as c-di-GMP production can be monitored *in vitro*, DGC-containing GCS proteins have provided an opportunity to dissect the effects of ligand binding on catalysis, linking these effects to structural rearrangements [6, 9, 11-12, 20, 27, 43].

4.1 *E. coli* DosC

The *E. coli* GCS, termed *Ec*DosC (for *E. coli* Direct oxygen sensor Cyclase), has an ~150 amino acid middle domain linking the N-terminal globin and C-terminal diguanylate cyclase domains [4,

[16]. The proximal histidine is conserved within the heme pocket; however only a tyrosine (Y43) is found as a distal hydrogen bond donor (a hydrogen bonding residue is not present at the homologous position of T95 in HemAT-*Bs*; Fig. 2) [4]. *Ec*DosC is able to bind all of the typical diatomic ligands and can stabilize O_2 binding (K_d = 15 μM), although the binding affinity is decreased by ~10-fold as compared to HemAT-*Bs*, potentially due to the altered hydrogen bonding interactions. Mutation of Y43 to alanine or leucine results in lack of measurable O_2 binding, while the Y43F and Y43W variants significantly weaken O_2 affinity, further highlighting the importance of Y43 in stabilizing O_2 bound to the heme. In addition, in contrast to HemAT-*Bs*, *Ec*DosC does not exhibit biphasic O_2 dissociation kinetics, suggesting differences in the heme pocket, likely due to the single hydrogen bond donor, and/or protein conformation [4].

While a second distal pocket hydrogen bonding residue is not present in *Ec*DosC, a distal pocket leucine (L65) plays an important role in stabilizing bound O_2. Mutation of L65 results in a spin state change for Fe^{III}, from 5-coordinate high spin for WT to 6-coordinate low spin for L65G/Q/T variants with H_2O as the axial ligand. L65 also serves to limit access to the distal pocket and ligand rebinding within the pocket, with mutations resulting in increased ligand association rates. In addition, blocking water access to the distal pocket decreases auto-oxidation of the heme, stabilizing the Fe^{II}-O_2 state of *Ec*DosC [3].

4.1.1 Ec*DosC Domain Characterization and Structures*

Similar to HemAT-*Bs*, full-length *Ec*DosC was found to be primarily dimeric in solution and to form a moderately elongated shape, with small percentages of larger oligomeric species. Characterization of individual domain constructs of *Ec*DosC, including crystal structures of each

domain, have provided the first insights into the possible three dimensional structure of a full-length GCS protein (Fig. 6 and 7) [12]. The *Ec*DosC globin domain forms a dimer both in solution and in the crystal structure, suggesting a role in dimerization of the full-length protein. Overall, the *Ec*DosC globin structure is quite similar to the HemAT-*Bs* globin, with a sensor globin fold and buried heme pocket; however, there are a few subtle differences between the globin domains. In contrast to the HemAT-*Bs* globin dimer, the *Ec*DosC globin dimer has a smaller buried surface area (1300 $Å^2$ vs. 1800 $Å^2$) and does not have any dimer contacts between the Z helices. Given that HemAT-*Bs* globin forms monomeric species in solution while *Ec*DosC globin forms dimers, subtle differences in sequence at the interface, rather than total buried surface area, must lead to the altered dimer affinity. In addition, the size of the heme pocket also is decreased for *Ec*DosC globin, as compared to HemAT-*Bs* globin (685 $Å^3$ and 995 $Å^3$, respectively) [12].

A truncation of the *Ec*DosC middle domain was crystallized to yield structures of middle domain dimers, which consisted of nearly entirely α-helical content (Fig. 6B) [12]. The dimeric middle domains formed an anti-parallel four-helix bundle, with the A and E helices lining the interface and forming a coiled-coil structure, and the N- and C-termini of the middle domain on the same of the assembly. While the middle domain structures were solved in two crystal forms ($P12_11$ and $P2_12_12_1$), only minor differences were observed, with a small rotation (~1.5°) and shift of the E-D-C-B helices of one monomer, as related to the other monomer, which did not provide significant insight into changes upon O_2 binding and activation. However, as the middle domains of DGC-containing GCS proteins are not homologous to previously characterized domains, these

structures provide information that has allowed for generation of full-length GCS protein models [12].

Overall, the structure of the *Ec*DosC DGC domain is quite similar to previously reported structures of unrelated DGC-containing proteins and consists of five strands of anti-parallel β-sheets, surrounded by five α-helices (Fig. 6C). The DGC domain forms a monomer both in solution and in the crystal structure, which is also analogous to DGC domains from PleD and WspR.

From the individual domain structures and solution data on the full-length protein, a model of the full-length *Ec*DosC protein has been proposed (Fig. 7). In this model, which was developed using docking protocols and manual positioning, each individual domain is depicted as a homo-dimer, and the globin domain is proposed to bind to one end of the middle domain, with the DGC domain bound to the other end of the middle domain. The resultant full-length *Ec*DosC model depicts a dimer that exhibits an elongated shape, as was observed for the full-length protein in solution. However, as 21 amino acids of the middle domain (residues 155-176) are not present in any of the structures and were modeled as unstructured linker, future studies will be required to refine the current understanding of *Ec*DosC structure. Nevertheless, the individual domain structures and full-length *Ec*DosC model provide the first experimental evidence regarding the architecture of DGC-containing GCS proteins and suggest that contacts within the middle domain may be required for signal transduction [12].

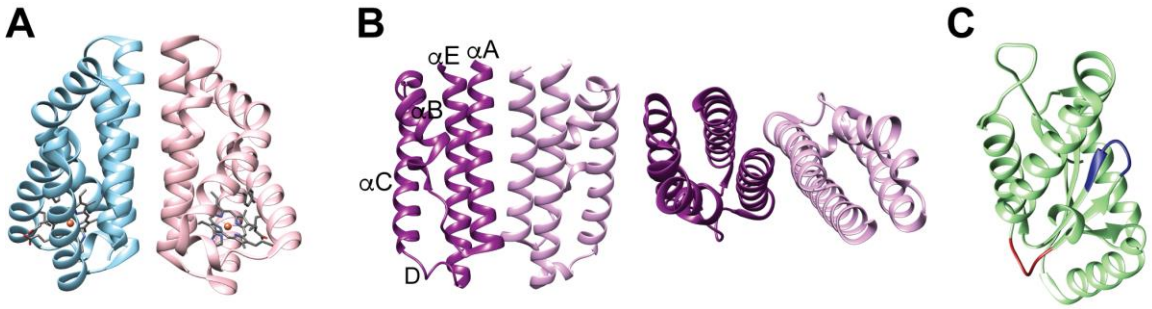

Figure 6. The isolated domains of EcDosC structures [12]. (a) Globin domain dimer, residues 8-154 (PDB: 4ZVA). (b) Front and top view of middle domain dimer, residues 177–292 (PDB: 4ZVC). (c) Monomeric DGC output domain with GGDEF domain highlighted in blue and RxxD domain highlighted in red, residues 297-460 (PBD: 4ZVG)

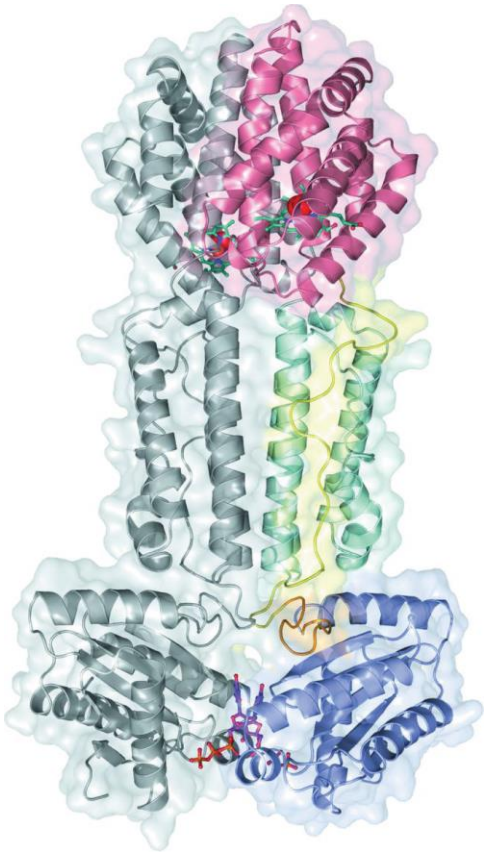

Figure 7. Full length, dimer model determined by docking experiments and reprinted with permission from Tarnawski, et al. 2015. Based on sedimentation data indicating an elongated dimer protein, EcDosC is shown in a linear model. Globin domains are shown with O_2 bound to the heme (O_2 was modeled as bound to the iron). Missing residues 155-176 are indicated by the yellow random coli. Only the two front α-helices of each middle domains (a 4 α- helix bundle) are shown. The DGC output domains are shown with GTP-γS bound.

4.1.2 Heme Ligand-Dependent Diguanylate Cyclase Kinetics

The production of a measurable product by *Ec*DosC, c-di-GMP, has allowed for in depth interrogation of the effects of various ligation and oxidation states on activity of the catalytic domain. Both the Fe^{II} unliganded and Fe^{II}-NO forms of *Ec*DosC resulted in nearly undetectable cyclase activity, while Fe^{III} yielded the greatest rate of catalysis, and Fe^{II}-O_2 and Fe^{II}-CO resulted in intermediate rates of c-di-GMP production (Table 2). These results raised the possibility that *Ec*DosC senses cellular O_2 levels through a change in redox state of the heme iron, with O_2 binding and auto-oxidation resulting in formation of Fe^{III} *Ec*DosC. However, the auto-oxidation rate of *Ec*DosC was found to be quite slow (0.0092 min^{-1}; sperm whale myoglobin = 0.001 min^{-1}), suggesting that O_2 binding to *Ec*DosC in *E. coli* likely does not cause oxidation of the heme on the time scales relevant for O_2-dependent c-di-GMP production and signaling and that *Ec*DosC reversibly senses diatomic O_2 [4].

Truncation of *Ec*DosC to yield a middle domain-DGC domain construct resulted in a protein that dimerizes in solution and exhibited ~4-fold lower diguanylate cyclase activity, as compared to full-length Fe^{III} *Ec*DosC. However, the isolated DGC domain construct (lacking both the globin

Table 2. Diguanylate cyclase GCS ligand-dependent activity.

Protein	k_{cat} Fe(II)	k_{cat} Fe(III)	k_{cat} Fe(II)-O_2	k_{cat} Fe(II)-CO	k_{cat} Fe(II)-NO	Ref.
*Ec*DosC (YddV)	<0.001[a]	0.066[a] 0.124[b]	0.022[a] 0.066[b]	0.022[a]	<0.001[a]	a4 b12
*Bpe*GReg	0.18	N.R.	0.59	0.23	0.38	[11]
*Pcc*GCS	0.29	N.R.	0.73	0.43	0.51	[11]

N.R. – not reported

and middle domains) forms monomeric species in solution and does not exhibit catalytic activity (DGC domains must dimerizes for cyclase activity), suggesting that globin/middle domains increase dimer affinity and are required to activate cyclase activity, rather than solely serving to repress activity of the cyclase domain [12].

4.1.3 Ec*DosC Signaling in* E. coli

*Ec*DosC (also annotated as *yddV*) is located within the DOS operon with *Ec*DosP, a heme-PAS domain protein with c-di-GMP phosphodiesterase activity. *Ec*DosC has been demonstrated to be predominantly expressed during the entry into stationary phase and is one of the most highly expressed DGC-containing proteins in *E. coli*. Furthermore, overexpression of *Ec*DosC results in increased intracellular c-di-GMP levels and increased biofilm mass. *Ec*DosC and *Ec*DosP have been shown work in concert to alter c-di-GMP levels in response to O_2 concentration, resulting in altered *E. coli* biofilm formation through regulation of curli-encoding genes and the operon that produces extracellular poly-*N*-acetylglucosamine. Both curli and poly-*N*-acetylglucosamine are key components of *E. coli* biofilms, demonstrating direct control of O_2-dependent biofilm formation by *Ec*DosC [4, 26].

In addition to roles in biofilm formation, *Ec*DosC and *Ec*DosP have been demonstrated to associate *in vivo*, forming hetero-oligomers. Furthermore, *Ec*DosC/*Ec*DosP hetero-oligomers interact with part of the *E. coli* RNA processing machinery, termed the degradosome, and co-purify as a ~1.3 MDa complex. Proteins and RNA were found within the complex, including RNase E, enolase, and PNPase. PNPase exhibits both 3'-polynucleotide polymerase activity and 3'-to-5'-exoribonuclease activity, depending on the other proteins within the degradosome

complex. When isolated in the *Ec*DosC/*Ec*DosP complex, PNPase was demonstrated to synthesize RNA tails under anaerobic conditions, but not under aerobic conditions. Furthermore, purified PNPase activity is directly regulated by c-di-GMP levels. Given the differential O_2-dependent activation of *Ec*DosC and *Ec*DosP, these data strongly suggest that *Ec*DosC and *Ec*DosP together regulate O_2-dependent levels of c-di-GMP within the immediate vicinity of PNPase to control PNPase activity and RNA processing. While the molecular details of the protein-protein interactions and particular mRNAs processed by the *Ec*DosC/*Ec*DosP complex are unknown, these studies suggest that O_2-dependent c-di-GMP production by GCS proteins likely controls downstream pathways, including mRNA degradation, in addition to biofilm formation [46].

4.2 *Bordetella pertussis Bpe*GReg

Characterization of the GCS from *B. pertussis*, the causative agent of whooping cough (termed *Bpe*GReg; *B. pertussis* Globin Regulator), has demonstrated that the GCS protein is involved in regulating biofilm formation, similar to the homologous proteins in *E. coli* and *S. putrefaciens*. *Bpe*GReg exhibits sequence similarity to *Ec*DosC, with the conserved proximal histidine, but contains two distal pocket hydrogen bonding residues (Y43 and S68) in homologous positions to the hydrogen bonding residues in HemAT-*Bs* (Fig. 2). Mutational analysis of the *Bpe*GReg globin domain has demonstrated that Y43 provides the primary hydrogen bond stabilizing bound O_2, while S68 provides secondary hydrogen bonding contacts. These results suggest a heme pocket conformation of the distal tyrosine similar to the *Ec*DosC pocket with a single hydrogen bond donor, but are in contrast to the arrangement of the HemAT-*Bs* pocket that contains two hydrogen bonding residues. As HemAT-*Bs* and *Bpe*GReg globins exhibit modest amino acid

sequence similarity (32%), differences in the globin domains and/or the middle and output domains likely are responsible for altering the globin domain conformations/structure and affecting ligand binding [5, 9, 11, 27].

Within the *Bpe*GReg middle domain, mutation of a histidine (H225) that shows sequence conservation within DGC-containing GCS proteins that have ~140 amino acid middle domains, such as *Ec*DosC and DosD, resulted in loss of diguanylate cyclase activity. From the *Ec*DosC middle domain crystal structures, H225 likely is not involved in packing of the four helix bundle, but could potentially be interacting with the DGC domain, if the full-length *Ec*DosC model is correct and *Bpe*GReg adopts a similar conformation [5, 9, 11, 27]. If so, these putative interactions could be important for correct orientation of the DGC domains. However, insights from the full-length *Ec*DosC model may not be directly applicable to *Bpe*GReg because, in contrast to *Ec*DosC, *Bpe*GReg forms a mixture of oligomeric states in solution, consisting of monomeric, dimeric, and tetrameric species. In addition, altering protein or salt concentration does not affect the equilibrium between the oligomeric states and isolated oligomers are slow to re-equilibrate, suggesting that maintaining the ratio of oligomeric states may be involved in protein function [11].

4.2.1 Effect of Ligand Binding on Oligomerization

Similar to HemAT-*Bs*, full-length *Bpe*GReg exhibits biphasic O_2 dissociation kinetics, yielding O_2 binding constants of 117 and 900 nM, and potentially allowing *B. pertussis* to respond to a wide range of environmental O_2 concentrations. Characterization of individual oligomeric states yielded differences in the O_2 dissociation rates (Table 1), although no significant differences in the amplitudes of each rate were observed [11]. Furthermore, the *Bpe*GReg globin domain

construct exclusively forms monomers in solution and exhibited monophasic O_2 dissociation kinetics, as well as an increased rate of O_2 association. Taken together, these data support a model in which different oligomerization states result in differences in structure and/or conformation of the globin heme pocket, thereby changing O_2 binding affinity [27].

4.2.2 Ligand-Dependent Cyclase Activation

*Bpe*GReg contains the conserved RxxD motif that binds c-di-GMP and results in product inhibition; however, *in vitro* enzyme kinetics could be measured by including a c-di-GMP phosphodiesterase in the reaction mixture to eliminate product inhibition. Analysis of enzyme assays demonstrated that *Bpe*GReg Fe^{II}-O_2 yields the greatest rate of c-di-GMP production, with Fe^{II} resulting in the lowest rate, and Fe^{II}-NO and Fe^{II}-CO yielding intermediate rates. In contrast to *Ec*DosC, *Bpe*GReg Fe^{II} results in measurable c-di-GMP production (k_{cat} = 0.18 min^{-1}) and O_2 binding only yields ~3-fold increase in c-di-GMP synthesis, demonstrating that the conformational changes associated with O_2 dissociation are not sufficient to fully inactivate the cyclase domain. Binding of O_2 also resulted in a decreased K_M for GTP, which suggests that the O_2 binding event is propagated throughout the protein, potentially altering both the orientation of the active sites and the GTP binding pocket [9, 11].

Binding of O_2 within the *Bpe*GReg globin domain results in changes to the ratio of oligomeric states of the protein (Fig.8A), suggesting a mechanism for O_2-dependent regulation of cyclase activity. The different oligomeric states of *Bpe*GReg exhibit differential rates of c-di-GMP production; tetrameric assemblies have ~5-fold greater cyclase activity than dimeric assemblies,

when standardized to monomer concentration. Furthermore, the inhibitory (RxxD) site also affects oligomerization state; c-di-GMP binding shifts the equilibrium away from tetrameric assemblies, while mutations (AxxD) that eliminate product binding allow *Bpe*GReg to access higher oligomeric states and alter O_2 dissociation kinetics. Therefore, the shifts in oligomeric state upon O_2 or c-di-GMP binding alter both O_2 affinity and c-di-GMP production, which may allow the organism to adapt its response based on environmental O_2 and intracellular c-di-GMP concentrations [9, 11].

4.3 *Pectobacterium carotovorum Pcc*GCS

P. carotovorum is a plant pathogen that displays O_2-dependent virulence and results in rotting of infected plant hosts. The GCS protein within *P. carotovorum*, termed *Pcc*GCS (*P. carotovorum* ssp. *carotovorum* Globin Coupled Sensor) displays 51% amino acid sequence similarity to *Bpe*GReg (35% amino acid identity), with intact DGC active site (GGDEF) and inhibitory site (RxxD) motifs, as well as heme pocket proximal histidine, distal tyrosine, and distal serine residues within the globin domain. Heme ligand-dependent activation of cyclase activity follows

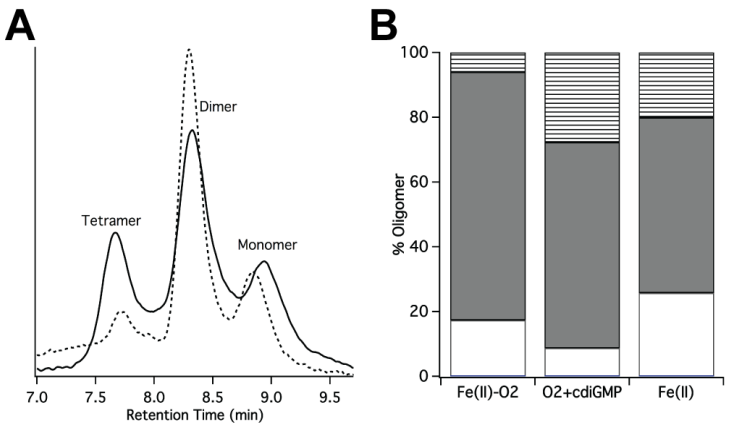

Figure 8. O_2 and c-di-GMP binding alter oligomerization of *Bpe*GReg (dashed line, Fe^{II}; solid line, Fe^{II}-O_2) (A) and *Pcc*GCS (white bar, dimer; grey bar, tetramer; striped bar, ≥ octamer) (B) [9].

the same trend as for *Bpe*GReg, with *Pcc*GCS Fe^{II}-O_2 exhibiting the greatest activity, and *Pcc*GCS also was demonstrated to exhibit biphasic O_2 dissociation rates [11].

However, despite these similarities, *Pcc*GCS exhibits key differences when compared to other DGC-containing GCS proteins. *Pcc*GCS exists as a mixture of dimer-tetramer-octamer/high molecular weight (HMW) species, in contrast to *Bpe*GReg (monomer-dimer-tetramer) and *Ec*DosC (primarily dimer). The globin domain likely affects the differences in the smallest accessible oligomeric state; *Bpe*GReg globin is monomeric in solution, while *Pcc*GCS and *Ec*DosC globin domains are dimeric in solution. The various GCS proteins also exhibit differences in O_2 binding kinetics, with *Ec*DosC exhibiting markedly slower O_2 association and faster O_2 dissociation rates, as compared to *Bpe*GReg and *Pcc*GCS. However, ~1.5-2-fold differences in O_2 dissociation rates were observed for *Bpe*GReg and *Pcc*GCS, suggesting that poorly understood differences in the protein structures/conformations result in pronounced differences in O_2 affinity and O_2-dependent cyclase activation [11, 27].

4.3.1 Pcc*GCS Ligand-Dependent Activity*

The different oligomeric states of *Pcc*GCS were demonstrated to exhibit different cyclase activities, similar to *Bpe*GReg, with tetrameric assemblies exhibiting greater activity than dimeric assemblies. As both O_2 binding to the globin domain and c-di-GMP binding to the DGC domain shifted the *Pcc*GCS oligomer equilibrium (Fig. 8B), as was observed for *Bpe*GReg, ligand-dependent changes in GCS oligomerization may be a mechanism used to control activity of a subset of DGC-containing proteins. Signal-dependent changes in oligomerization of unrelated DGC-containing proteins previously also has been observed, suggesting that

oligomerization state changes likely are a more general strategy to control c-di-GMP production. The lack of O_2-dependent changes in oligomerization observed for *Ec*DosC might be due to formation of hetero-oligomers with *Ec*DosP or involve changes in the hetero-oligomer composition, suggesting that DGC-containing GCS proteins may have evolved different activation mechanisms depending on their cellular partners.[9, 11].

5. Stressosome-Related GCS Proteins

Within *B. subtilis*, a protein, termed RsbR, has been characterized that contains a non-heme globin domain linked to a STAS (sulfate transport and anti-anti-σ factor) domain as part of the *rsb* operon involved in stressosome formation and signaling [47-49]. The non-heme globin maintains the sensor globin structure, but mutations within the globin heme pocket (including mutation of the proximal histidine) yielded a protein with a sensor globin fold that is unable to bind heme [49]. The stressosome was demonstrated to be an ~1.8 MDa complex in *B. subtilis* formed by RsbR, RsbS (STAS domain that makes protein-protein interactions with the RsbR STAS domain), and RsbT (kinase, phosphorylates RsbR/RsbS) that senses environmental stress and, through a phosphorylation cascade involving additional *rsb* proteins, eventually resulting in activation of σ^B and transcription of over 150 genes [50-54]. However, the ligand(s) that bind and activate the *B. subtilis* non-heme globin-containing RsbR are unknown. Therefore, bioinformatic searches that predicted the occurrence of GCS proteins with STAS output domains suggested that these proteins could provide insights into both stressosome signaling and function of additional GCS proteins [48-49, 52].

5.1 *Vibrio brasiliensis* GCS

RsbR protein sequences that contained the heme pocket proximal histidine and distal tyrosine were identified in the genomes of multiple *Vibrio* species [55], including *V. vulnificus*, which causes skin and soft tissue infections with a ~50% mortality rate [56]. Expression of the RsbR protein from *V. brasiliensis* (*Vb*RsbR), a non-pathogenic relative, found that the protein was heme bound and able to bind diatomic ligands [10]. When *Vb*RsbR was incubated with *Vb*RsbS and *Vb*RsbT, the other components of the stressosome complex, the proteins were able to associate, although complex formation did not depend on heme ligation state. Similar to HemAT-*Bs* and DGC-containing GCS proteins, *Vb*RsbR exhibited biphasic O_2 dissociation kinetics. In addition, O_2 dissociation from *Vb*RsbR was affected by formation of the *Vb*RsbR/S/T complex; however, in contrast to *Bpe*GReg and *Pcc*GCS, the amplitudes of each rate were altered upon complex formation (Table 1) but the dissociation rates were not affected. While the molecular details behind these differences are unknown, these data further support the roles of subtle differences between sensor globin domains, as well as the middle and output domains of GCS proteins, in affecting globin ligand binding kinetics and affinity [10].

Phosphorylation activity of *Vb*RsbT was demonstrated to be controlled by ligation/oxidation state of *Vb*RsbR; Fe^{II}-O_2 and Fe^{III} *Vb*RsbR inhibited *Vb*RsbT phosphoryl transfer to *Vb*RsbR and *Vb*RsbS, while inclusion of *Vb*RsbR Fe^{II}-CO and Fe^{II}-NO exhibited intermediate activity, and *Vb*RsbR Fe^{II} resulted in maximum phosphoryl transfer [10]. As phosphorylation of both RsbR and RsbS was absolutely required for activation of the *B. subtilis* Rsb signaling pathway [52-53], these studies suggest that *Vb*RsbR controls O_2-dependent stressosome signaling in *V. brasiliensis* (Fig.

9). These data also suggest that the conformational changes that occur in *Vb*RsbR upon O_2 association/dissociation are sufficient for signal propagation to multiple binding partners [10].

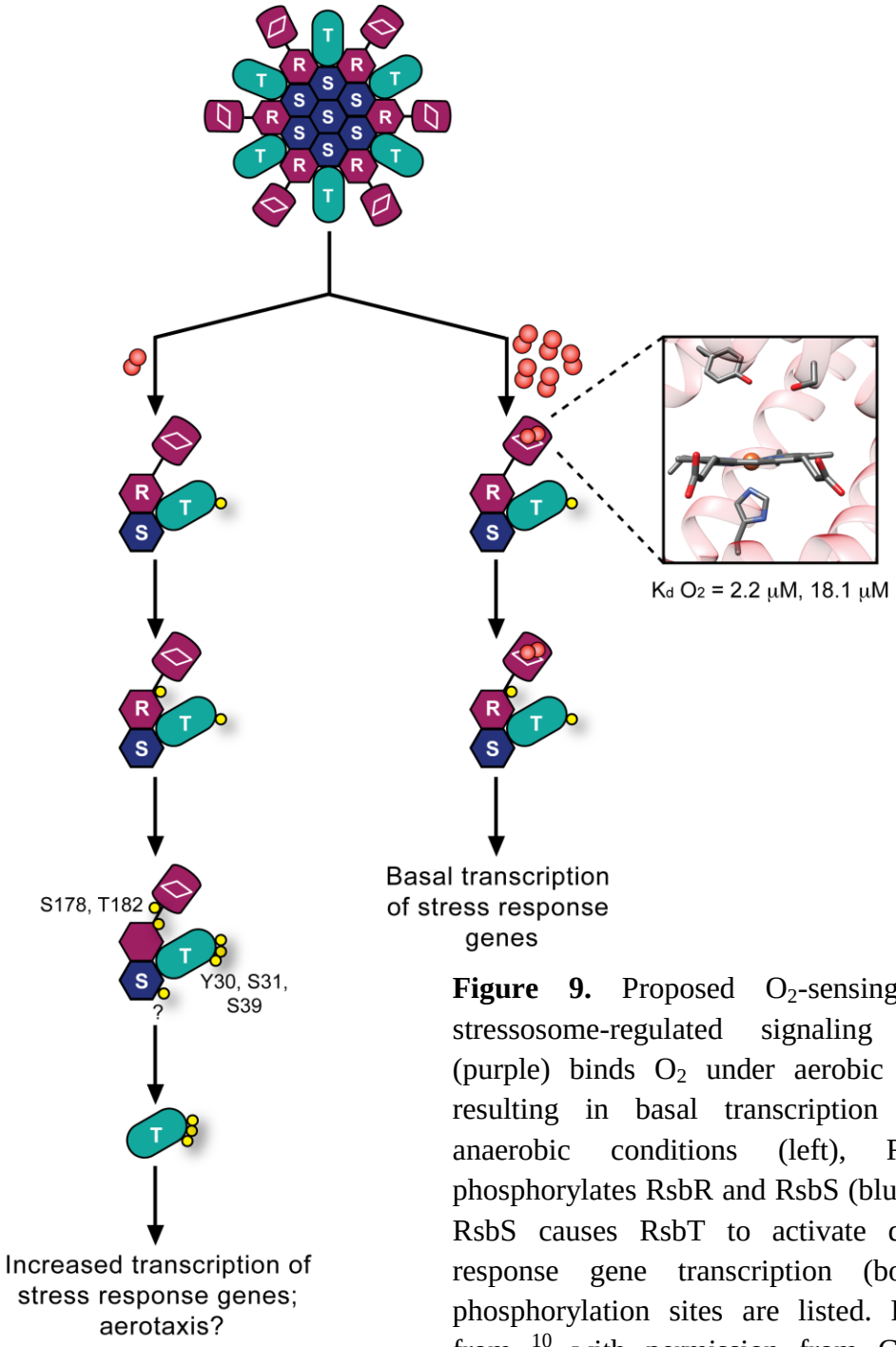

Figure 9. Proposed O_2-sensing *V. brasiliensis* stressosome-regulated signaling pathway. RsbR (purple) binds O_2 under aerobic conditions (right) resulting in basal transcription level (top). For anaerobic conditions (left), RsbT (turquoise) phosphorylates RsbR and RsbS (blue). Phosphorylated RsbS causes RsbT to activate downstream stress response gene transcription (bottom). Identified phosphorylation sites are listed. Figure reproduced from [10] with permission from Creative Commons

The sites of phosphorylation on *Vb*RsbR were found to occur on the short α-helical linker connecting the globin and STAS domains (*Vb*RsbR S178, T182) [10]. Previous mutational studies on the homologous linker from *B. subtilis* RsbR had demonstrated that linker is required for activation of downstream responses [57], suggesting that phosphorylation of the *Vb*RsbR linker may lead to the changes in protein conformation that activate *Vb*RsbT. Additional phosphorylation sites were identified in *Vb*RsbT on an α-helix (*Vb*RsbT Y30, S31, S39) that, in homologous proteins, is involved in kinase dimerization and interactions with downstream kinases, highlighting a possible mechanism to transmit the O_2 binding event to downstream proteins [10].

6. Conclusion

The diversity of output domains and physiological functions regulated by GCS proteins suggests that future studies will continue to find broad roles for these proteins in controlling O_2-dependent signaling pathways and phenotypes. The recent characterization of GCS proteins with catalytically active output domains has provided the basis for our understanding of how the O_2 binding signal is transmitted from the globin domain to control output domain function and has laid the groundwork for future studies into the rearrangements that occur upon ligand binding. Given that characterized GCS proteins have been demonstrated to play roles in biofilm formation, motility, and cell growth, the GCS protein family likely will be found to control additional important pathways and may serve as targets for novel tools to control O_2-dependent signaling.

7. Chapter References

1. Zhang, W.; Phillips, G. N., Jr., Structure of the oxygen sensor in *Bacillus subtilis*: signal transduction of chemotaxis by control of symmetry. *Structure* **2003,** *11* (9), 1097-110.

2. Zhang, W.; Olson, J. S.; Phillips, G. N., Jr., Biophysical and kinetic characterization of HemAT, an aerotaxis receptor from *Bacillus subtilis*. *Biophys. J.* **2005,** *88* (4), 2801-14.

3. Nakajima, K.; Kitanishi, K.; Kobayashi, K.; Kobayashi, N.; Igarashi, J.; Shimizu, T., Leu65 in the heme distal side is critical for the stability of the Fe(II)-O2 complex of YddV, a globin-coupled oxygen sensor diguanylate cyclase. *J. Inorg. Biochem.* **2012,** *108*, 163-70.

4. Kitanishi, K.; Kobayashi, K.; Kawamura, Y.; Ishigami, I.; Ogura, T.; Nakajima, K.; Igarashi, J.; Tanaka, A.; Shimizu, T., Important roles of Tyr43 at the putative heme distal side in the oxygen recognition and stability of the Fe(II)-O2 complex of YddV, a globin-coupled heme-based oxygen sensor diguanylate cyclase. *Biochemistry* **2010,** *49* (49), 10381-93.

5. Wan, X.; Tuckerman, J. R.; Saito, J. A.; Freitas, T. A.; Newhouse, J. S.; Denery, J. R.; Galperin, M. Y.; Gonzalez, G.; Gilles-Gonzalez, M. A.; Alam, M., Globins synthesize the second messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* **2009,** *388* (2), 262-70.

6. Shimizu, T.; Huang, D.; Yan, F.; Stranava, M.; Bartosova, M.; Fojtikova, V.; Martinkova, M., Gaseous O2, NO, and CO in signal transduction: structure and function relationships of heme-based gas sensors and heme-redox sensors. *Chem. Rev.* **2015,** *115* (13), 6491-533.

7. Gilles-Gonzales, M.-A., Gonzalez, G., Heme-based sensors: defining characteristics, recent developments, and regulatory hypothesese. *J. Inorg. Biochem.* **2005,** *99*, 1-22.

8. Martinkova, M.; Kitanishi, K.; Shimizu, T., Heme-Based Globin-Coupled Oxygen Sensors: Linking Oxygen Binding to Functional Regulation of Diguanylate Cyclase, Histidine Kinase and Methyl-accepting Chemotaxis. *J. Biol. Chem.* **2013,** *288*, 27702-27711.

9. Burns, J. L.; Rivera, S.; Deer, D. D.; Joynt, S. C.; Dvorak, D.; Weinert, E. E., Oxygen and c-di-GMP Binding Control Oligomerization State Equilibria of Diguanylate Cyclase-Containing Globin Coupled Sensors. *Biochemistry* **2016,** *55*, 6642-6651.

10. Jia, X.; Wang, J.-b.; Rivera, S.; Duong, D.; Weinert, E. E., An O2-Sensing Stressosome from a Gram Negative Bacterium. *Nature Commun.* **2016,** *7*, 12381.

11. Burns, J. L.; Deer, D. D.; Weinert, E. E., Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. *Mol. Biosyst.* **2014,** *10* (11), 2823-6.

12. Tarnawski, M.; Barends, T. R.; Schlichting, I., Structural analysis of an oxygen-regulated diguanylate cyclase. *Acta Crystallogr D Biol Crystallogr* **2015,** *71* (Pt 11), 2158-77.

13. Hou, S.; Freitas, T.; Larsen, R. W.; Piatibratov, M.; Sivozhelezov, V.; Yamamoto, A.; Meleshkevitch, E. A.; Zimmer, M.; Ordal, G. W.; Alam, M., Globin-coupled sensors: a class of heme-containing sensors in Archaea and Bacteria. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (16), 9353-8.

14. Hou, S.; Larsen, R. W.; Boudko, D.; Riley, C. W.; Karatan, E.; Zimmer, M.; Ordal, G. W.; Alam, M., Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* **2000,** *403* (6769), 540-4.

15. Hou, S.; Belisle, C.; Lam, S.; Piatibratov, M.; Sivozhelezov, V.; Takami, H.; Alam, M., A globin-coupled oxygen sensor from the facultatively alkaliphilic Bacillus halodurans C-125. *Extremophiles* **2001,** *5* (5), 351-4.

16. Tuckerman, J. R.; Gonzalez, G.; Sousa, E. H.; Wan, X.; Saito, J. A.; Alam, M.; Gilles-Gonzalez, M. A., An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* **2009,** *48* (41), 9764-74.

17. Roy, J.; Sen Santara, S.; Adhikari, A.; Mukherjee, A.; Adak, S., Control of catalysis in globin coupled adenylate cyclase by a globin-B domain. *Arch. Biochem. Biophys.* **2015,** *579*, 85-90.

18. Roy, J.; Sen Santara, S.; Bose, M.; Mukherjee, S.; Saha, R.; Adak, S., The ferrous-dioxy complex of Leishmania major globin coupled heme containing adenylate cyclase: the role of proximal histidine on its stability. *Biochim. Biophys. Acta* **2014,** *1844* (3), 615-22.

19. Sen Santara, S.; Roy, J.; Mukherjee, S.; Bose, M.; Saha, R.; Adak, S., Globin-coupled heme containing oxygen sensor soluble adenylate cyclase in Leishmania prevents cell death during hypoxia. *Proc Natl Acad Sci U S A* **2013,** *110* (42), 16790-5.

20. Wu, C.; Cheng, Y. Y.; Yin, H.; Song, X. N.; Li, W. W.; Zhou, X. X.; Zhao, L. P.; Tian, L. J.; Han, J. C.; Yu, H. Q., Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. *Sci. Rep.* **2013,** *3*, 1945-1951.

21. Sawai, H.; Yoshioka, S.; Uchida, T.; Hyodo, M.; Hayakawa, Y.; Ishimori, K.; Aono, S., Molecular oxygen regulates the enzymatic activity of a heme-containing diguanylate cyclase (HemDGC) for the synthesis of cyclic di-GMP. *Biochim. Biophys. Acta* **2010,** *1804* (1), 166-72.

22. Fojtikova, V.; Stranava, M.; Vos, M. H.; Liebl, U.; Hranicek, J.; Kitanishi, K.; Shimizu, T.; Martinkova, M., Kinetic Analysis of a Globin-Coupled Histidine Kinase, AfGcHK: Effects of the Heme Iron Complex, Response Regulator, and Metal Cations on Autophosphorylation Activity. *Biochemistry* **2015,** *54* (32), 5017-29.

23. Kitanishi, K.; Kobayashi, K.; Uchida, T.; Ishimori, K.; Igarashi, J.; Shimizu, T., Identification and functional and spectral characterization of a globin-coupled histidine kinase from Anaeromyxobacter sp. Fw109-5. *J. Biol. Chem.* **2011,** *286* (41), 35522-34.

24. Stranava, M.; Man, P.; Skalova, T.; Kolenko, P.; Blaha, J.; Fojtikova, V.; Martinek, V.; Dohnalek, J.; Lengalova, A.; Rosulek, M., Coordination and redox state-dependent structural changes of the heme-based oxygen sensor *Af*GcHK associated with intraprotein signal transduction. *J. Biol. Chem.* **2017**, jbc. M117. 817023.

25. Donné, J.; Maes, L.; Cos, P.; Dewilde, S., The role of the globin-coupled sensor YddV in a mature E. coli biofilm population. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2016,** *1864* (7), 835-839.

26. Tagliabue, L.; Maciag, A.; Antoniani, D.; Landini, P., The yddV-dos operon controls biofilm formation through the regulation of genes encoding curli fibers' subunits in aerobically growing Escherichia coli. *FEMS Immunol. Med. Microbiol.* **2010,** *59* (3), 477-84.

27. Rivera, S.; Burns, J. L.; Vansuch, G. E.; Chica, B.; Weinert, E. E., Globin domain interactions control heme pocket conformation and oligomerization of globin coupled sensors. *J. Inorg. Biochem.* **2016,** *164*, 70-76.

28. Vinogradov, S. N.; Tinajero-Trejo, M.; Poole, R. K.; Hoogewijs, D., Bacterial and archaeal globins - A revised perspective. *Biochim. Biophys. Acta* **2013,** *1834*, 1789-1800.

29. Thijs, L.; Vinck, E.; Bolli, A.; Trandafir, F.; Wan, X.; Hoogewijs, D.; Coletta, M.; Fago, A.; Weber, R. E.; Van Doorslaer, S.; Ascenzi, P.; Alam, M.; Moens, L.; Dewilde, S., Characterization of a globin-coupled oxygen sensor with a gene-regulating function. *J. Biol. Chem.* **2007,** *282* (52), 37325-40.

30. Pesce, A.; Thijs, L.; Nardini, M.; Desmet, F.; Sisinni, L.; Gourlay, L.; Bolli, A.; Coletta, M.; Van Doorslaer, S.; Wan, X.; Alam, M.; Ascenzi, P.; Moens, L.; Bolognesi, M.; Dewilde, S., HisE11 and HisF8 provide bis-histidyl heme hexa-coordination in the globin domain of Geobacter sulfurreducens globin-coupled sensor. *J. Mol. Biol.* **2009,** *386* (1), 246-60.

31. Zhang, W.; Phillips, G. N., Jr., Crystallization and X-ray diffraction analysis of the sensor domain of the HemAT aerotactic receptor. *Acta Crystallogr D Biol Crystallogr* **2003,** *59* (Pt 4), 749-51.

32. Collins, K. D.; Lacal, J.; Ottemann, K. M., Internal sense of direction: sensing and signaling from cytoplasmic chemoreceptors. *Microbiol. Mol. Biol. Rev.* **2014,** *78* (4), 672-684.

33. Szurmant, H.; Ordal, G. W., Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol. Mol. Biol. Rev.* **2004,** *68* (2), 301-319.

34. Cannistraro, V. J.; Glekas, G. D.; Rao, C. V.; Ordal, G. W., Cellular stoichiometry of the chemotaxis proteins in Bacillus subtilis. *J. Bacteriol.* **2011,** *193* (13), 3220-3227.

35. Hölscher, T.; Bartels, B.; Lin, Y.-C.; Gallegos-Monterrosa, R.; Price-Whelan, A.; Kolter, R.; Dietrich, L. E.; Kovács, Á. T., Motility, chemotaxis and aerotaxis contribute to competitiveness during bacterial pellicle biofilm development. *J. Mol. Biol.* **2015,** *427* (23), 3695-3708.

36. Aono, S.; Kato, T.; Matsuki, M.; Nakajima, H.; Ohta, T.; Uchida, T.; Kitagawa, T., Resonance Raman and ligand binding studies of the oxygen-sensing signal transducer protein HemAT from *Bacillus subtilis*. *J. Biol. Chem.* **2002,** *277* (16), 13528-38.

37. Ohta, T.; Yoshimura, H.; Yoshioka, S.; Aono, S.; Kitagawa, T., Oxygen-sensing mechanism of HemAT from *Bacillus subtilis:* a resonance Raman spectroscopic study. *J. Am. Chem. Soc.* **2004,** *126* (46), 15000-1.

38. Pinakoulaki, E.; Yoshimura, H.; Daskalakis, V.; Yoshioka, S.; Aono, S.; Varotsis, C., Two ligand-binding sites in the O2-sensing signal transducer HemAT: Implications for ligand recognition/discrimination and signaling. *Proceedings of the National Academy of Sciences* **2006,** *103* (40), 14796-14801.

39. Yoshimura, H.; Yoshioka, S.; Mizutani, Y.; Aono, S., The formation of hydrogen bond in the proximal heme pocket of HemAT-Bs upon ligand binding. *Biochem. Biophys. Res. Commun.* **2007,** *357* (4), 1053-1057.

40. El-Mashtoly, S. F.; Gu, Y.; Yoshimura, H.; Yoshioka, S.; Aono, S.; Kitagawa, T., Protein conformation changes of HemAT-Bs upon ligand binding probed by ultraviolet resonance Raman spectroscopy. *J. Biol. Chem.* **2008,** *283* (11), 6942-6949.

41. Yoshida, Y.; Ishikawa, H.; Aono, S.; Mizutani, Y., Structural dynamics of proximal heme pocket in HemAT-Bs associated with oxygen dissociation. *Biochim. Biophys. Acta* **2012,** *1824* (7), 866-72.

42. El-Mashtoly, S. F.; Kubo, M.; Gu, Y.; Sawai, H.; Nakashima, S.; Ogura, T.; Aono, S.; Kitagawa, T., Site-specific protein dynamics in communication pathway from sensor to signaling domain of oxygen sensor protein, HemAT-Bs: Time-resolved Ultraviolet Resonance Raman Study. *J. Biol. Chem.* **2012,** *287* (24), 19973-84.

43. Lambry, J.-C.; Stranava, M.; Lobato, L.; Martinkova, M.; Shimizu, T.; Liebl, U.; Vos, M. H., Ultrafast spectroscopy evidence for picosecond ligand exchange at the binding site of a heme protein: heme-based sensor YddV. *The journal of physical chemistry letters* **2015,** *7* (1), 69-74.

44. Schirmer, T., C-di-GMP synthesis: structural aspects of evolution, catalysis and regulation. *J. Mol. Biol.* **2016,** *428* (19), 3683-3701.

45. Hengge, R., Principles of c-di-GMP signalling in bacteria. *Nature Rev. Microbiol.* **2009,** *7* (4), 263-73.

46. Tuckerman, J. R.; Gonzalez, G.; Gilles-Gonzalez, M. A., Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. *J. Mol. Biol.* **2011,** *407* (5), 633-9.

47. Marles-Wright, J.; Lewis, R. J., The stressosome: molecular architecture of a signalling hub. *Biochem. Soc. Trans.* **2010,** *38* (4), 928-33.

48. Chen, C. C.; Yudkin, M. D.; Delumeau, O., Phosphorylation and RsbX-dependent dephosphorylation of RsbR in the RsbR-RsbS complex of *Bacillus subtilis*. *J Bacteriol* **2004,** *186* (20), 6830-6.

49. Murray, J. W.; Delumeau, O.; Lewis, R. J., Structure of a nonheme globin in environmental stress signaling. *Proc Natl Acad Sci U S A* **2005,** *102* (48), 17320-5.

50. Marles-Wright, J.; Grant, T.; Delumeau, O.; van Duinen, G.; Firbank, S. J.; Lewis, P. J.; Murray, J. W.; Newman, J. A.; Quin, M. B.; Race, P. R.; Rohou, A.; Tichelaar, W.; van Heel, M.; Lewis, R. J., Molecular architecture of the "stressosome," a signal integration and transduction hub. *Science* **2008,** *322* (5898), 92-6.

51. Price, C. W.; Fawcett, P.; Cérémonie, H.; Su, N.; Murphy, C. K.; Youngman, P., Genome-wide analysis of the general stress response in Bacillus subtilis. *Mol. Microbiol.* **2001,** *41* (4), 757-774.

52. Gaidenko, T. A.; Price, C. W., Genetic evidence for a phosphorylation-independent signal transduction mechanism within the Bacillus subtilis stressosome. *PloS one* **2014,** *9* (3), e90741.

53. Kim, T.-J.; Gaidenko, T. A.; Price, C. W., In vivo phosphorylation of partner switching regulators correlates with stress transmission in the environmental signaling pathway of Bacillus subtilis. *J. Bacteriol.* **2004,** *186* (18), 6124-6132.

54. Eymann, C.; Schulz, S.; Gronau, K.; Becher, D.; Hecker, M.; Price, C. W., In vivo phosphorylation patterns of key stressosome proteins define a second feedback loop that limits activation of *Bacillus subtilis* sigmaB. *Mol. Microbiol.* **2011,** *80* (3), 798-810.

55. Pané-Farré, J.; Lewis, R. J.; Stülke, J., The RsbRST stress module in bacteria: a signalling system that may interact with different output modules. *J. Mol. Microbiol. Biotechnol.* **2005,** *9* (2), 65-76.

56. Williams, T. C.; Blackman, E. R.; Morrison, S. S.; Gibas, C. J.; Oliver, J. D., Transcriptome sequencing reveals the virulence and environmental genetic programs of Vibrio vulnificus exposed to host and estuarine conditions. *PLoS One* **2014,** *9* (12), e114376.

57. Gaidenko, T. A.; Bie, X.; Baldwin, E. P.; Price, C. W., Two surfaces of a conserved interdomain linker differentially affect output from the RST sensing module of the *Bacillus subtilis* stressosome. *J Bacteriol* **2012,** *194* (15), 3913-21.

Chapter 2: Introduction to Globin Coupled Sensor Signaling

Adapted from: Rivera, S.; Burns, J.L.; Vansuch, G.E.; Chica, B.; Weinert, E.E. (2016) Globin Domain Interactions Control Heme Pocket Conformation and Oligomerization of Globin Coupled sensors. *J. Inorg. Biochem. 164*, 70-76. DOI: 10.1016/j.jinorgbio.2016.08.016, with permission form Elsevier.

Chapter 2:
Characterizing GCS Globin Domain Interactions

Chapter Overview:

1. Introduction

Bacteria have evolved sophisticated systems to sense changes in their environment and to respond, optimizing their survival [1-2]. The mechanism by which bacteria sense changing conditions and transmit those signals into metabolic and phenotypic responses is an active area of research. The ability of bacteria to adjust to changing O_2 concentrations [3-8] suggests signaling pathway(s) that can sense extracellular O_2 and adjust intracellular chemistry. A class of heme proteins termed globin coupled sensors is predicted to serve as bacterial environmental O_2 sensors within a large number of bacteria [5, 9-11]. Globin coupled sensors (GCS) consist of a sensor globin domain linked to an output domain by a variable middle domain. Output domains vary widely and include methyl accepting chemotaxis proteins (MCP) [12-14], histidine kinases [15], anti-anti-sigma factors [17], and diguanylate cyclases [5, 18-21], with methyl accepting chemotaxis proteins and diguanylate cyclases being the most prevalent.

Within the cell, diguanylate cyclases are found fused to many types of regulatory domains and are responsible for cyclizing two GTP to form c-di-GMP, a bacterial secondary messenger that

controls biofilm formation [22-23]. Diguanylate cyclases have been demonstrated to be catalytically active as dimers, with the cyclization reaction occurring across the dimer interface[24-28], regardless of their regulatory domain. Work on the diguanylate cyclase-containing GCS proteins from *E. coli* (*Ec*DosC) [19, 21, 29], *Desulfotalea psychrophila* (HemDGC) [20], *Bordetella pertussis* (*Bpe*GReg, the causative agent of whooping cough) [18, 30], and *Pectobacerium carotovorum* ssp. *carotovorum* (*Pcc*GCS, a plant pathogen) [18] has demonstrated that cyclase activity is increased upon binding of O_2 to the globin domain, as compared with the Fe^{II} unligated state. Further characterization of GCS proteins from *B. pertussis* and *P. carotovorum* found that full length GCS proteins form mixtures of oligomeric states (monomer-dimer-tetramer and dimer-tetramer-octamer, respectively) [18]. For both *Pcc*GCS and *Bpe*GReg, the tetrameric assemblies were found to exhibit the highest cyclase activity (on a per monomer basis), individual oligomeric assemblies were slow to re-equilibrate (>18 hrs.), and oligomer percentages were not dependent on protein or salt concentration, suggesting that *Pcc*GCS and *Bpe*GReg exist as mixtures of kinetically trapped oligomers [18]. Furthermore, a shift toward tetrameric assemblies, which exhibit increased catalytic activity, was triggered by O_2 binding to the heme [18], highlighting the role of the sensor globin domain in controlling protein conformation/oligomerization. Characterization of *Ec*DosC [29] and HemDGC [20] also showed that the proteins formed primarily dimeric and tetrameric assemblies, respectively, highlighting that GCS oligomerization may be conserved.

Structures of isolated sensor globin domains from *Bacillus subtilis* (HemAT-*Bs*, MCP output domain)[31] and *Ec*DosC [29] GCS proteins have been solved and shown to form dimers in the crystals (homology model of *Pcc*Globin based on HemAT-*Bs* is shown in Fig. 1). The globin domains maintain a primarily α-helical globin fold with dimer interfaces that consist of two

helices from each monomer, arranged in coiled-coil packing with key hydrophobic and electrostatic interactions helping to stabilize the dimer [29, 31]. Within the heme pocket, a conserved proximal histidine ligates to the heme, and a distal tyrosine helps to stabilize bound ligands (Fig. 1A). The HemAT-*Bs* globin distal pocket has a second distal pocket hydrogen bond donor, a threonine, that also interacts with bound ligands [31-33], while the *Ec*DosC globin only has a distal pocket tyrosine [19, 29]. In contrast, the prototypical histidyl-ligated heme protein, myoglobin, contains a distal pocket histidine that coordinates bound ligands. Changes in oxidation/ligation states resulted in subtle rearrangements within the sensor globin structures, such as slight rotation

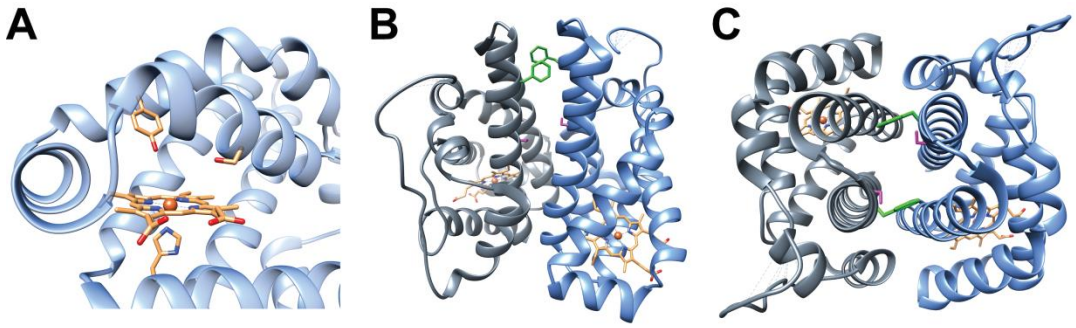

Figure 1. Homology model of *Pcc*Globin (based on PDB ID: 1OR4 [16]). A) Heme pocket with proximal histidine and distal tyrosine and serine highlighted. B) Dimer interface with F148R shown in green. C) Alternate view of the putative dimer interface.

at the dimer interface, which are hypothesized to lead to the observed ligand-dependent changes in catalytic activity [31]. However, to date, the effects of dimerization on GCS globin domain structure and conformation are still unknown.

Previous spectroscopic characterization of HemAT-*Bs* observed biphasic O_2 dissociation kinetics [33], which, through resonance Raman studies [32], were attributed to multiple open and closed conformations of the heme pocket. In addition, O_2 dissociation from dimeric and tetrameric assemblies of *Pcc*GCS and *Bpe*GReg was found to be biphasic [18], again suggesting multiple

conformations, and possibly hydrogen bonding patterns, within GCS heme pockets. To investigate the role of sensor globin domains in dimerization of full-length GCS proteins and the effects of globin-globin interactions on O_2 binding kinetics and hydrogen bonding interactions, isolated globin domains from *Pcc*GCS and *Bpe*GReg were investigated to determine if globin domains are important determinants of full-length GCS oligomerization and the effects of oligomerization on the heme pocket.

2. Experimental Results and Discussion

Previous studies found that GCS proteins from *P. carotovorum* and *B. pertussis* exist as mixtures of oligomeric states, with the dimeric and tetrameric assemblies having different catalytic activity, but consistent O_2 dissociation rates, for each respective GCS [18]. However, the interfaces between domains and molecular effects of oligomerization are currently unknown. To identify the effects of oligomerization on the heme pocket and ligand binding, truncations were made to *Pcc*GCS and *Bpe*GReg to yield isolated globin domains (*Pcc*Globin and *Bpe*Globin, respectively). In addition, mutations of putative distal pocket residues (tyrosine and serine; *Bpe*Globin(S68A) and (Y43A), *Pcc*Globin(S82A) and (Y57A)) were generated to investigate the roles of hydrogen bonding patterns in O_2 dissociation kinetics (Fig. 1A). Finally, mutations were introduced at the putative globin dimer interface (*Pcc*Globin(F148R) and (S155R)) to weaken globin-globin interactions and further investigate the role(s) of globin dimerization on ligand binding properties (Fig. 1B and 1C).

2.1. Heme Spectra of Globin Constructs

*Pcc*Globin and *Bpe*Globin WT and mutant constructs were expressed, purified, and characterized by UV-visible spectroscopy. Both proteins exhibit spectra that are characteristic of histidyl-ligated heme proteins and are nearly identical to the respective full-length proteins [18, 30] (Fig. 2A, 2B) and to other previously characterized sensor globins [14, 20-21, 30]. Similarly, ferrous *Pcc*Globin(S82A), *Bpe*Globin(S68A), and *Pcc*Globin(F148R) were able to bind all of the diatomic ligands (CO, NO, O_2) and exhibited spectra nearly identical to their respective WT globin constructs. In contrast, both tyrosine mutants, *Pcc*Globin(Y57F) and *Bpe*Globin(Y43F), rapidly oxidized when exposed to O_2, with no detectable stable O_2 binding, identifying the distal pocket tyrosine residues as the predominant hydrogen bond donors.

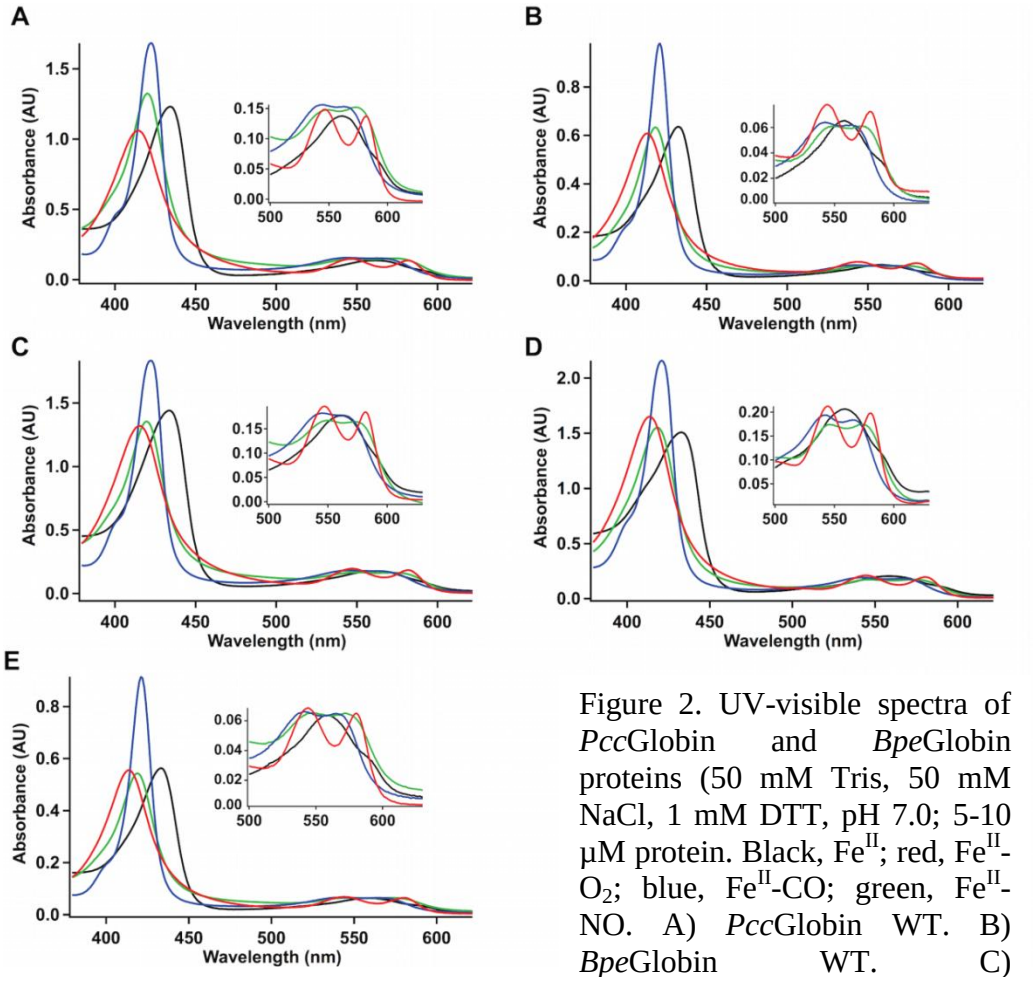

Figure 2. UV-visible spectra of *Pcc*Globin and *Bpe*Globin proteins (50 mM Tris, 50 mM NaCl, 1 mM DTT, pH 7.0; 5-10 µM protein. Black, Fe^{II}; red, Fe^{II}-O_2; blue, Fe^{II}-CO; green, Fe^{II}-NO. A) *Pcc*Globin WT. B) *Bpe*Globin WT. C)

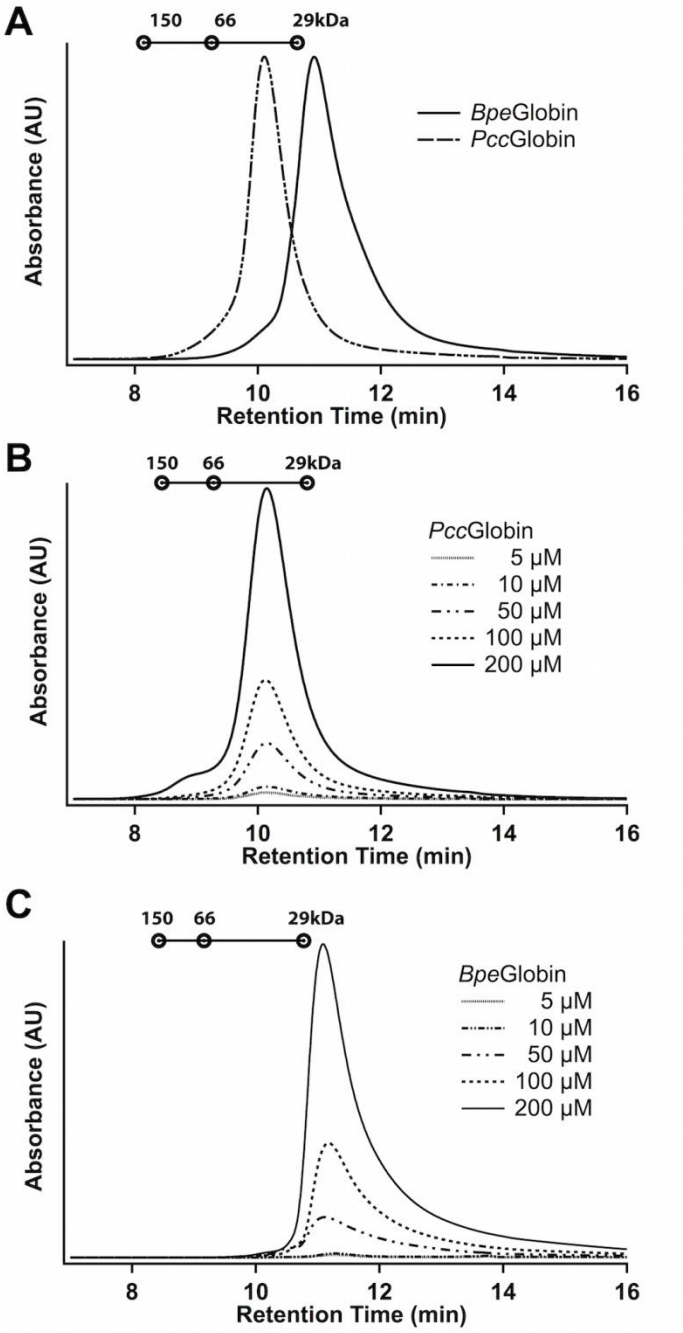

Figure 3. *Pcc*Globin and *Bpe*Globin oligomerization states (50 mM Tris, 50 mM NaCl, 1 mM DTT, pH 7.0).

2.2. Oligomeric States of WT Globin Constructs

Due to previous studies demonstrating that GCS proteins oligomerize [18], analytical gel filtration (AGF) was performed to determine the accessible oligomerization states. Both WT globin proteins were found to form predominantly (>90%) one oligomeric state, dimer for *Pcc*Globin (with <5% tetramer) and monomer for *Bpe*Globin (with <5% dimer) (Fig. 3A). Retention times for globular standards (molecular weight listed) are plotted above each graph. A) Comparison of Fe^{II}-O_2 *Pcc*Globin (dashed, dimer) and Fe^{II}-O_2 *Bpe*Globin (solid, monomer). B) *Pcc*Globin (Fe^{II}-O_2) oligomerization state does not depend on concentration. C) *Bpe*Globin (Fe^{II}-O_2) oligomerization state does not exhibit concentration dependence.

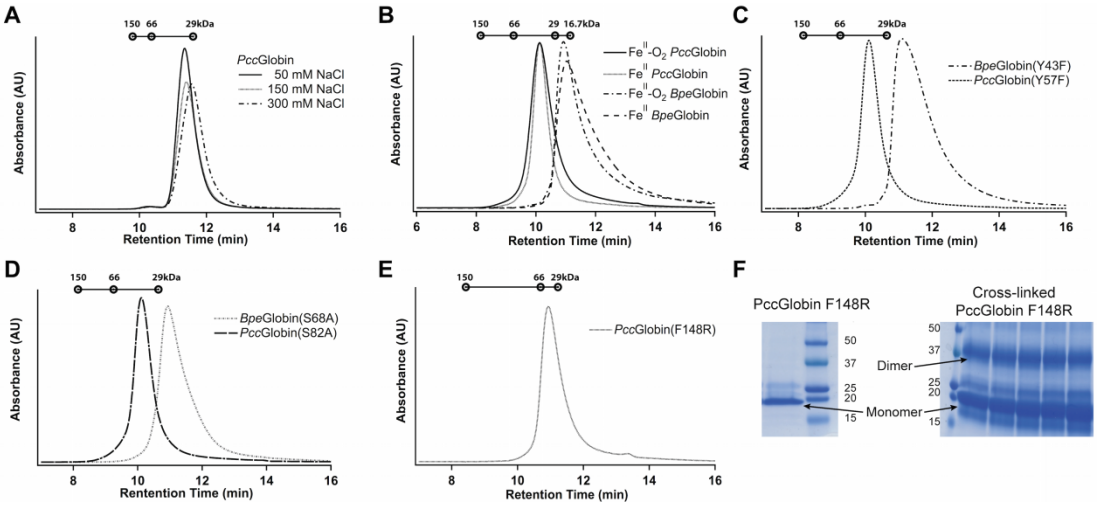

Figure 4. Salt, O_2, and mutation-dependent oligomerization. Retention times for globular standards (molecular weight listed) are plotted above each graph. A) *Pcc*Globin Fe^{II}-O_2 remains dimeric at a range of salt concentrations. B) *Pcc*Globin Fe^{II} exists as a dimer and *Bpe*Globin Fe^{II} exists as a monomer. C) Mutation of the distal pocket tyrosine does not affect oligomerization state (*Pcc*Globin and *Bpe*Globin Fe^{III}, due to rapid oxidation in air). D) Fe^{II}-O_2 *Pcc*Globin and *Bpe*Globin distal pocket serine mutants maintain WT oligomerization states. E) *Pcc*Globin(F148R) (Fe^{II}-O_2) mutation weakens dimer affinity. F) Non-crosslinked *Pcc*Globin runs as a monomer by denaturing PAGE (left), while addition of BS^3 crosslinking reagent results in formation of a clear dimer band (right).

Similar to full-length *Pcc*GCS and *Bpe*GReg proteins, *Pcc*Globin and *Bpe*Globin (36% identical, 61% similar) access different oligomeric states, despite the lack of obvious differences in amino acid similarity at the putative dimer interface. The dimerization affinities of WT *Pcc*Globin and *Bpe*Globin were interrogated by changing protein and salt concentrations. From 5-200 μM protein concentration, only minor shifts (<10%) in the oligomerization state of the proteins were observed and *Bpe*Globin did not form significant amounts of dimeric species even at the highest protein concentrations (Fig. 3B, 3C). Furthermore, a small change was observed for *Pcc*Globin at different salt concentrations (~6 kDa based on retention time), but the *Pcc*Globin dimer interface could not be disrupted, even at 300 mM NaCl (Fig. 4A).

These results show that the globin domains exhibit similar properties to the full-length GCS proteins, namely that oligomerization state is not significantly modulated by protein concentration or ionic strength [18]. Additionally, neither *Pcc*Globin nor *Bpe*Globin shifted oligomeric states when examined under anaerobic conditions; therefore O_2 binding is not responsible for altering protein conformation to allow globin dimerization (Fig. 4B). Similarly, anaerobic AGF analysis of full-length *Pcc*GCS demonstrated that O_2 binding controlled the dimer-tetramer-octamer/high molecular weight (HMW) equilibrium, but Fe^{II} *Pcc*GCS did not yield any measurable monomer formation.[19] Therefore, it is likely that O_2 binding causes conformational changes that propagated through the full-length GCS proteins, altering protein-protein interactions and altering oligomer affinity without changing globin domain dimerization.

These data suggest that the globin domains are key determinants of the ability of full-length GCS proteins to exist as monomeric proteins. *Bpe*Globin is monomeric and *Bpe*GReg is observed in the monomeric state and exists as a monomer-dimer-tetramer mixture [18]. In contrast, *Pcc*Globin is dimeric, and *Pcc*GCS has not been observed as a monomer; instead, *Pcc*GCS exists as a dimer-tetramer-octamer/HMW mixture [18]. Strong interactions between globin domains in *Pcc*GCS likely preclude formation of monomeric *Pcc*GCS, while weak interactions in *Bpe*Globin readily allow *Bpe*GReg monomer formation. As less than 5% of *Pcc*Globin forms tetrameric assemblies, additional interactions clearly are required to generate higher order oligomers (tetramers and octamers); however, globin domain dimer affinity controls the accessibility of monomeric full-length proteins.

2.3. Characterization of Globin Mutants

As the oligomerization state of *Pcc*Globin and *Bpe*Globin could not be altered by protein or salt concentration, weakening the globin-globin interface was probed through mutagenesis. Using a homology model based on crystal structures of other sensor globins [29, 31, 34], residues putatively lining the interface were identified (Fig. 1B, 1C, S2). Phenylalanine148 in *Pcc*Globin was chosen for mutagenesis as it is proposed to face directly into the coiled-coil of the dimer interface. Introduction of the F148R mutation in *Pcc*Globin resulted in a shift to monomeric protein, as assessed by AGF (Fig. 4E). However, further analysis using protein-protein cross-linking found that *Pcc*Globin F148R still was able to form a dimer in solution (>35% dimer, based on cross-linking gel integration), although with significantly reduced dimer affinity (Fig. 4F).

*Pcc*Globin(S82A) and *Bpe*Globin(S68A) remained dimeric and monomeric, respectively, with no change as compared to their corresponding WT globin (Fig. 4D). In addition, from the AGF traces, both Tyr mutations remained in the same oligomerization state (dimer and monomer, respectively) as the respective WT globin (Fig. 4C). Taken together, these studies demonstrate that mutations within the heme pocket that alter hydrogen-bonding patterns do not affect oligomerization state of the globin domains. In addition, *Pcc*Globin dimer affinity can be modulated through introduction of charged residues at the putative dimer interface, which supports the homology model structure and further highlights the high affinity of *Pcc*Globin dimers.

2.4. Effects of Dimerization on Oxygen Binding Kinetics

Previous work on full-length *Pcc*GCS and *Bpe*GReg found oligomerization-dependent effects on both cyclase activity and the two O_2 dissociation rates (Table 1) [18]. The different oligomerization states of the isolated globin domains provided a model system to interrogate the role of the globin-globin interface in altering O_2 binding kinetics. Comparing k_{off} rates (k_1 and k_2**)** of previous obtained *Pcc*GCS and *Bpe*GReg [18] to *Pcc*Globin and *Bpe*Globin, respectively, it can be noted that dissociation rates do not significantly change from full-length protein to globin truncation (Table 1). However, *Bpe*Globin exhibits a loss of the fast rate (k_2),

while *Pcc*Globin exhibits a large decrease in fast rate percentage (44% vs. 21% for *Pcc*GCS and *Pcc*Globin, respectively). The complete loss of k_2 for *Bpe*Globin, compared to the biphasic O_2 dissociation kinetics of *Pcc*Globin, suggests that the globin domain dimer interface is required for effective formation of the conformation(s) that lead to the fast rate. In addition, the decrease in the amplitude of k_2 observed for *Pcc*Globin demonstrates that the remaining two domains in the full-length GCS proteins play important roles in altering heme pocket conformation by contributing to a state with decreased O_2 stabilization, thereby contributing to the fast rate. Removal of the middle and output domains of HemAT-*Bs* also resulted in changes in O_2 dissociation and loss of the heme pocket closed conformation [32-33], supporting roles for these domains in affecting the heme pocket.

Table 1. O_2 binding kinetics (k_{on}, O_2 association rate; k_1 and k_2, O_2 dissociation rates ; % k_1 and % k_2, amplitude of each respective O_2 dissociation rate). Significant oligomeric states (≥10%) in the solution of each protein during binding measurements are listed.

Protein	**Oligomeric state(s)**	k_{on} ($\mu M^{-1} s^{-1}$)	k_1 (s^{-1})	k_2 (s^{-1})	% k_1[a]	% k_2[a]	K_D (nM)	
*Bpe*GReg	Monomer, dimer, tetramer	7.0[b]	0.82 ± 0.01[c]	6.3 ± 0.1[c]	39[c]	60[c]	117	900
*Bpe*Globin	Monomer	17.6 ± 0.8	1.18 ± 0.04	-	100	-	67	-
*Bpe*Globin (S68A)	Monomer	20.3 ± 0.5	2.08 ± 0.01	-	100	-	102	-
*Bpe*Globin (Y43F)	Monomer	-	-	-	-	-	-	-
*Pcc*GCS	Dimer, tetramer, HMW	7.0 ± 0.3[d]	0.56 ± 0.01[c]	3.87 ± 0.08[c]	56[c]	44[c]	78	538
*Pcc*GCS Dimer	Dimer	N.D.	0.67 ± 0.02 [c]	4.7 ± 0.3 [c]	56[c]	44[c]	N.D.	N.D.
*Pcc*Globin	Dimer	6.1 ± 0.3	0.66 ± 0.01	5.8 ± 0.1	79	21	108	944
*Pcc*Globin (S82A)	Dimer	7.9 ± 0.5	3.02 ± 0.01	-	100	-	382	-
*Pcc*Globin (Y57F)	Dimer	-	-	-	-	-	-	-
*Pcc*Globin (F148R)	Monomer, dimer	7.4 ± 0.1	1.22 ± 0.01	6.98 ± 0.07	86	14	165	943

- Not observed
N.D., not determined
HMW, high molecular weight (octamer and larger)
[a] All errors for % k_1 and % k_2 are ≤ 5%
[b] Ref. *[30]*
[c] Ref. *[18]*
[d] Ref. *[17]*

The effect of salt concentration on *Pcc*Globin O_2 dissociation rates also was measured to determine if the higher salt concentrations, which can be accessed using stopped flow as compared to AGF (1 M vs. 300 mM, respectively), would weaken salt bridges putatively stabilizing the globin dimer interface. Despite the high salt concentration, no change in O_2

dissociation rates or rate percentages was observed (Table 2). These data further support the AGF data that show that the *Pcc*Globin dimer interface has high affinity and is a key determinant of dimer formation in full-length *Pcc*GCS.

Table 2. *Pcc*Globin WT salt-dependent O_2 dissociation rates.

Salt concentration	k_1 (s^{-1})	k_2 (s^{-1})	% k_1	% k_2
50 mM NaCl	0.681 ± 0.002	5.48 ± 0.08	82.88 ± 0.5	17.12 ± 0.5
300 mM NaCl	0.598 ± 0.001	3.5 ± 0.1	79.39 ± 0.8	20.61 ± 0.8
500 mM NaCl	0.584 ± 0.002	4.64 ± 0.08	82.88 ± 0.4	17.12 ± 0.4
750 mM NaCl	0.543 ± 0.002	4.0 ± 0.2	82.27 ± 1.2	17.73 ± 1.2
1000 mM NaCl	0.563 ± 0.001	5.01 ± 0.05	80.67 ± 0.5	19.33 ± 0.5

Measurement of *Pcc*Globin(F148R) O_2 dissociation rates in varying salt concentrations found that 0 mM NaCl resulted in k_1 and k_2 percentages similar to WT *Pcc*Globin, likely due to increased dimer formation, while inclusion of salt decreased the percentage of k_2 (Tables 1 and 3), potentially by stabilizing monomeric *Pcc*Globin(F148R). These data highlight the importance of the globin dimer interface in controlling conformations of the heme pocket and O_2 affinity within globin coupled sensors. An additional mutation (S155R) putatively located at the center of the dimer interface was generated to further disrupt *Pcc*Globin dimerization However, this mutation resulted in unstable protein that readily aggregated, prohibiting analysis of the effects of dimerization on O_2 binding kinetics.

To further characterize the effects of globin dimerization on the heme pocket, O_2 association rates were measured (Table 1). *Pcc*Globin exhibited an association rate (k_{on}) almost identical to full-length *Pcc*GCS (6.1 $\mu M^{-1}s^{-1}$ and 7.2 $\mu M^{-1}s^{-1}$ [17], respectively), demonstrating that removing the middle and diguanylate cyclase domains, but maintaining the globin dimer interface, does not significantly affect O_2

Table 3**.** *Pcc*Globin(F148R) salt-dependent O_2 dissociation rates.

Salt Concentration	**k_1 (s^{-1})**	**k_2 (s^{-1})**	**% k_1**	**% k_2**
0 mM NaCl	1.06 ± 0.002	6.38 ± 0.04	80.5 ± 2.4	19.5 ± 2.4
50 mM NaCl	1.224 ± 0.002	6.98 ± 0.07	85.9 ± 0.5	14.1 ± 0.5
150 mM NaCl	1.203 ± 0.002	7.43 ± 0.09	86.2 ± 1.1	13.8 ± 1.4
300 mM NaCl	1.226 ± 0.003	6.97 ± 0.08	85.5 ± 1.9	14.5 ± 1.9

entry into the heme pocket. In contrast, *Bpe*Globin exhibited a markedly faster O_2 association rate, as compared to full-length *Bpe*GReg (17.6 $\mu M^{-1}s^{-1}$ vs. 7.0 $\mu M^{-1}s^{-1}$ [30], respectively). The biphasic O_2 dissociation rates yield calculated O_2 affinities of 108 nM and 944 nM for *Pcc*Globin, as compared to 78 nM and 538 nM for full-length *Pcc*GCS. In contrast, *Bpe*Globin yields a single 67 nM O_2 affinity, while O_2 affinities of 117 nM and 900 nM were calculated for full-length *Bpe*GReg. Dual O_2 affinities in both full-length proteins may allow the organisms to sense a wide range of O_2 concentrations, allowing the bacteria to modulate their responses to adapt to moderate and then low O_2 levels.

As truncating *Pcc*GCS did not result in a change in O_2 association rate, it is unlikely that the change in rate for *Bpe*Globin is due to the loss of the middle and cyclase domains. Instead, loss of the globin dimerization interface may be responsible for the 2.5-fold increase in the rate of O_2 binding. Given that crystal structures of homologous sensor globins solved as dimers found that entry to the heme pocket was away from the globin-globin interface [29, 31] (Fig. 1), the changes in k_{on} are unlikely to be due to O_2 access being blocked by the dimer interface. Instead, both the O_2 association and dissociation rate data for *Bpe*Globin suggest that loss of the globin dimer interface results in significant conformational changes to the heme pocket, increasing accessibility and altering conformations/dynamics within the pocket.

2.5. Involvement of Distal Pocket Residues in Biphasic O_2 Dissociation Rates

Based on crystal structures and sequence alignments, *Pcc*Globin and *Bpe*Globin are proposed to have two amino acids (Y57/S82 and Y43/S68, respectively) within the distal pocket that serve as hydrogen bond donors and stabilize O_2 binding (Fig. 1A, S2) [16, 18-19, 29]. Within HemAT-*Bs*, both the distal pocket tyrosine and threonine contribute to O_2 stabilization, while *Ec*DosC only has a distal tyrosine to provide hydrogen bonding [16, 19, 32-33]. Therefore, mutations of both distal pocket residues within *Pcc*Globin and *Bpe*Globin were analyzed to determine which residue is involved in the fast O_2 dissociation rate (k_2).

Given the rapid autoxidation, O_2 association and dissociation rates for *Pcc*Globin(Y57F) and *Bpe*Globin(Y43F) could not be obtained. In contrast, both serine mutants, *Pcc*Globin(S82A) and *Bpe*Globin(S68A), were able to bind O_2; however, even though *Pcc*Globin(S82A) forms a dimer (Fig. 4D), the O_2 dissociation rate simplified to a single exponential (Table 1), suggesting that the distal pocket serine is involved in interactions with bound O_2 and/or distal pocket residues that lead to biphasic kinetics and the fast O_2 dissociation rate (k_2). Mutation of *Pcc*Globin distal pocket serine also leads to an increase in k_1, which is not surprising, given that a secondary hydrogen-bonding residue has been removed, and is similar to the increase observed for *Bpe*Globin(S68A). The O_2 association rates for the serine mutants of *Pcc*Globin and *Bpe*Globin did not change relative to the WT proteins [17, 30] (Table 1), further confirming that the mutations did not significantly affect the heme pocket. These data support the hypothesis that globin domain dimerization leads to conformation(s) of the distal pocket that adjust the position of the

hydrogen bonding residues, leading to increased interaction with the distal pocket serine and resulting in biphasic O_2 dissociation (Fig. 5).

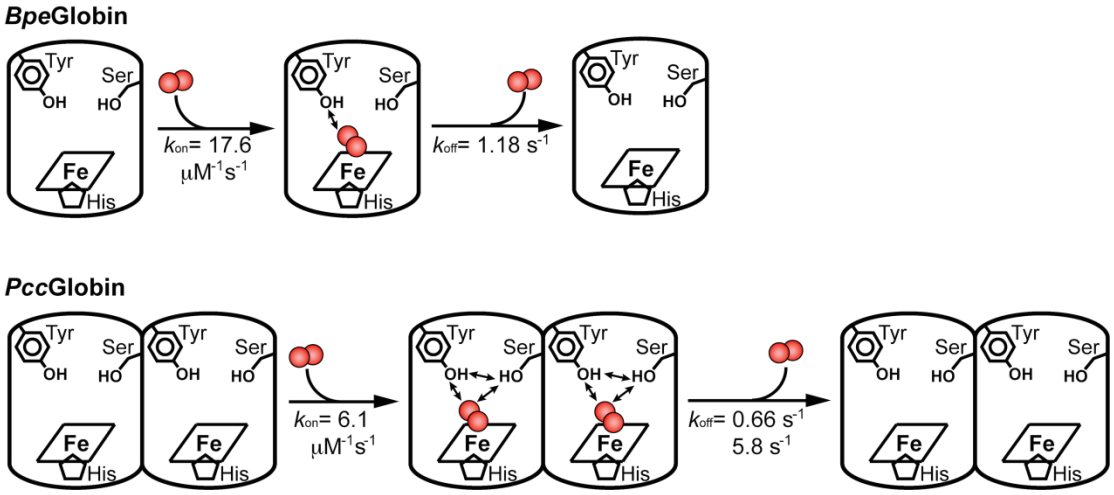

Figure 5. Overview of the effects of globin dimerization on O_2 binding kinetics. Monomeric *Bpe*Globin (top) exhibits a single O_2 dissociation rate, with the distal pocket tyrosine serving as the primary hydrogen bond donor. In contrast, dimeric *Pcc*Globin (bottom) exhibits two O_2 dissociation rates, the faster rate requiring the presence of the distal pocket serine. Dimerization causes a conformational change in the distal pocket, positioning the serine such that it affects O_2 dissociation through interactions with bound O_2 and/or the distal pocket tyrosine.

3. Conclusions

In conclusion, globin domains from *Pcc*GCS and *Bpe*GReg exist in different oligomeric states, despite the relatively high amino acid sequence similarity at the putative dimer interface. These differences in oligomerization result in altered properties of the heme domains. Globin domain dimerization within globin coupled sensors leads to biphasic O_2 dissociation, likely through changes in the conformation of the heme pocket. Of the two distal pocket hydrogen bond donors, tyrosine provides direct hydrogen bonding, while a serine residue serves as a secondary hydrogen bond donor. In addition, *Pcc*Globin distal pocket serine makes key interactions involved in the second, fast O_2 dissociation rate observed in dimeric sensor globins, suggesting

that dimerization changes the accessible heme pocket conformations and position of the serine. These data provide insight into the oligomerization determinants of the full-length GCS proteins and information about the oligomerization-dependent changes to the heme pocket GCS globin domains.

4. Chapter References

1. Karatan, E.; Watnick, P., Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol. Mol. Biol. Rev.* **2009,** *73* (2), 310-47.

2. Srivastava, D.; Waters, C. M., A tangled web: regulatory connections between quorum sensing and cyclic Di-GMP. *J Bacteriol* **2012,** *194* (17), 4485-93.

3. Cypionka, H., Oxygen respiration by desulfovibrio species. *Annu. Rev. Microbiol.* **2000,** *54*, 827-48.

4. Sass, H.; Berchtold, M.; Branke, J.; Konig, H.; Cypionka, H.; Babenzien, H. D., Psychrotolerant sulfate-reducing bacteria from an oxic freshwater sediment, description of Desulfovibrio cuneatus sp. nov. and Desulfovibrio litoralis sp. nov. *Syst Appl Microbiol* **1998,** *21* (2), 212-219.

5. Wu, C.; Cheng, Y. Y.; Yin, H.; Song, X. N.; Li, W. W.; Zhou, X. X.; Zhao, L. P.; Tian, L. J.; Han, J. C.; Yu, H. Q., Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. *Sci. Rep.* **2013,** *3*, 1945-1951.

6. Bauer, C. E.; Elsen, S.; Bird, T. H., Mechanisms for redox control of gene expression. *Annu Rev Microbiol* **1999,** *53*, 495-523.

7. Gilles-Gonzalez, M. A., Oxygen signal transduction. *IUBMB Life* **2001,** *51* (3), 165-73.

8. Green, J.; Crack, J. C.; Thomson, A. J.; LeBrun, N. E., Bacterial sensors of oxygen. *Curr Opin Microbiol* **2009,** *12* (2), 145-51.

9. Shimizu, T.; Huang, D.; Yan, F.; Stranava, M.; Bartosova, M.; Fojtikova, V.; Martinkova, M., Gaseous O2, NO, and CO in signal transduction: structure and function relationships of heme-based gas sensors and heme-redox sensors. *Chem. Rev.* **2015,** *115* (13), 6491-533.

10. Vinogradov, S. N.; Tinajero-Trejo, M.; Poole, R. K.; Hoogewijs, D., Bacterial and archaeal globins - A revised perspective. *Biochim. Biophys. Acta* **2013,** *1834*, 1789-1800.

11. Martinkova, M.; Kitanishi, K.; Shimizu, T., Heme-Based Globin-Coupled Oxygen Sensors: Linking Oxygen Binding to Functional Regulation of Diguanylate Cyclase, Histidine Kinase and Methyl-accepting Chemotaxis. *J. Biol. Chem.* **2013,** *288*, 27702-27711.

12. Hou, S.; Belisle, C.; Lam, S.; Piatibratov, M.; Sivozhelezov, V.; Takami, H.; Alam, M., A globin-coupled oxygen sensor from the facultatively alkaliphilic Bacillus halodurans C-125. *Extremophiles* **2001,** *5* (5), 351-4.

13. Hou, S.; Freitas, T.; Larsen, R. W.; Piatibratov, M.; Sivozhelezov, V.; Yamamoto, A.; Meleshkevitch, E. A.; Zimmer, M.; Ordal, G. W.; Alam, M., Globin-coupled sensors: a class of heme-containing sensors in Archaea and Bacteria. *Proc. Natl. Acad. Sci. USA* **2001,** *98* (16), 9353-8.

14. Hou, S.; Larsen, R. W.; Boudko, D.; Riley, C. W.; Karatan, E.; Zimmer, M.; Ordal, G. W.; Alam, M., Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* **2000,** *403* (6769), 540-4.

15. Kitanishi, K.; Kobayashi, K.; Uchida, T.; Ishimori, K.; Igarashi, J.; Shimizu, T., Identification and functional and spectral characterization of a globin-coupled histidine kinase from Anaeromyxobacter sp. Fw109-5. *J. Biol. Chem.* **2011,** *286* (41), 35522-34.

16. Zhang, W.; Phillips, G. N., Jr., Crystallization and X-ray diffraction analysis of the sensor domain of the HemAT aerotactic receptor. *Acta Crystallogr D Biol Crystallogr* **2003,** *59* (Pt 4), 749-51.

17. Jia, X.; Wang, J.-b.; Rivera, S.; Duong, D.; Weinert, E. E., An O2-Sensing Stressosome from a Gram Negative Bacterium. *Nature Commun.* **2016,** *7*, 12381.

18. Burns, J. L.; Deer, D. D.; Weinert, E. E., Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. *Mol. Biosyst.* **2014,** *10* (11), 2823-6.

19. Kitanishi, K.; Kobayashi, K.; Kawamura, Y.; Ishigami, I.; Ogura, T.; Nakajima, K.; Igarashi, J.; Tanaka, A.; Shimizu, T., Important roles of Tyr43 at the putative heme distal side in the oxygen recognition and stability of the Fe(II)-O2 complex of YddV, a globin-coupled heme-based oxygen sensor diguanylate cyclase. *Biochemistry* **2010,** *49* (49), 10381-93.

20. Sawai, H.; Yoshioka, S.; Uchida, T.; Hyodo, M.; Hayakawa, Y.; Ishimori, K.; Aono, S., Molecular oxygen regulates the enzymatic activity of a heme-containing diguanylate cyclase (HemDGC) for the synthesis of cyclic di-GMP. *Biochim. Biophys. Acta* **2010,** *1804* (1), 166-72.

21. Tuckerman, J. R.; Gonzalez, G.; Sousa, E. H.; Wan, X.; Saito, J. A.; Alam, M.; Gilles-Gonzalez, M. A., An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* **2009,** *48* (41), 9764-74.

22. Hengge, R., Principles of c-di-GMP signalling in bacteria. *Nature Rev. Microbiol.* **2009,** *7* (4), 263-73.

23. Romling, U.; Galperin, M. Y.; Gomelsky, M., Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* **2013,** *77* (1), 1-52.

24. Chan, C.; Paul, R.; Samoray, D.; Amiot, N. C.; Giese, B.; Jenal, U.; Schirmer, T., Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* **2004,** *101* (49), 17084-9.

25. Deepthi, A.; Liew, C. W.; Liang, Z. X.; Swaminathan, K.; Lescar, J., Structure of a diguanylate cyclase from Thermotoga maritima: insights into activation, feedback inhibition and thermostability. *PLoS One* **2014,** *9* (10), e110912.

26. De, N.; Navarro, M. V.; Raghavan, R. V.; Sondermann, H., Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. *J. Mol. Biol.* **2009,** *393* (3), 619-33.

27. Paul, R.; Abel, S.; Wassmann, P.; Beck, A.; Heerklotz, H.; Jenal, U., Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J. Biol. Chem.* **2007,** *282* (40), 29170-7.

28. Wassmann, P.; Chan, C.; Paul, R.; Beck, A.; Heerklotz, H.; Jenal, U.; Schirmer, T., Structure of BeF3- -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* **2007,** *15* (8), 915-27.

29. Tarnawski, M.; Barends, T. R.; Schlichting, I., Structural analysis of an oxygen-regulated diguanylate cyclase. *Acta Crystallogr D Biol Crystallogr* **2015,** *71* (Pt 11), 2158-77.

30. Wan, X.; Tuckerman, J. R.; Saito, J. A.; Freitas, T. A.; Newhouse, J. S.; Denery, J. R.; Galperin, M. Y.; Gonzalez, G.; Gilles-Gonzalez, M. A.; Alam, M., Globins synthesize the second messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* **2009,** *388* (2), 262-70.

31. Zhang, W.; Phillips, G. N., Jr., Structure of the oxygen sensor in *Bacillus subtilis*: signal transduction of chemotaxis by control of symmetry. *Structure* **2003,** *11* (9), 1097-110.

32. Ohta, T.; Yoshimura, H.; Yoshioka, S.; Aono, S.; Kitagawa, T., Oxygen-sensing mechanism of HemAT from *Bacillus subtilis:* a resonance Raman spectroscopic study. *J. Am. Chem. Soc.* **2004,** *126* (46), 15000-1.

33. Zhang, W.; Olson, J. S.; Phillips, G. N., Jr., Biophysical and kinetic characterization of HemAT, an aerotaxis receptor from *Bacillus subtilis*. *Biophys. J.* **2005,** *88* (4), 2801-14.

34. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **2015,** *10* (6), 845-58.

Chapter 3: Structural Insights into Globin Coupled Sensors

Adapted from: Rivera, S.; Paul, P.G.; Hoffer, E.D.; Vansuch, G.E.; Metzler, C.L.; Dunham, C.M.; Weinert, E.E. (2018) Structural Insights into Oxygen-Dependent Signal Transduction within Globin Coupled Sensors. *Inorg. Chem. 57,* 14386-14395. DOI: 10.1021/acs.inorgchem.8b02584, with permission form ©2019 American Chemical Society.

Chapter 3:

Structural Insights into Globin Coupled Sensors

Chapter Overview:

1. Introduction
2. Experimental Results and Discussion
   2.1 Structure of Ferrous-Oxy and Mixed *Bpe*Globin Dimers
   2.2 Comparison to Other Crystalized GCS Proteins
   2.3 Heme Distortion in *Bpe*Globin Influenced by O_2 Binding
   2.4 Heme Distortion Comparison to Other GCS Proteins
   2.5 Interactions between Distal Residues and Bound Ligand
   2.6 Examination of Structural Water
3. Conclusion
4. Chapter References

1. 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.[7-8] As the presence of molecular oxygen (O_2) 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 O_2 levels.[1, 9-11] The globin coupled sensor (GCS) family of bacterial heme sensor proteins monitors O_2 levels and modulates downstream pathways for adaption,[7] including O_2-dependent biofilm formation,[10, 12-14] motility,[12, 14-15] and virulence.[12, 14-16] Because biofilms are a primary defense against antibiotic treatment, host immune responses, and environmental stress, understanding the mechanisms by which O_2 levels regulate GCS activity has the potential to identify new antibacterial targets.[7-8, 17]

GCS proteins are widely distributed throughout bacteria, as well as in a number of archaea, suggesting their widespread importance in controlling O_2-dependent bacterial phenotypes.[19] These proteins consist of a 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-σ factor domains, and kinases (Fig. 1A).[1, 5, 9, 16, 19-23] 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.[12, 15] 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.[1, 5, 10-11, 16, 24-25]

While the ligand-dependent enzymatic functions of these proteins have been investigated, the mechanism of activation for GCS proteins remains unknown.[1, 5, 11, 16, 26] This is in part due to subtle differences between closely related GCS proteins.[7] GCS proteins containing a DGC domain must dimerize to exhibit activity,[27-28] however GCS proteins from *Bordetella pertussis* (*Bpe*GReg) and *Pectobacterium carotovorum* (*Pcc*GCS) exist in different mixtures of oligomeric states (*i.e.* monomer-dimer-tetramer and dimer-tetramer-octamer, respectively),[1, 6, 11, 24, 2] while the GCS from *Desulfotalea psychrophile* (HemDGC) is only found as a tetramer.[11] For both *Bpe*GReg and *Pcc*GCS, the tetrameric assemblies with O_2 bound exhibit the highest activity.[1-2] Furthermore, differences in dimerization affinity of the isolated globin domains were demonstrated to alter O_2 binding kinetics,[3] suggesting that oligomeric state alters the heme pocket environment and thus plays an important role in signaling.[2, 6-7] 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 (*Af*GcHK), 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 activation.[29] 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)-O_2 state is prone to autoxidation.[6, 25, 29] However, the Fe(II)-O_2 state is typically the high activity form of GCS proteins, so this ligation state is the most informative for elucidating biologically relevant signaling mechanisms.[7] *Bpe*GReg and *Pcc*GCS have greater oxygen binding stability than other GCS proteins, such as *Ec*DosC, HemAT-*Bs* and *Af*GcHK, as demonstrated by their lower oxygen dissociation rates and their ability to remain in the Fe(II)-O_2 state for extended periods.[1, 4-5, 16, 30] Furthermore, the isolated globin domains from *Bpe*GReg and *Pcc*GCS (*Bpe*Globin and *Pcc*Globin) have been shown to be even more stable than their full length counterparts.[3] For example, whereas *Bpe*GReg and *Pcc*GCS can remain Fe(II)-O_2 at 4° C for ~18 hours before showing significant autoxidation, *Bpe*Globin and *Pcc*Globin remain fully Fe(II)-O_2 after 72 hours at 20° C.[1-3] Consequently, *Bpe*Globin and *Pcc*Globin are excellent systems to study the mechanism of O_2 sensing in these important signaling proteins.

To probe the mechanism of O_2-dependent activation and the molecular basis for subtle differences between closely related GCS proteins, X-ray crystallography and FTIR were performed on *Bpe*GReg, *Pcc*GCS, 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.

2. Experimental Results and Discussion

2.1 Structure of Ferrous-Oxy and Mixed Bpe*Globin Dimers*

The globin domain of *Bpe*GReg crystallized in the spacegroup C121 and contained three monomeric units of *Bpe*Globin in the asymmetric unit (Fig. 1, Table 1). However, the biological unit is a dimer as assessed by Pisa[31] and based on previous biochemical studies.[2-3] There are two species of biological dimers formed within the crystal lattice. The first dimer contains two monomers with O_2 bound to the heme (hereafter referred to as the ferrous-oxy dimer) while the second dimer consists of one O_2-bound and one Fe(III)-H_2O monomer (hereafter referred to as the mixed dimer) (Fig. 1B). *Bpe*Globin adopts the classic globin fold of an eight α-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 O_2 and is engaged in a direct hydrogen bond with heme distal pocket residues, suggesting a potential role in ligand binding and stabilization (Table 2).

This is the first reported crystal structure of a bacterial sensor globin domain in the Fe(II)-O_2 ligation state, offering a novel opportunity to examine a signaling domain in its native active form (Fig. 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)-O_2 *Bpe*GReg is stable for at least 18 hrs with no significant oxidation, aggregation, or oligomerization changes.[1] The exceptionally low autoxidation and O_2 dissociation rates of *Bpe*GReg likely enabled the extended stability of the Fe(II)-O_2 protein necessary for its structural determination.[1]

2.2 Comparison to Other Crystalized GCS Proteins

To investigate the conservation of potentially relevant signaling residues, the structure of *Bpe*Globin was compared to previously published structures of the globin domains from *Af*GcHK,[29] HemAT-*Bs*,[25, 32] and *Ec*DosC[6]. Alignments of the globin domains of HemAT-*Bs*, *Af*GcHK and *Ec*DosC with *Bpe*Globin show similar helical arrangement as well as a conserved hydrogen bonding Tyr residue within the distal pocket. The crystal structure of *Bpe*Globin reveals shorter distances between the distal hydrogen bonding residues (Tyr43 and Ser68) than *Ec*DosC (Tyr43), HemAT-*Bs* (Tyr70 and Thr95), and *Af*GcHK (Tyr45 and Thr71; Table 2).

Table 1. X-ray structure statistics.

	***Bpe*Globin**
Wavelength	1.000
Resolution range	35.97 - 2.3 (2.38 - 2.30)
Space group	C 1 2 1
Unit cell	145.16, 49.64, 68.14, 90, 110.59, 90
Total reflections	169919 (18748)
Unique reflections	18911 (2040)
Multiplicity	9.0 (9.2)
Completeness (%)	92.14 (99.66)
Mean I/sigma(I)	26.08 (19.21)
Wilson B-factor	38.06
R-merge	0.07781 (0.1159)
R-meas	0.08293 (0.1227)
R-pim	0.02807 (0.03971)
CC1/2	0.993 (0.994)
CC*	0.998 (0.999)
Reflections used in refinement	18882 (2033)
Reflections used for R-free	1890 (205)
R-work (%)	16.14 (19.75)
R-free (%)	25.27 (29.43)
CC(work)	0.969 (0.917)
CC(free)	0.953 (0.817)
Number of non-hydrogen atoms	3986
macromolecules	3664
ligands	136
solvent	186
Protein residues	469
RMS(bonds)	0.008
RMS(angles)	1.12
Ramachandran favored (%)	98.27
Ramachandran allowed (%)	1.73
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	1.32
Clashscore	5.08
Average B-factor	46.12
macromolecules	45.94
ligands	46.52
solvent	49.25

*Statistics for the highest-resolution shell are shown in parentheses.

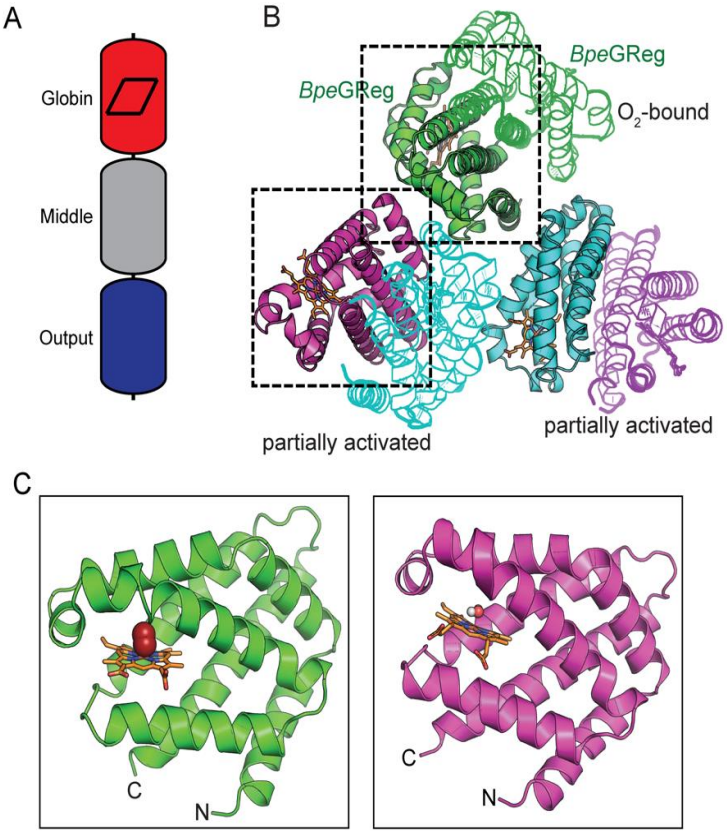

Figure 1. Structure of *Bpe*Globin. (A) Schematic of GCS domain architecture (left) and structure of the *Bpe*Globin. (B) Orientation of the two different dimeric forms. Each monomer forms associations via salt bridges along helices A, B, and E in the asymmetric unit (solid pink, green and blue). The ferrous-oxy dimer is shown in green and the mixed dimer is shown in pink and blue. (C) Close ups for both the Fe(II)-O_2 (green insert, left) and the Fe(III)-H_2O (pink insert, right) show heme pockets without significant differences (RMSD = 0.57 Å).

Additionally, the distances between *Bpe*Globin 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 *Af*GcHK also contain water molecules, which are ~4.6 and ~5 Å from the heme irons, respectively.[25, 29, 32] In the case of HemAT-*Bs*, this ordered water has been shown to interact with bound O_2 through resonance Raman studies.[33] The water molecule for *Af*GcHK has not been further examined as of yet. However, no water was found within the distal pocket of *Ec*DosC due to the presence of Leu92, which acts as a hydrophobic gate preventing the water from entering the heme pocket.[30] Furthermore, *Ec*DosC does not contain a second distal pocket hydrogen-bond

Table 2. Inter-residue distances in Å, including non-ligated water molecules.

Protein/Residues	**Chain A**	**Chain B**	**Chain C**	**Chain D**
***Bpe*GReg (PDB: 6M9A)**	Fe(II)	Fe(II)	*Fe(III)	
Fe - His99	2.1 ±0.4	1.9 ±0.5	2.0 ±0.5	-
Fe - Tyr43	4.5 ±0.4	4.9 ±0.5	4.7 ±0.5	-
Fe - Ser68	6.3 ±0.4	6.2 ±0.5	6.0 ±0.5	-
Fe- Trp72	8.1 ±0.4	8.0 ±0.5	8.3 ±0.4	-
Tyr43 - Ser68	4.5 ±0.4	4.1 ±0.5	4.3 ±0.5	-
O_2- Fe(II)	1.8 ±0.4	1.9 ±0.5	-	-
O_2- Tyr43	2.3 ±0.4	2.4 ±0.5	-	-
O_2- Ser68	5.1 ±0.4	5.2 ±0.5	-	-
H_2O - Tyr43	2.9 ±0. 4	2.5 ±0.4	2.5 ± 0.4	-
H_2O - Ser68	2.2 ±0.4	2.1 ±0.4	2.3 ± 0.4	-
H_2O - Fe	4.7 ±0. 4	3.5 ±0.4	4.2 ± 0.4	-
HemAT-*Bs*[25, 32]				
Fe(III)-CN (PDB: 1OR4)				
Fe(III) - His180	2.0	2.0	-	-
Fe(III) - Tyr70	5.7	5.6	-	-
Fe(III) - Thr95	7.1	6.9	-	-
Tyr70 - Thr95	5.7	5.7	-	-
H_2O - Tyr70	4.0	3.6	-	-
H_2O - Thr95	2.7	2.8	-	-
H_2O - Fe(III)	4.8	4.5	-	-
Fe(III) (PDB: 1OR6)			-	-
Fe(III) - His180	2.8	2.1	-	-
Fe(III) - Tyr70	5.8	6.3	-	-
Fe(III) - Thr95	6.8	7.6	-	-
Tyr70 - Thr95	5.3	5.9	-	-
***Ec*DosC[6]**				
Fe(II) (PDB: 4ZVB)				
Fe(II) - His99	2.2	2.2	2.2	2.2
Fe(II) - Tyr43	7.0	7.2	7.0	7.0
Fe(III) (PDB: 4ZVA)				
Fe(III) - His99	2.2	2.2	-	-
Fe(III) - Tyr43	6.6	6.7	-	-

donor, suggesting that the water may contribute to the different catalytic, ligand binding, and oligomerization behaviors of the various GCS proteins.[6]

2.3 Heme Distortion in Bpe*Globin influenced by O_2 binding*

Heme deformation is a critical component in the signaling mechanism of several heme-based sensor protein families.[18, 34-35] To determine if heme distortion could be involved in signal transduction within GCS proteins, the heme cofactors of different monomers of the *Bpe*Globin dimers were compared to each other as well as other GCS hemes (Fig. 2). 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.[18, 34-35] All crystallized heme groups of *Bpe*Globin have similar saddling (B_{2u}) and propellering (A_{1u}) deformations (Fig. 2A, 2B; Table 3). However, the direction and magnitude of the doming (A_{2u}) of the Fe(II)-O_2 hemes are more similar to each other than to the

Table 3. Heme distortion modes in *Bpe*Globin in Å. 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.

Protein	**Doop (Over All)**	**B2u (Saddle)**	**B1u (Ruffle)**	**A2u (Dome)**	**Eg(x) (Wave)**	**Eg(y) (Wave)**	**A1u (Propeller)**
Homo Dimer Fe(II)-O2	1.21	-0.22	1.13	-0.27	-0.24	-0.05	0.01
Mixed Dimer Fe(II)-O2	1.11	-0.13	-1.03	-0.32	0.09	-0.20	<0.01
Mixed Dimer Fe(III)-H_2O	0.90	-0.18	0.85	0.15	-0.08	0.14	<0.01

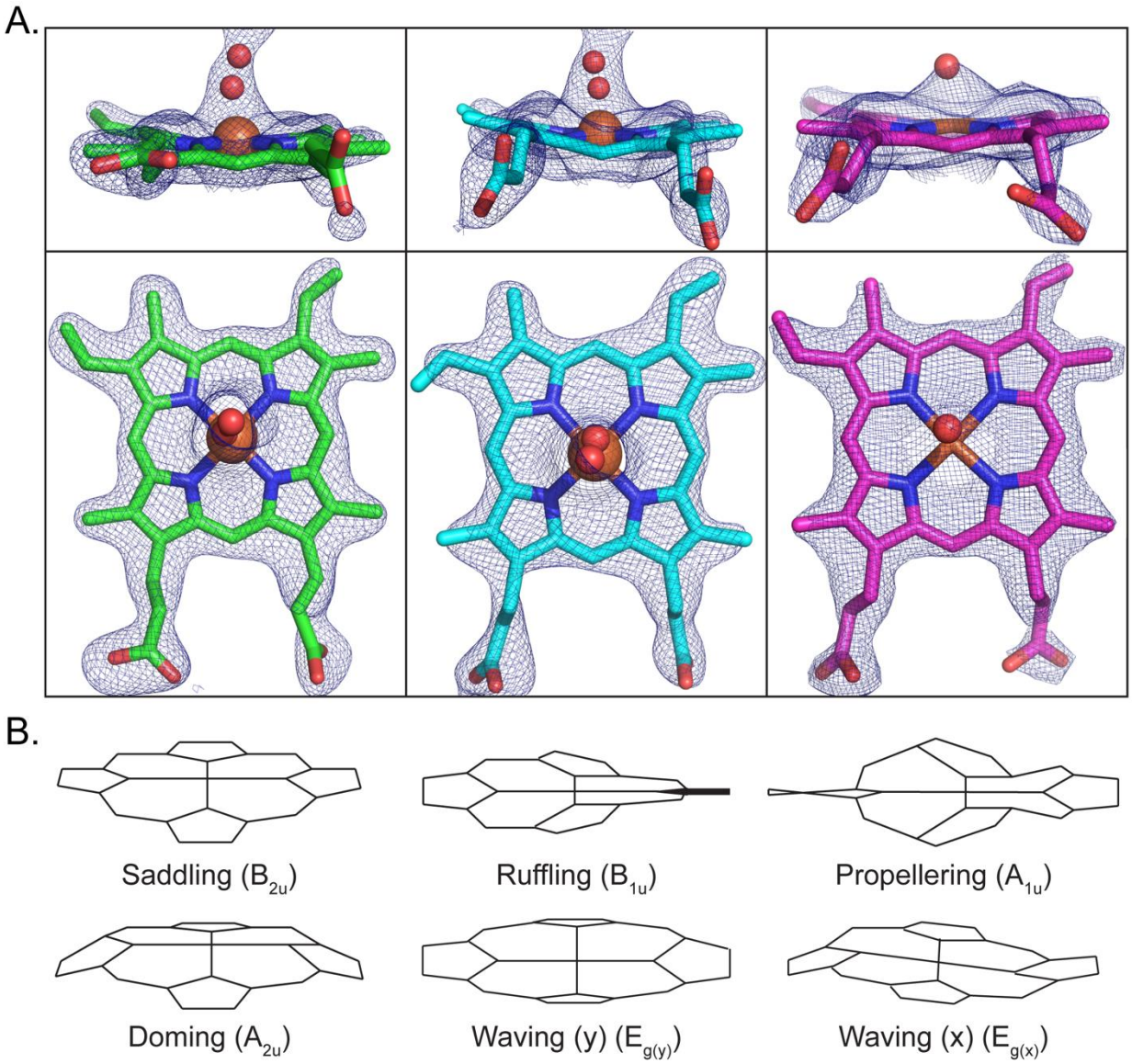

Figure 2. Heme deformations in *Bpe*Globin. (A) O_2-bound heme groups from both dimers exhibit significant degree of ruffling, while the Fe(III)-H_2O heme exhibits doming in the opposite direction of the O_2-bound hemes. The combination of all six types of deformation yield the overall heme distortion shown. (From left to right) Fe(II)- O_2 heme from the ferrous-oxy dimer; Fe(II)-O_2 heme from the mixed dimer; Fe(III)-H_2O 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 O_2-bound units. The $2F_o$-F_c electron density is contoured at 1.5 I/σ. (B) Heme cofactors can take on a number of deformation modes, including saddling (B_{2u}), ruffling (B_{1u}), propellering (A_{1u}), doming (A_{2u}), or waving ($E_{g(x)}$, $E_{g(y)}$) that may influence ligand-dependent signaling (adapted from Shelnutt, et al.[18]).

Fe(III)-H_2O heme. Additionally, the change in ruffling value (B_{1u}) of the O_2-bound hemes from the ferrous oxy dimer and the mixed dimer suggests that heme deformation for *Bpe*GReg is ligand dependent and that heme distortion modes may play a role in GCS signaling. Since DGC activity is dimerization dependent and the ligation state of the heme within GCS proteins affects oligomeric state, it is hypothesized that the ferrous-oxy dimer represents the active, O_2-bound state, while the mixed dimer represents a transition between active and inactive states (Fe(III)-H_2O is a low activity state of the enzyme). Specifically, the difference in heme conformation between Fe(II)-O_2 and Fe(III)-H_2O 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.

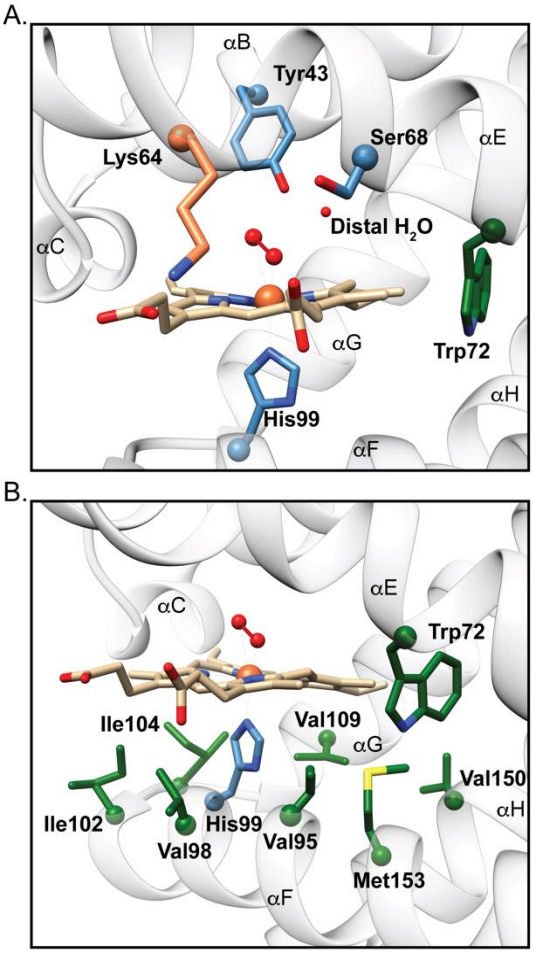

Figure 3. Heme pocket residues in *Bpe*Globin. (A) The distal hydrogen bonding residues Ser68 and Tyr43 are shown, along with the proximal His99 and the heme pocket residues Lys64 and Trp72. (B) Hydrophobic residues on the heme proximal side are shown. These residues are thought to propagate ligand-mediated heme distortions throughout the protein.

Heme distortion within *Bpe*Globin is likely due to interactions with heme pocket residues; in the proximal pocket, most residues located within 5 Å are hydrophobic, with the majority being isoleucines and valines (Fig. 3). These residues are well-positioned on the F helix to influence the globin

dimer interface, changes in which may represent a potential signaling mechanism. [2, 14, 29] 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 *Af*GcHK via HDX-MS.[29] In addition to these hydrophobic proximal residues, Trp72 of *Bpe*Globin sits at the heme edge on the E helix within appropriate distance for van der Waals contact with the heme. This Trp is conserved in *Ec*DosC and *Af*GcHK, while HemAT-*Bs* contains a histidine residue in the same position.

The location of Trp72 on the E helix of *Bpe*GReg suggests another potential route for the changes in heme distortion to be propagated through the protein by steric interaction. As Ser68 and Lys64 also are located on the E helix (Fig 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 O_2 binding[3] while Lys64 forms a hydrogen bond with heme propionate 6 in only the Fe(II)-O_2 monomers, suggesting a currently unexplored role for these residues in DGC activation. Recent crosslinking data also has demonstrated that the globin E helix of *Pcc*GCS 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 *Pcc*GCS and *Bpe*GReg (36%), it is inferred that the E helix of *Bpe*GReg 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.

2.4 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.[18, 34, 36] Accordingly, the distortions measured in *Bpe*Globin and *Ec*DosC are more similar than to HemAT-*Bs*. Specifically, the Fe(III)-H_2O heme of *Bpe*Globin has highly similar D_{oop} to all units of *Ec*DosC Fe(III)-H_2O (~0.87 and ~0.96, respectively). Additionally, heme groups from *Bpe*Globin and *Ec*DosC both show high D_{oop} and B_{1u} (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 D_{oop} and B_{1u}, respectively) indicating less overall distortion and ruffling. It should be noted that the Fe(III)-H_2O form of *Bpe*Globin was the used in this comparison with other sensor globins in the Fe(III)-H_2O 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[18, 37] and that conserved heme pocket residues assist in signal transmission within DGC-containing GCS proteins,[3, 7] potentially by modulating heme ruffling (B_{1u}) to transmit the ligation status of the heme-Fe into the globin fold. Indeed, comparison of the pockets of *Pcc*GCS, *Ec*DosC, and HemAT-*Bs* indicate that HemAT-*Bs* has the lowest sequence identity for these residues (12.5%, as compared to 75% for *Pcc*GCS and *Ec*DosC). In contrast, *Af*GcHK exhibits 16.5% sequence identity to *Bpe*GReg and has a distinct output domain (histidine kinase). While some of the *Af*GcHK chains exhibit similar distortion trends to *Bpe*Globin and other Fe(III) globins, a universal trend within the *Af*GcHK 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.

While previous studies on GCS proteins have not examined the role 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.[5, 38-40] 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 *Bpe*GReg and *Pcc*GCS, which have identified the Fe(II)-O2 state as having the highest enzyme activity. EcDosP, a PAS domain heme sensor, also exhibits high activity in the Fe(II)-O2, Fe(II)-CO, and Fe(II)-NO states, but low activity in the high spin Fe(II) unligated state.[38-39] Previous studies on *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 upwards out of the heme plane by 0.4-0.5 Å, while ligation pushes the Fe back down towards the plane.[40-42] However, as EcDosC, the GCS from *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.[5] 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.

2. 5 Interactions between Distal Residues and Bound Ligand

In order to verify putative interactions between bound ligands and protein residues identified from the *Bpe*Globin crystal structure, FTIR was performed on CO-bound wild-type and *Bpe*GReg/*Bpe*Globin protein variants, as well as the closely related GCS from *P. carotovorum*,

*Pcc*GCS/*Pcc*Globin (Fig. 4).[1] *Pcc*GCS is often examined in tandem with *Bpe*GReg 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.[1-3]

Furthermore, as *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.[14] Although the sensor globins from HemAT-*Bs* and *Ec*Dos*C* have previously been crystallized, they contain lower sequence identity to *Bpe*GReg as compared to *Pcc*GCS (14%, 34% and 36%, respectively). *Pcc*GCS exhibits the FTIR spectrum expected for a globin, with peaks at 1922 cm^{-1} and 1961 cm^{-1} corresponding to CO stretches with a hydrogen bonding interaction and apolar pocket, respectively (Fig. 4B).[43-44] Surprisingly, *Bpe*GReg exhibits three CO stretching frequencies, 1925 cm^{-1}, 1964 cm^{-1}, and 1972 cm^{-1} (Fig. 4A), which likely correspond to a hydrogen bond, apolar interaction, and electrostatic interaction with a negative charge, respectively.[44] FTIR scanning with isotopically labeled ^{13}CO resulted in a downfield shift of all three peaks in the spectrum of *Bpe*GReg (1925 cm^{-1}, 1964 cm^{-1}, and 1972 cm^{-1}) and two peaks in the spectrum of *Pcc*Globin (1922 cm^{-1} and 1961 cm^{-1}) (Fig. 4A, B), verifying that these stretching frequencies correspond to the CO stretch. In addition, FTIR spectra of full-length *Bpe*GReg and *Pcc*GCS were consistent with the corresponding globin spectra (Fig. 4C, D), demonstrating that the deletion of the middle and DGC domains does not substantially affect interactions between bound CO and the proteins.

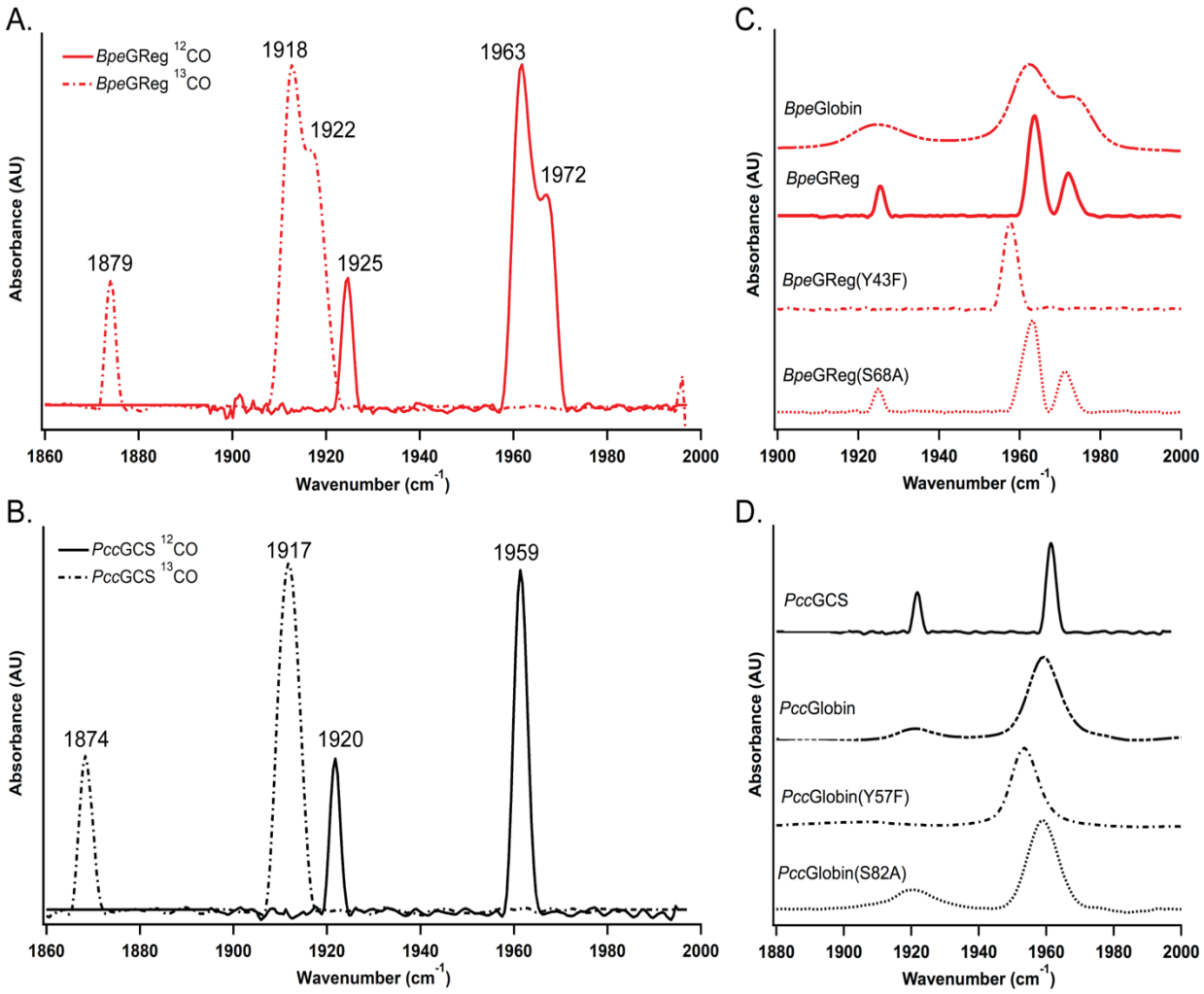

Figure 4: CO stretches for the various GCS proteins. (A) *Bpe*GReg Fe(II)-^{12}CO (solid red) vs Fe(II)-^{13}CO (dashed red). (B) *Pcc*GCS Fe(II)-^{12}CO (solid black) vs Fe(II)-^{13}CO (dashed black). (C) Fe(II)-^{12}CO peaks of *Bpe*Globin, *Bpe*GReg, *Bpe*GReg(Y43F) and *Bpe*GReg(S68A) (D) Fe(II)-^{12}CO peaks of *Pcc*Globin, *Pcc*Globin (Y57F) and *Pcc*Globin (S82A).

The observation of a 1972/1973 cm^{-1} stretching frequency for *Bpe*GReg and *Bpe*Globin, respectively, was intriguing given the lack of charged amino acids within the distal pocket (Fig. 1). Furthermore the observation of the 1972/1973 cm^{-1} stretch only in *Bpe*GReg/Globin suggests subtle differences in domain structure and/or conformation, as compared to *Pcc*GCS, supporting previously identified differences in ligand binding kinetics.[1] 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 *Bpe*GReg/*Bpe*Globin; Tyr57Phe for *Pcc*GCS; Tyr57Phe and Ser82Ala for *Pcc*Globin). For both proteins, only mutation of the distal Tyr produced an appreciable change in the FTIR spectrum; spectra of both *Bpe*GReg/Globin (Tyr43Phe) and *Pcc*GCS/Globin (Tyr57Phe) lack the stretches at 1925 cm^{-1} and 1922/1921 cm^{-1}, respectively, suggesting that the distal tyrosine provides a hydrogen bond to bound CO (Fig. 4D). Furthermore, the apolar stretch at 1964/1962 cm^{-1} shifted for both *Bpe*GReg/Globin (Tyr43Phe) and *Pcc*GCS/Globin (Tyr57Phe), indicating a change in the heme pocket environment upon mutation of Tyr43 (Fig. 4D). This shift may be due to subtle rearrangements that allow the partial positive charge in the center of the phenyl π-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,[3] suggests that the distal serine plays a role in O_2 binding, likely through a hydrogen bond network within the heme pocket.

2. 6 Examination of Structural Water

The observation of a *Bpe*GReg/Globin CO stretching frequency at 1972 cm^{-1} suggested that the bound water may be involved in modulating the distal pocket environment. While the crystal structure identified a hydrogen bond between O_2 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 *Bpe*Globin crystal structure, FTIR was performed on wild type *Bpe*Globin at pH values ranging from 5.0 to 8.0 (Fig. 5). As the buffer pH was decreased, the peak at 1975 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 D_2O and $H_2{}^{18}O$ labeled water (Fig. 6). Within the *Bpe*Globin FTIR spectrum, the 1975 cm^{-1} peak shifts and loses intensity when in the $H_2{}^{18}O$ buffer (Fig. 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 H_2O, spatially oriented through interactions with the distal Tyr43.

While *Pcc*GCS/*PccGlobin* does not exhibit a peak in the 1970 cm^{-1} region like *Bpe*GReg/*Bpe*Globin, 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 *Bpe*GReg and *Pcc*GCS suggests that *Pcc*GCS/*Pcc*Globin 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 *Pcc*Globin, and thus *Pcc*GCS, FTIR was performed on wild type *Pcc*Globin over a pH range of 7.0 to 9.0 (Fig. 5B). Like *Bpe*Globin, *Pcc*Globin remained stable at all tested pH values, although it was noted that stability significantly decreased above pH 9. The spectrum of *Pcc*Globin 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.

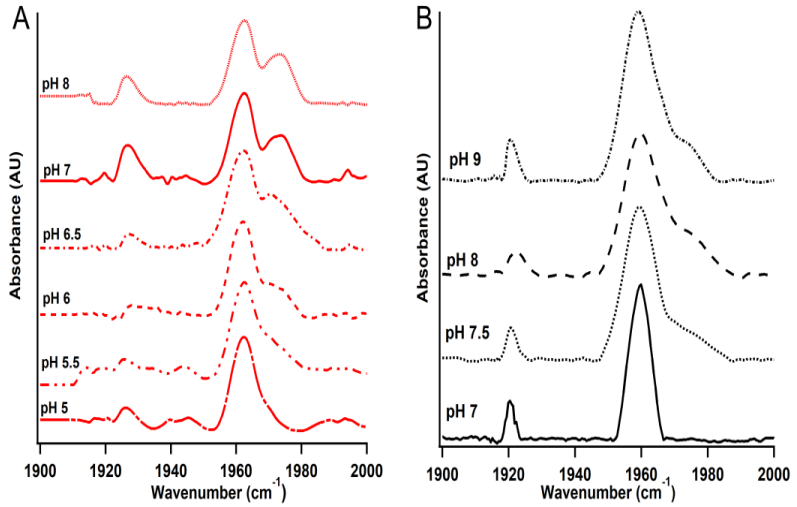

Figure 5. Effect of pH on CO stretching frequencies. (A) *Bpe*Globin Fe(II)-CO ranging from pH 5.0 -8.0. (B) *Pcc*Globin Fe(II)-CO ranging from pH 7.0 to 9.0.

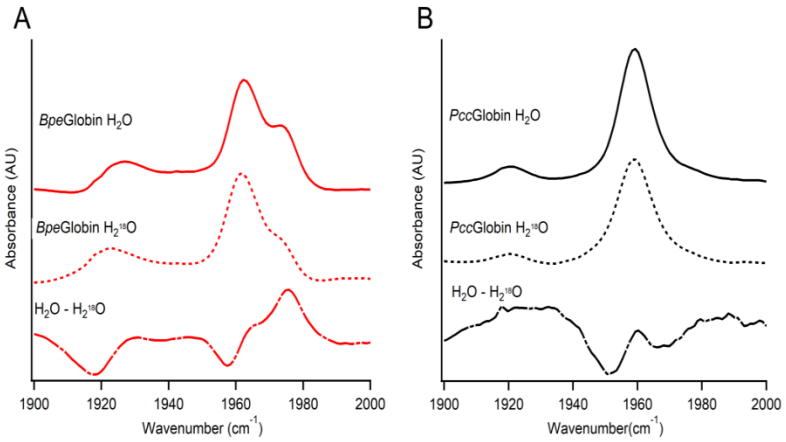

Figure 6. FTIR examination of the structured water with difference spectra shown below. (A) FTIR spectra of *Bpe*Globin Fe(II)-CO in H_2O (solid red) and $H_2{}^{18}O$ (dashed red) with the difference spectra shown below. (B) FTIR spectra of *Pcc*Globin Fe(II)-CO in H_2O (solid black) and $H_2{}^{18}O$ (dashed black) with the difference spectra shown below.

FTIR spectra of PccGlobin were also taken in D_2O and $H_2{}^{18}O$ labeled water to support the distal water moiety (Fig. 5B, 6B). Unlike in the $H_2{}^{18}O$ spectrum of *Bpe*Globin, that of *Pcc*Globin does not demonstrate a significant shift in any of its peaks due to the lack of structural water-dependent stretches in the *Pcc*Globin Fe(II)-CO spectrum (Fig. 6). This difference in the protonation state of the ordered water between *Bpe*GReg and *Pcc*GCS, which must be due to differences in heme pocket environments, likely contributes to the observed differences in O_2 binding kinetics (mono/bi-exponential dissociation rates) and O_2-dependent cyclase activity of these GCS proteins.[1]

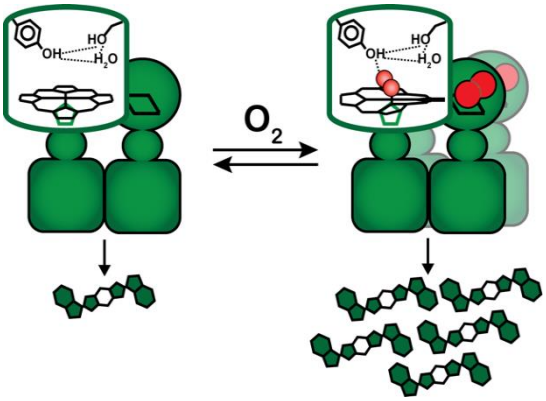

Figure 7. Overview of proposed heme pocket interactions involved in O_2-dependent signal transduction. The tetramer, which has high diguanylate cyclase activity (c-di-GMP depicted in green), exists in the Fe(II)-O_2 ligation state (right) with a deformed heme and a proposed hydrogen bond network based on current and previous work.[1-6] Dissociation of O_2 and/or heme autoxidation results in a less distorted heme and formation of a dimer with low diguanylate cyclase activity (left).

3. Conclusion

In conclusion, *Bpe*Globin is the first isolated GCS sensor globin domain to be crystallized in the native active Fe(II)-O_2 ligation state. Differences in ligation state correlate with heme distortion, suggesting a role in transmission of the O_2 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 *Bpe*GReg, a 1975 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 *Pcc*GCS distal pocket remains protonated. In both proteins, the distal pocket Tyr is required for the 1975 and 1920 cm^{-1} stretching frequencies, 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 O_2 binding signal in GCS proteins (Fig 7).

4. Chapter References

1. Burns, J. L.; Deer, D. D.; Weinert, E. E., Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. *Mol. Biosyst.* **2014,** *10* (11), 2823-6.

2. Burns, J. L.; Rivera, S.; Deer, D. D.; Joynt, S. C.; Dvorak, D.; Weinert, E. E., Oxygen and c-di-GMP Binding Control Oligomerization State Equilibria of Diguanylate Cyclase-Containing Globin Coupled Sensors. *Biochemistry* **2016,** *55*, 6642-6651.

3. Rivera, S.; Burns, J. L.; Vansuch, G. E.; Chica, B.; Weinert, E. E., Globin domain interactions control heme pocket conformation and oligomerization of globin coupled sensors. *J. Inorg. Biochem.* **2016,** *164*, 70-76.

4. Aono, S.; Kato, T.; Matsuki, M.; Nakajima, H.; Ohta, T.; Uchida, T.; Kitagawa, T., Resonance Raman and ligand binding studies of the oxygen-sensing signal transducer protein HemAT from *Bacillus subtilis*. *J. Biol. Chem.* **2002,** *277* (16), 13528-38.

5. Kitanishi, K.; Kobayashi, K.; Kawamura, Y.; Ishigami, I.; Ogura, T.; Nakajima, K.; Igarashi, J.; Tanaka, A.; Shimizu, T., Important roles of Tyr43 at the putative heme distal side in the oxygen recognition and stability of the Fe(II)-O2 complex of YddV, a globin-coupled heme-based oxygen sensor diguanylate cyclase. *Biochemistry* **2010,** *49* (49), 10381-93.

6. Tarnawski, M.; Barends, T. R.; Schlichting, I., Structural analysis of an oxygen-regulated diguanylate cyclase. *Acta Crystallogr D Biol Crystallogr* **2015,** *71* (Pt 11), 2158-77.

7. Walker, J. A.; Rivera, S.; Weinert, E. E., Mechanism and Role of Globin Coupled SensorSignaling. *Adv Microb Physiol* **2017,** *71*, 133-169.

8. Shimizu, T.; Huang, D.; Yan, F.; Stranava, M.; Bartosova, M.; Fojtikova, V.; Martinkova, M., Gaseous O2, NO, and CO in signal transduction: structure and function

relationships of heme-based gas sensors and heme-redox sensors. *Chem. Rev.* **2015,** *115* (13), 6491-533.

9. Green, J.; Rolfe, M. D.; Smith, L. J., Transcriptional regulation of bacterial virulence gene expression by molecular oxygen and nitric oxide. *Virulence* **2014,** *5* (8), 794-809.

10. Wu, C.; Cheng, Y. Y.; Yin, H.; Song, X. N.; Li, W. W.; Zhou, X. X.; Zhao, L. P.; Tian, L. J.; Han, J. C.; Yu, H. Q., Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. *Sci. Rep.* **2013,** *3*, 1945-1951.

11. Sawai, H.; Yoshioka, S.; Uchida, T.; Hyodo, M.; Hayakawa, Y.; Ishimori, K.; Aono, S., Molecular oxygen regulates the enzymatic activity of a heme-containing diguanylate cyclase (HemDGC) for the synthesis of cyclic di-GMP. *Biochim. Biophys. Acta* **2010,** *1804* (1), 166-72.

12. Romling, U.; Galperin, M. Y.; Gomelsky, M., Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* **2013,** *77* (1), 1-52.

13. Chan, C.; Paul, R.; Samoray, D.; Amiot, N. C.; Giese, B.; Jenal, U.; Schirmer, T., Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* **2004,** *101* (49), 17084-9.

14. Burns, J. L.; Jariwala, P. B.; Briggs, L.; Fontaine, B. M.; Weinert, E. E., Oxygen-Dependent Globin Coupled Sensor Signaling Modulates Motility and Virulence of the Plant Pathogen *Pectobacterium carotovorum*. *ACS Chem. Biol.* **2017,** *12*, 2070-2077.

15. Hengge, R., Principles of c-di-GMP signalling in bacteria. *Nature Rev. Microbiol.* **2009,** *7* (4), 263-73.

16. Wan, X.; Tuckerman, J. R.; Saito, J. A.; Freitas, T. A.; Newhouse, J. S.; Denery, J. R.; Galperin, M. Y.; Gonzalez, G.; Gilles-Gonzalez, M. A.; Alam, M., Globins synthesize the second

messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* **2009,** *388* (2), 262-70.

17. Kalia, D.; Merey, G.; Nakayama, S.; Zheng, Y.; Zhou, J.; Luo, Y.; Guo, M.; Roembke, B. T.; Sintim, H. O., Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chem. Soc. Rev.* **2012,** *42* (1), 305-41.

18. Shelnutt, J. A.; Song, X. Z.; Ma, J. G.; Jia, S. L.; Jentzen, W.; Medforth, C. J., Nonplanar porphyrins and their significance in proteins. *Chem. Soc. Rev.* **1998,** *27* (1), 31-41.

19. Vinogradov, S. N.; Tinajero-Trejo, M.; Poole, R. K.; Hoogewijs, D., Bacterial and archaeal globins - A revised perspective. *Biochim. Biophys. Acta* **2013,** *1834*, 1789-1800.

20. Borriello, G.; Werner, E.; Roe, F.; Kim, A. M.; Ehrlich, G. D.; Stewart, P. S., Oxygen limitation contributes to antibiotic tolerance of *Pseudomonas aeruginosa* in biofilms. *Antimicrob. Agents Chemother.* **2004,** *48* (7), 2659-64.

21. Phippen, B. L.; Oliver, J. D., Role of anaerobiosis in capsule production and biofilm formation in Vibrio vulnificus. *Infect. Immun.* **2015,** *83* (2), 551-9.

22. Roberts, R. R.; Hota, B.; Ahmad, I.; Scott, R. D., 2nd; Foster, S. D.; Abbasi, F.; Schabowski, S.; Kampe, L. M.; Ciavarella, G. G.; Supino, M.; Naples, J.; Cordell, R.; Levy, S. B.; Weinstein, R. A., Hospital and societal costs of antimicrobial-resistant infections in a Chicago teaching hospital: implications for antibiotic stewardship. *Clin. Infect. Dis.* **2009,** *49* (8), 1175-84.

23. Sommer, F.; Backhed, F., The gut microbiota-masters of host development and physiology. *Nat Rev. Microbiol.* **2013,** *11* (4), 227-38.

24. Zhang, W.; Olson, J. S.; Phillips, G. N., Jr., Biophysical and kinetic characterization of HemAT, an aerotaxis receptor from *Bacillus subtilis*. *Biophys. J.* **2005,** *88* (4), 2801-14.

25. Zhang, W.; Phillips, G. N., Jr., Structure of the oxygen sensor in *Bacillus subtilis*: signal transduction of chemotaxis by control of symmetry. *Structure* **2003,** *11* (9), 1097-110.

26. Fojtikova, V.; Stranava, M.; Vos, M. H.; Liebl, U.; Hranicek, J.; Kitanishi, K.; Shimizu, T.; Martinkova, M., Kinetic Analysis of a Globin-Coupled Histidine Kinase, Af GcHK: Effects of the Heme Iron Complex, Response Regulator, and Metal Cations on Autophosphorylation Activity. *Biochemistry* **2015,** *54* (32), 5017-5029.

27. Tamayo, R.; Pratt, J. T.; Camilli, A., Roles of Cyclic Diguanylate in the Regulation of Bacterial Pathogenesis. *Ann. Rev. Microbiol.* **2007,** *61*, 131-148.

28. Paul, R.; Abel, S.; Wassmann, P.; Beck, A.; Heerklotz, H.; Jenal, U., Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J. Biol. Chem.* **2007,** *282* (40), 29170-7.

29. Stranava, M.; Man, P.; Skalova, T.; Kolenko, P.; Blaha, J.; Fojtikova, V.; Martinek, V.; Dohnalek, J.; Lengalova, A.; Rosulek, M., Coordination and redox state-dependent structural changes of the heme-based oxygen sensor *Af*GcHK associated with intraprotein signal transduction. *J. Biol. Chem.* **2017**, jbc. M117. 817023.

30. Nakajima, K.; Kitanishi, K.; Kobayashi, K.; Kobayashi, N.; Igarashi, J.; Shimizu, T., Leu65 in the heme distal side is critical for the stability of the Fe(II)-O2 complex of YddV, a globin-coupled oxygen sensor diguanylate cyclase. *J. Inorg. Biochem.* **2012,** *108*, 163-70.

31. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007,** *372* (3), 774-797.

32. Zhang, W.; Phillips, G. N., Jr., Crystallization and X-ray diffraction analysis of the sensor domain of the HemAT aerotactic receptor. *Acta Crystallogr D Biol Crystallogr* **2003,** *59* (Pt 4), 749-51.

33. Ohta, T.; Yoshimura, H.; Yoshioka, S.; Aono, S.; Kitagawa, T., Oxygen-sensing mechanism of HemAT from *Bacillus subtilis:* a resonance Raman spectroscopic study. *J. Am. Chem. Soc.* **2004,** *126* (46), 15000-1.

34. Ma, J. G.; Zhang, J.; Franco, R.; Jia, S. L.; Moura, I.; Moura, J. J.; Kroneck, P. M.; Shelnutt, J. A., The structural origin of nonplanar heme distortions in tetraheme ferricytochromes c3. *Biochemistry* **1998,** *37* (36), 12431-42.

35. Graves, A. B.; Graves, M. T.; Liptak, M. D., Measurement of Heme Ruffling Changes in MhuD Using UV–vis Spectroscopy. *J. Phys. Chem. B* **2016,** *120* (16), 3844-3853.

36. Ma, J. G.; Laberge, M.; Song, X. Z.; Jentzen, W.; Jia, S. L.; Zhang, J.; Vanderkooi, J. M.; Shelnutt, J. A., Protein-induced changes in nonplanarity of the porphyrin in nickel cytochrome c probed by resonance Raman spectroscopy. *Biochemistry* **1998,** *37* (15), 5118-28.

37. Jentzen, W.; Ma, J. G.; Shelnutt, J. A., Conservation of the conformation of the porphyrin macrocycle in hemoproteins. *Biophys. J.* **1998,** *74* (2), 753-763.

38. Tanaka, A.; Takahashi, H.; Shimizu, T., Critical role of the heme axial ligand, Met95, in locking catalysis of the phosphodiesterase from Escherichia coli (Ec DOS) toward cyclic diGMP. *J. Biol. Chem.* **2007,** *282* (29), 21301-21307.

39. 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 di-GMP: Roles of Met95, Arg97, and Phe113 of the putative heme distal side in catalytic regulation and ligand binding. *Biochemistry* **2008,** *47* (50), 13438-13446.

40. Kitanishi, K.; Kobayashi, K.; Uchida, T.; Ishimori, K.; Igarashi, J.; Shimizu, T., Identification and functional and spectral characterization of a globin-coupled histidine kinase from Anaeromyxobacter sp. Fw109-5. *J. Biol. Chem.* **2011,** *286* (41), 35522-34.

41. Miyatake, H.; Mukai, M.; Adachi, S.-i.; Nakamura, H.; Tamura, K.; Iizuka, T.; Shiro, Y.; Strange, R. W.; Hasnain, S. S., Iron coordination structures of oxygen sensor FixL characterized by Fe K-edge extended X-ray absorption fine structure and resonance Raman spectroscopy. *J. Biol. Chem.* **1999,** *274* (33), 23176-23184.

42. Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N., Mechanisms of Ligand Recognition in Myoglobin. *Chem. Rev.* **1994,** *94* (3), 699-714.

43. Pavlou, A.; Martinkova, M.; Shimizu, T.; Kitanishi, K.; Stranava, M.; Loullis, A.; Pinakoulaki, E., Probing the ligand recognition and discrimination environment of the globin-coupled oxygen sensor protein YddV by FTIR and time-resolved step-scan FTIR spectroscopy. *Phys. Chem. Chem. Phys.* **2015,** *17* (26), 17007-15.

44. Li, X.-Y., Ibers, J.A., Sessler, J.L., Spiro, T.G., How far can proteins bend the FeCO unit? Distal, polar and steric effects in heme proteins and models. *J. Am. Chem. Soc.* **1994,** *116*, 162-176.

Chapter 4: Determining Heme Residue Effects on Kinetic Turnover in GCSs

Chapter 4

Determining Heme Residue Effects on Kinetic Turnover in GCSs

Chapter Overview:

1. Introduction

Molecular oxygen (O_2) concentration is important for bacterial survival due to the hazard of reactive oxygen species-mediated host immune responses and its importance in metabolism.[1-4] Because of this, bacteria have evolved a family of O_2-sensing hemoproteins termed globin coupled sensors (GCS),[5] which are implicated in the regulation of a variety of O_2-mediated tasks,[7] such as motility,[8-10] virulence,[8-11] and biofilm formation.[3, 8, 9, 12] GCS proteins consist of a heme-containing globin domain at the N-terminal connected by a highly variable middle domain to output domain(s), which include methyl-accepting chemotaxis proteins, phosphodiesterases, sulfate transport and anti-anti-σ factor (STAS) domains, kinases, and diguanylate cyclases (DGCs).[1, 2, 11, 13-18] The diguanylate cyclase-containing globin coupled sensors (DGC-GCS) have

recently been the subject of increased study, as diguanylate cyclases are the enzymes responsible for the synthesis of the small molecule c-di-GMP, the primary second messenger involved in biofilm formation.[8, 10] However, despite important roles for these proteins in modulating O_2-mediated biofilm formation and virulence,[3, 9, 11, 19] there is currently no comprehensive model detailing the mechanism by which the O_2 binding signal is communicated from the globin domain to the DGC output domain.

The greatest impediment to a detailed mechanism is the lack of a full-length structural information for any GCS protein, leading to uncertainty regarding residues involved in signaling, as well as the three-dimensional shape of these proteins. Consequently, much of the previous work on the transduction mechanism has been developed from the crystallized globin domains of a few GCS proteins, including HemAT-*Bs*,[20, 21], *Ec*DosC,[22]*Af*GcHK,[23] and *Bpe*GReg.[24] These structures and accompanying biochemical studies have suggested three heme pocket residues as the major contributors to the signaling mechanism: the primary hydrogen bond-donating distal Tyr, a secondary distal H-bond donor (Ser or Thr), and the proximal His that ligates the heme iron center (Fig 1).[15, 23-26]

Previous work has demonstrated that both the distal Tyr and Ser are important for O_2 binding and the distal Tyr can act as a heme pocket hydrophobic gate in response to ligands.[25, 26] Ligand-specific conformational changes of these heme pocket residues have also been cited as evidence of a H-bond-mediated signaling mechanism.[26] Furthermore, the placement of the proximal His on the globin F8 helix has implicated ligand-mediated changes in the His-Fe bond length in

modulating oligomerization state, which have been shown to have a strong effect on O_2 affinity and DGC activity.[2, 23, 25] Given the importance of these three residues in stabilizing ligand binding, it has been long expected that the residues would be involved in the intra-protein signaling mechanism of GCSs.

However, a recent crystal structure of the globin domain from the DGC-GCS *Bpe*GReg (*Bpe*Globin) highlighted a number of additional structural features potentially involved in the signaling mechanism of these proteins.[24] In particularly, comparison of *Bpe*Globin structures in different ligation states suggested a mechanism whereby the ligand binding signal is propagated through interactions between a distorted heme cofactor and heme edge residues, including a Trp

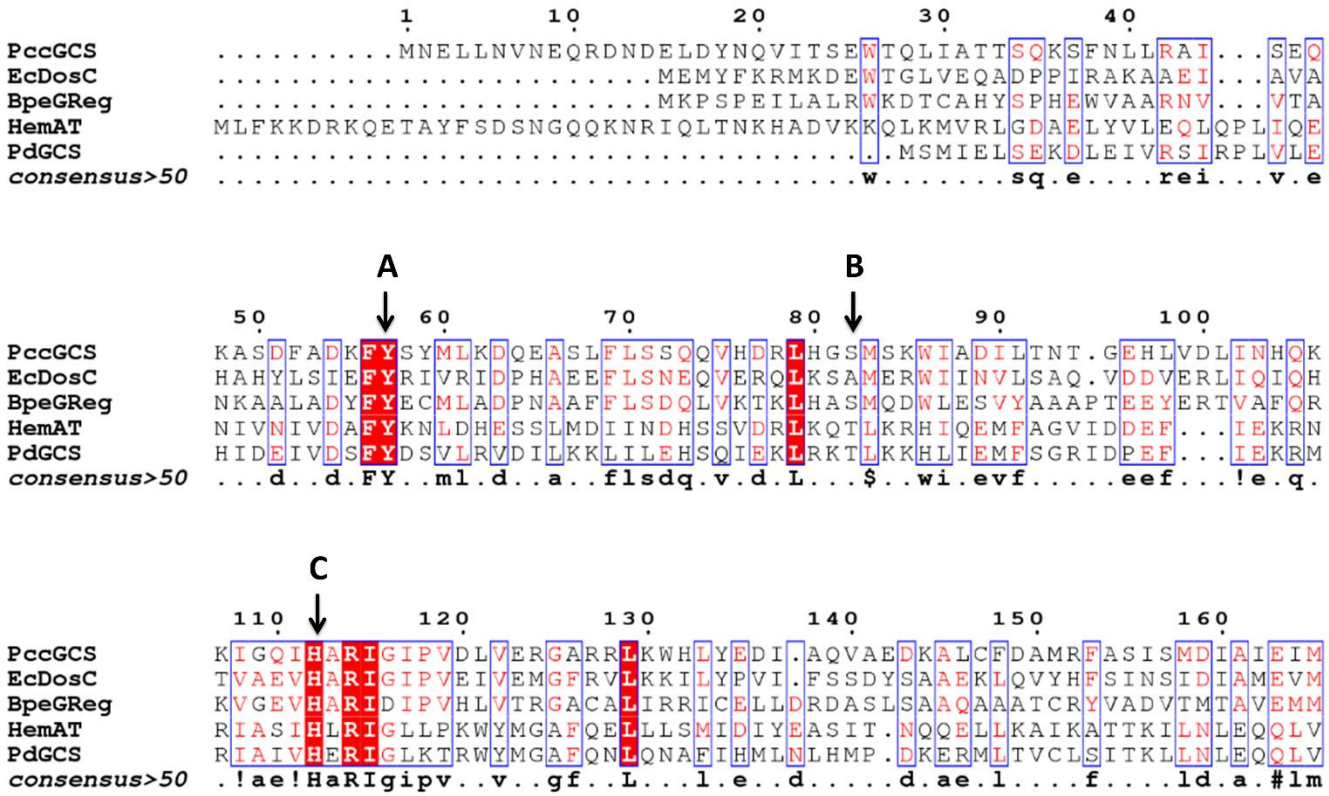

Figure 1. Sequence Alignment of *Pcc*GCS, *Ec*DosC, *Bpe*GReg, HemAT-*Bs* and *Pd*GCS; representing a variety of GCSs from different bacteria. Highlighted are (A) distal Tyr, (B) distal Ser/Thr and (C) proximal His.

on the E helix, which seems to be engaged in a T-shaped π-stack with the heme cofactor, and Lys also on the E helix, which forms a H-bond with the heme propionate 6 only in the O_2-bound form of the protein (Fig. 2).[24] As a follow up to these findings, heme pocket features were probed in the DGC-GCS from the bacterial soft rot pathogen *Pectobacterium carotovorum* (*Pcc*GCS), which is closely related to *Bpe*GReg. Using a combination of site directed mutagenesis, ligand binding studies, enzyme kinetics, and heme spectroscopy, distal pocket hydrogen bond donating residues have been demonstrated to have very little effect on DGC activation. In contrast, heme edge residues have been identified as key determinants of the signaling mechanism. In addition, the heme edge residues play key roles in controlling redox potential and ligand affinity of GCS proteins, providing new insights into heme protein ligand selectivity and activation.

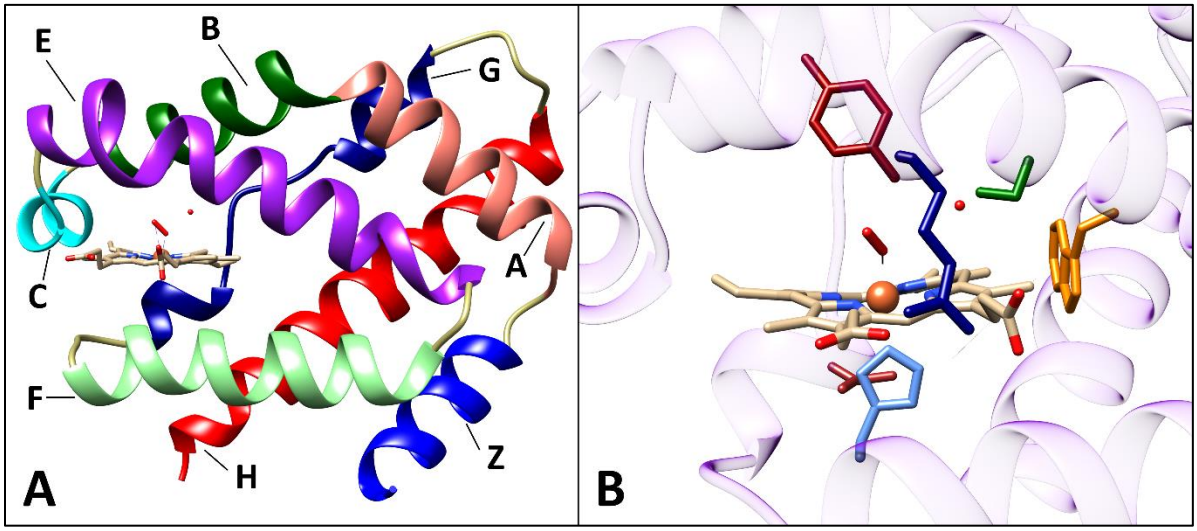

Figure 2. Homology model of *Pcc*Globin. (A) Globin domain with labeled helices (B) Heme pocket with highlighted residues; Tyr57 (Red), Ser82 (Green), Trp86 (yellow), Arg78 (dark blue), His112 (light blue), and Val122 (orange). Homology model is based on BpeGlobin (PDB:6M9A), generated by Phyer2.[6]

2. Experimental Results and Discussion

2.1 Effects of Heme Pocket Mutations on Ligand Binding

Previous work has established the close relationship between O_2 dissociation rate, oligomeric state, and catalytic output in GCS proteins.[27] Thus, obtaining the O_2 dissociation rates of *Pcc*GCS heme pocket residue variants is important for relating the roles of the putatively important residues in the signaling mechanism to the roles they play in O_2 binding (Table 1). It is clear that Tyr57 and His112 are both required for O_2 binding, as the Y57F and H112G mutants were both unable to be assayed due to rapid autoxidation under aerobic conditions.

In contrast to previous work on isolated globin domain variants, the full-length *Pcc*GCS S82A variant shows only a modest increase in both the slow and fast O_2 dissociation rates (k_1 and k_2, respectively) compared to wild type. Furthermore, while the *Pcc*Globin S82A variant exhibits a complete shift to the slow dissociation rate, the full-length variant shows a decrease in the percentage of the slow dissociation rate. This implies that the distal serine does not play the same role in determining the biphasic dissociation kinetics of the full-length protein that it does in the isolated globin. The discrepancy between the isolated globin domain S82A variant and the corresponding full-length may be explained by the three-dimensional structure of the full-length protein. Recent evidence has shown that the E-helix, on which the distal serine resides, is in close proximity to the DGC domain (Walker et al. *Manuscript in preparation*). As a result, the presence of this domain in the full-length may act as a physical barrier to O_2 dissociation that is otherwise absent in the isolated globin domain, thus compensating for the role of the Ser in O_2 binding.

Table 1. Oxygen Dissociation Rates

Protein	k_1	k_2	% k_1	% k_2
*Pcc*GCS	0.56	3.87	56	44
*Pcc*GCS(Y57F)	N.A.	N.A.	N.A	N.A.
*Pcc*GCS(R78G)	TBD	TBD	TBD	TBD
*Pcc*GCS(S82A)	1.2 ± 0.4	5.7 ± 0.7	39 ± 2	61 ± 2
*Pcc*GCS(W86A)	2.8 ± 0.3	13.1 ± 0.4	53 ± 4	47 ± 4
*Pcc*GCS(W86H)	1.7 ± 0.1	20 ± 1	42 ± 2	58 ± 2
*Pcc*GCS(H112G)	N.A.	N.A.	N.A	N.A.
*Pcc*GCS(V122Y)	0.68 ± 0.01	4.4 ± 0.03	65 ± 2	35 ± 2
*Pcc*GCS(W86H/V122Y)	3.0 ± 0.4	13.2 ± 0.2	57 ± 5	43 ± 5

Recently, residues lining the heme edge have arisen as potentially important players in the signaling mechanism of these proteins. Thus, we generated variants of *Pcc*GCS with mutations to Trp86, Val122, and Arg78, which form Van der Waals and H-bonding interactions with the heme cofactor, respectively. In contrast to the distal or proximal residues, assaying variants with mutations to the heme-edge Trp residue yielded a range of results, depending on the mutation. Both the W86H and W86A variants exhibited an increase in the slow dissociation rate, yielding a weaker O_2 affinity; however W86A dissociation rate and percentage attributed to the slow rate are higher than W86H. This indicates that the presence of an aromatic residue at this position is important O_2 stability and therefore O_2 affinity. Another heme edge residue mutant, *Pcc*GCS(V122Y), yielded very little change in the dissociation kinetics, but in the case of the V122Y/W86H double mutant, we observed the greatest increase in the slow rate of any of the variants and a large increase in the fast rate. Both the W86 and V122 locations do not directly interact with the bound ligand, O_2, but are within distance to interact with the heme. The significant changes to oxygen dissociation suggests altered porphyrin distortions, resulting in decreased stability of bound O_2.

2.2 Heme Pocket Residues Control DGC Activation

In order to test the importance of heme pocket residues in DGC activation, enzyme kinetic assays were performed on wild-type *Pcc*GCS and heme pocket variants in the unbound (Fe(II)), ferrous-oxy (Fe(II)-O_2), and ferrous nitrous (Fe(II)-NO) forms. Wild-type *Pcc*GCS was assayed alongside protein variants as a positive control, and all K_M and k_{cat} values were the same as previously reported (Table 2).[2, 25] Surprisingly, neither of the variants with mutations to the distal H-bonding residues (Y57F and S82A) had a strong effect on the DGC k_{cat}. Of particular note, the fold activation of the DGC enzyme for both variants was almost exactly the same as that of wild type, indicating that differences in activity may be due to structural perturbations rather than functional ones.

Table 2. Kinetic Rates for Various GCS

Protein	Fe(II)		Fe(II)-O_2		Fe(II)-NO	
	k_{cat} (min^{-1})	Km (μM)	k_{cat} (min^{-1})	Km (μM)	k_{cat} (min^{-1})	Km (μM)
*Pcc*GCS	0.29 ± 0.01	62 ± 3	0.73 ± 0.01	31 ± 6	0.51 ± 0.01	N.D.
*Pcc*GCS(Y57F)	0.25 ± 0.01	85 ± 16	N.A	N.A.	0.42 ± 0.01	61 ± 9
*Pcc*GCS(R78G)	0.48 ± 0.01	77 ± 12	TBD	TBD	0.33 ± 0.01	100 ± 9
*Pcc*GCS(S82A)	0.21 ± 0.01	145 ± 15	0.52 ± 0.01	73 ± 4	TBD	TBD
*Pcc*GCS(W86A)	0.64 ± 0.01	43 ± 4	TBD	TBD	0.60 ± 0.01	35 ± 2
*Pcc*GCS(W86H)	0.45 ± 0.01	193 ± 30	2.33 ± 0.05	23 ± 6	1.02 ± 0.3	52 ± 8
*Pcc*GCS(H112G)	N.A.	N.A.	N.A	N.A.	0.34 ± 0.05	127 ± 24
*Pcc*GCS(V122Y)	0.23 ± 0.01	69 ± 8	0.81 ± 0.2	121 ± 21	0.61 ±0.01	36 ± 12
*Pcc*GCS(W86H/V122Y)	0.31 ± 0.01	56 ± 8	1.00 ± 0.3	64 ± 16	0.93 ± 0.01	25 ± 5

As with the Y57F variant, the H112G variant is also unable to stably bind O_2, but additionally, it is challenging to isolate this variant to be in the Fe(II) unligated state, as imidazole is required to prevent heme dissociation. However, the Fe(II)-NO state can be assayed, yielding a k_{cat} of 0.34 min^{-1}, which is not significantly lower than the wild type k_{cat} of 0.51 min^{-1}, indicating that this residue also, likely, does not play a direct role in signaling. It should be noted, though, that previous work as well as the present one confirms that Tyr57 and His112 both participate in ligand and protein stability, evidenced by the rapid autoxidation and aggregation of both variants in the Fe-O_2 state.

For the heme edge residues, the W86H variant exhibited a large increase in DGC activity in all ligation states compared to wild type. It is possible that this mutation changes the electronics of the heme in such a way that it distorts into a more ‘on’ position, leading to the observed increase in activity. Furthermore, the fold-change of the W86H variant is twice that of wild type, indicating that this residue modulates the responsiveness of the DGC domain in addition to just its activity. The V122Y variant exhibits a more modest increase in the Fe-O_2 and Fe-NO k_{cat} values than the W86H variant. The double mutant W86H/V122Y resulted in faster rates compared to V122Y, but surprisingly, the Fe-O_2 and Fe-NO rates are nearly the same (1.00 and 0.93 min^{-1}, respectively), unlike any other variant or the wild type. This is intriguing, as it suggests ligand selectivity is controlled by heme deformation, and thus, Trp86 and Val122 are a key residue in the signaling pathway.

2.3 Heme Pocket Residues Influence Oligomeric States

Ligand-dependent oligomeric state changes are a well-documented phenomenon in these proteins, with previous work from our lab showing that O_2 binding drives the oligomeric equilibrium of these proteins towards the more active tetrameric state in both *Pcc*GCS and the related *Bpe*GReg[2, 25]. To determine the role of heme pocket residues in affecting this ligand-dependent oligomeric shift, heme pocket variants of *Pcc*GCS were subjected to analytical gel filtration (AGF) HPLC. Wild-type *Pcc*GCS exists primarily as a tetramer (~75%), with small populations of dimer (~15%), octamer (~6%), and high molecular weight (HMW, ~4%) species (Table 3). The Y57F and S82A variant showed virtually no change in oligomeric populations.

Table 3. Oligomeric Percentages

Protein	Oligomeric State (%)			
	HMW	Oct	Tetra	Dimer
*Pcc*GCS	4.1	5.8	75.1	15.0
*Pcc*GCS(Y57F)	8.1	4.2	74.5	13.2
*Pcc*GCS(R78G)	4.5	5.6	88.9	-
*Pcc*GCS(S82A)	3.6	9.6	76.3	10.4
*Pcc*GCS(W86A)	-	-	100	-
*Pcc*GCS(W86H)	-	-	88.1	11.8
*Pcc*GCS(H112G)	-	6.4	93.6	-
*Pcc*GCS(V122Y)	-	-	100	-
*Pcc*GCS(W86H/V122Y)	-	-	100	-

In contrast, the H112G variant demonstrates a marked shift in the oligomeric equilibrium, with ~94% of the protein existing in the tetrameric state and ~6% in the octamer state. This large shift in the oligomeric state is surprising, given the low kinetic activity of this variant and the association between the tetrameric state and high DGC activity. However, as the proximal His residue is located on the F helix, which is connected to the G and H helices that comprise the

globin dimer interface (Figure 2A), it is reasonable that this mutation would have a large effect on the oligomeric states this protein populates.

Additional the heme edge mutations all show significant increases to the tetramer oligomerization state, suggesting an increase in kinetic activity. In the case of W86H, V122Y and the double mutant, faster kinetic rates are indeed observed. Once again, support the idea that Trp86 and Val122 are involved in the signaling pathway and additionally, alter the oligomeric state to increase the high-activity tetramer state.

2.4 Heme Edge Residues Alter Midpoint Potential

Although electrochemical analysis has been performed on other hemoproteins, this form of analysis has not been used for the GCS family of hemoproteins, thus presenting an additional of means of probing the transduction mechanism. (Table 4) Wild-type *Pcc*GCS has a redox potential (E^o) of -2.9 mV, which is similar to the redox potentials of a previously measured GCS protein, *Ec*DosC, which has a E^o of -17 mV.[15] In contrast, the *Pcc*GCS(Y57F) variant shows a significant increase in E^o compared to WT to 58 mV(Fig. 4). We attribute this increase to the loss of distal pocket water that previous work has indicated is present in *Pcc*GCS and *Pcc*Globin and is lost with the Y57F varient.[28] Previous work on heme proteins has shown redox potential increases when water/solvent is removed from the heme pocket.[29-31] This same trend can be observed when comparing the full-length, *Pcc*GCS, to its isolated globin, *Pcc*Globin which has a redox potential of -15 mV. As a direct result of the removal of the additional domains, *Pcc*Globin likely contains more solvent near or around the heme pocket and thus exhibts a significantly lower redox potential.

The W86H mutant also exhibits a significantly increased redox potential (71 mV). Taken together with studies performed on *P. dendritiformis* GCS *Pd*GCS (Patterson, et al. *Manuscript in prepartion*), which has a His residue at this position and a reduction potential of 405 mV, this data suggest that this heme edge position greatly modulates the heme reduction potential in these proteins. It is possible that the large shift in the k_{cat} of W86H is due to the change in electronics of the porphyrin π system brought about by the increased reduction potential. This change in electronic structure may result in greater porphyrin flexibility, allowing the porphyrin ring to twist into a greater 'on' position than is possible in the wild-type protein, resulting in the increased k_{cat}.

Table 4. Redox Potentials (mV vs SHE)

Protein	mV
*Pcc*GCS	-3 ± 6
*Pcc*GCS(Y57F)	58 ± 7
*Pcc*GCS(R78G)	TBD
*Pcc*GCS(S82A)	TBD
*Pcc*GCS(W86A)	43 ± 3
*Pcc*GCS(W86H)	71 ± 3
*Pcc*GCS(H112G)	N.A.
*Pcc*GCS(V122Y)	30 ± 4
*Pcc*GCS(W86H/V122Y)	35 ± 7
*Pcc*Globin	-15 ± 3

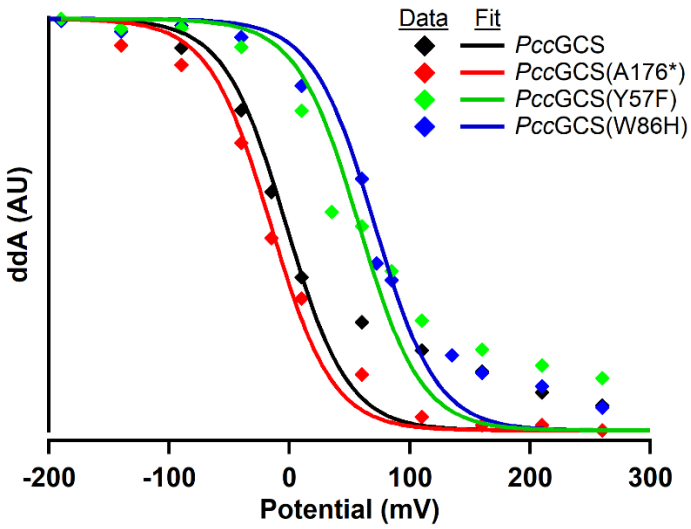

Figure 4. Redox data and fit for selected variants. *Pcc*GCS (black) can be compared to *Pcc*Globin (red), *Pcc*GCS(Y57F) (green) and *Pcc*GCS(W86H) (blue), showing the effect of the mutations on the midpoint.

The V122Y variant shows a more modest increase in reduction potential (30 mV), indicating that it does not exert as strong an effect on the heme cofactor as the Trp86 location. However, it also

follows the trend of a higher redox potential associates with higher kinetic rates. It is also of note that the V122Y is irreversible compared to the previously discussed variants. This means that the redox could only be performed going Fe(II) to Fe(III), and thus V122Y is unstable in the Fe(III) form. The double mutant shows a slight increase redox (35 mV). This data, taken with the oxygen dissociation rates and kinetic as compared to V122Y and W86H single varients, indicates that the V122Y mutant detracts from the major effect of the W86H mutant when combined.

3. Conclusion

In conclusion, we have determined that the roles of the distal Tyr, distal Ser and proximal His in *Pcc*GCS is mainly oxygen and heme binding stability, respectively. The distal Ser does not play an obvious role in the signaling pathway as its mutation does not significantly alter kinetics or oligomeric ratios. The main effect of S82A is the increases in oxygen dissociation rates compared to the WT. In both the Tyr57 and His112 cases, it is impossible to determine a role in the direct signaling path due to the inability of the Y57F and H112G mutations to form the high activity state. The Y57F mutation significantly increases the redox potential of the GCS, but there is no notable change in oligomeric ratios or kinetic rates. In the case of H112G, the oligomeric ratio is significantly shifted towards the high-activity, tetramer state. However, the ligation limitation of H112G restricts the ability to determine kinetic rate changes or redox potential.

The most notable data from this work is the identification and confirmation that heme edge residues significantly alter the properties of *Pcc*GCS. Examination of Trp86 and Val122 have shown that both locations alter oxygen binding, oligomeric distribution, and redox potentials.

The pi-stacking of the heme edge Trp is thought to explain the large effect mutating this single location has on the full-length protein. Additionally, V122Y and W86H/V112Y show the addition of an aromatic residue near the heme edge affects the ability to distinguish Fe(II)-O2 verse Fe(II)-NO. This work suggests that heme deformation in the major component of the signaling pathway in GCSs, resulting in the signaling pathway from the globin domain to the output domain.

4. Chapter References

1. Green, J.; Rolfe, M. D.; Smith, L. J. Transcriptional regulation of bacterial virulence gene expression by molecular oxygen and nitric oxide. *Virulence* 2014, 5 (8), 794-809 DOI: 10.4161/viru.27794.

2. Burns, J. L.; Deer, D. D.; Weinert, E. E. Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. *Mol. Biosyst.* 2014, 10 (11), 2823-6.

3. Wu, C.; Cheng, Y. Y.; Yin, H.; Song, X. N.; Li, W. W.; Zhou, X. X.; Zhao, L. P.; Tian, L. J.; Han, J. C.; Yu, H. Q. Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. *Sci. Rep.* 2013, 3, 1945-1951 DOI: 10.1038/srep01945.

4. Sawai, H.; Yoshioka, S.; Uchida, T.; Hyodo, M.; Hayakawa, Y.; Ishimori, K.; Aono, S. Molecular oxygen regulates the enzymatic activity of a heme-containing diguanylate cyclase (HemDGC) for the synthesis of cyclic di-GMP. *Biochim. Biophys. Acta* 2010, 1804 (1), 166-72 DOI: 10.1016/j.bbapap.2009.09.028.

5. Shimizu, T.; Huang, D.; Yan, F.; Stranava, M.; Bartosova, M.; Fojtikova, V.; Martinkova, M. Gaseous O2, NO, and CO in signal transduction: structure and function relationships of heme-based gas sensors and heme-redox sensors. *Chem. Rev.* 2015, 115 (13), 6491-533 DOI: 10.1021/acs.chemrev.5b00018.

6. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* 2015, 10 (6), 845.

7. Walker, J. A.; Rivera, S.; Weinert, E. E. Mechanism and Role of Globin Coupled SensorSignaling. *Adv Microb Physiol* 2017, 71, 133-169.

8. Romling, U.; Galperin, M. Y.; Gomelsky, M. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* 2013, 77 (1), 1-52 DOI: 10.1128/MMBR.00043-12.

9. Burns, J. L.; Jariwala, P. B.; Rivera, S.; Fontaine, B. M.; Briggs, L.; Weinert, E. E. Oxygen-Dependent Globin Coupled Sensor Signaling Modulates Motility and Virulence of the Plant Pathogen Pectobacterium carotovorum. *ACS chemical biology* 2017, 12 (8), 2070-2077 DOI: 10.1021/acschembio.7b00380.

10. Hengge, R. Principles of c-di-GMP signalling in bacteria. *Nature Rev. Microbiol.* 2009, 7 (4), 263-73 DOI: 10.1038/nrmicro2109.

11. Wan, X.; Tuckerman, J. R.; Saito, J. A.; Freitas, T. A.; Newhouse, J. S.; Denery, J. R.; Galperin, M. Y.; Gonzalez, G.; Gilles-Gonzalez, M. A.; Alam, M. Globins synthesize the second messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* 2009, 388 (2), 262-70 DOI: S0022-2836(09)00305-2 [pii]

10.1016/j.jmb.2009.03.015.

12. Chan, C.; Paul, R.; Samoray, D.; Amiot, N. C.; Giese, B.; Jenal, U.; Schirmer, T. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* 2004, 101 (49), 17084-9 DOI: 10.1073/pnas.0406134101.

13. Vinogradov, S. N.; Tinajero-Trejo, M.; Poole, R. K.; Hoogewijs, D. Bacterial and archaeal globins - A revised perspective. *Biochim. Biophys. Acta* 2013, 1834, 1789-1800 DOI: 10.1016/j.bbapap.2013.03.021.

14. Borriello, G.; Werner, E.; Roe, F.; Kim, A. M.; Ehrlich, G. D.; Stewart, P. S. Oxygen limitation contributes to antibiotic tolerance of *Pseudomonas aeruginosa* in biofilms. *Antimicrob. Agents Chemother.* 2004, 48 (7), 2659-64 DOI: 10.1128/AAC.48.7.2659-2664.2004.

15. Kitanishi, K.; Kobayashi, K.; Kawamura, Y.; Ishigami, I.; Ogura, T.; Nakajima, K.; Igarashi, J.; Tanaka, A.; Shimizu, T. Important roles of Tyr43 at the putative heme distal side in the oxygen recognition and stability of the Fe(II)-O2 complex of YddV, a globin-coupled heme-based oxygen sensor diguanylate cyclase. *Biochemistry* 2010, 49 (49), 10381-93 DOI: 10.1021/bi100733q.

16. Phippen, B. L.; Oliver, J. D. Role of anaerobiosis in capsule production and biofilm formation in Vibrio vulnificus. *Infect. Immun.* 2015, 83 (2), 551-9 DOI: 10.1128/IAI.02559-14.

17. Roberts, R. R.; Hota, B.; Ahmad, I.; Scott, R. D., 2nd; Foster, S. D.; Abbasi, F.; Schabowski, S.; Kampe, L. M.; Ciavarella, G. G.; Supino, M.; Naples, J.; Cordell, R.; Levy, S. B.; Weinstein, R. A. Hospital and societal costs of antimicrobial-resistant infections in a Chicago

teaching hospital: implications for antibiotic stewardship. *Clin. Infect. Dis.* 2009, 49 (8), 1175-84 DOI: 10.1086/605630.

18. Sommer, F.; Backhed, F. The gut microbiota-masters of host development and physiology. *Nat Rev. Microbiol.* 2013, 11 (4), 227-38 DOI: 10.1038/nrmicro2974.

19. Tucker, C.; Baroso, G.; Tan, P. Ventilator-associated pneumonia due to Shewanella putrefaciens. *Am J Health Syst Pharm* 2010, 67 (12), 1007-9 DOI: 10.2146/ajhp090344.

20. Zhang, W.; Phillips, G. N., Jr. Structure of the oxygen sensor in *Bacillus subtilis*: signal transduction of chemotaxis by control of symmetry. *Structure* 2003, 11 (9), 1097-110 DOI: S0969212603001692 [pii].

21. Zhang, W.; Phillips, G. N., Jr. Crystallization and X-ray diffraction analysis of the sensor domain of the HemAT aerotactic receptor. *Acta Crystallogr D Biol Crystallogr* 2003, 59 (Pt 4), 749-51 DOI: S0907444903003299 [pii].

22. Tarnawski, M.; Barends, T. R.; Schlichting, I. Structural analysis of an oxygen-regulated diguanylate cyclase. *Acta Crystallogr D Biol Crystallogr* 2015, 71 (Pt 11), 2158-77 DOI: 10.1107/S139900471501545X.

23. Greene, B. L.; Wu, C.-H.; Vansuch, G. E.; Adams, M. W.; Dyer, R. B. Proton inventory and dynamics in the Nia-S to Nia-C transition of a [NiFe] hydrogenase. *Biochemistry* 2016, 55 (12), 1813-1825.

24. Greene, B. L.; Vansuch, G. E.; Wu, C.-H.; Adams, M. W.; Dyer, R. B. Glutamate Gated Proton-Coupled Electron Transfer Activity of a [NiFe]-Hydrogenase. *J. Am. Chem. Soc.* 2016, 138 (39), 13013-13021.

25. Rivera, S.; Burns, J. L.; Vansuch, G. E.; Chica, B.; Weinert, E. E. Globin domain interactions control heme pocket conformation and oligomerization of globin coupled sensors. *J. Inorg. Biochem.* 2016, 164, 70-76 DOI: 10.1016/j.jinorgbio.2016.08.016.

26. Ohta, T.; Yoshimura, H.; Yoshioka, S.; Aono, S.; Kitagawa, T. Oxygen-sensing mechanism of HemAT from *Bacillus subtilis:* a resonance Raman spectroscopic study. *J. Am. Chem. Soc.* 2004, 126 (46), 15000-1 DOI: 10.1021/ja046896f.

27. Burns, J. L.; Rivera, S.; Deer, D. D.; Joynt, S. C.; Dvorak, D.; Weinert, E. E. Oxygen and c-di-GMP Binding Control Oligomerization State Equilibria of Diguanylate Cyclase-Containing Globin Coupled Sensors. *Biochemistry* 2016, 55, 6642-6651.

28. Rivera, S.; Young, P. G.; Hoffer, E. D.; Vansuch, G. E.; Metzler, C. L.; Dunham, C. M.; Weinert, E. E. Structural Insights into Oxygen-Dependent Signal Transduction within Globin Coupled Sensors. *Inorg. Chem.* 2018, 57 (22), 14386-14395.

29. Tezcan, F. A.; Winkler, J. R.; Gray, H. B. Effects of ligation and folding on reduction potentials of heme proteins. *J. Am. Chem. Soc.* 1998, 120 (51), 13383-13388.

30. Murgida, D. H.; Hildebrandt, P. Redox and redox-coupled processes of heme proteins and enzymes at electrochemical interfaces. *PCCP* 2005, 7 (22), 3773-3784.

31. Stellwagen, E. Haem exposure as the determinate of oxidation–reduction potential of haem proteins. *Nature* 1978, 275 (5675), 73.

Chapter 5: Conclusion - Signaling in Globin Coupled Sensors Proteins

Chapter 5: Conclusion

Signaling in Globin Coupled Sensors Proteins

Chapter Overview:

1. Introduction
2. Knowledge Gained
3. Proposed Future Work
4. Chapter References

1. Introduction

GCS are linked to bacteria motility, virulence, biofilm formation, and gene regulation. Depending on the environment of the original bacterium, globin coupled sensor proteins (GCSs) are activated by different gaseous ligands: O_2, CO or NO.[1-3] Bacterial infections occur under these various types of environment and often form biofilms at the environmental changes of these gases (i.e. at the aerobic/anaerobic border or at the NO_2/NO border).[1-5] As discussed in previous chapters, biofilms are a primary defense used by bacteria for growth and replication.[2] They are a major component for antibiotic resistance, and thus, information regarding their regulation and formation is highly sought after by the medical community. By investigating the role(s) of GCS signaling in controlling bacterial biofilm formation in response to changing O_2 levels, studies have the potential to identify novel methods to target bacterial resistance.

2. Knowledge Gained

Through this body of work, the goal has been to elucidate the molecular signaling pathway that allows the globin domain of GCS proteins to alter output domain activity. Previous studies had demonstrated that output domain activity was directly linked to the ligation state of the heme,

thus highlighting the gas binding signal in the regulation of the output domain of GCSs.[6, 7] However, the interactions required for sensing and transmitting the ligand binding signal were unknown.

The studies described in this doctoral thesis have provided information about key residues involved in stabilization of ligand binding and output domain activation, which have implications for the GCS community and the broader scientific community interested in bacterial sensing and signaling. We have demonstrated that the distal Tyr primary serves a role in stabilizing oxygen binding and has very little effect on signaling, highlighting the roles of other interactions within the globin domain in output domain activation. Through structural studies, we identified that heme distortion differs significantly between ligation states and, through interactions with heme pocket residues, provides the ligand dependent signal. Additionally, the crystallization of *Bpe*Globin helped identify additional amino acid locations, termed 'heme-edge residues', that were previous overlooked but play important roles in signal transduction. Investigation into the heme edge Trp (Trp86) of *Pcc*GCS shows the significance of the location which T-pi stacks with the heme to kinetic turn-over rates. Also, *Pcc*GCS(V122Y) has shown how placing aromatics at key locations around the heme edge affects ligand recognition and kinetic response.

Insights into the globin domain's role with oligomerization and oxygen binding have also been gained. We have observed shifts in slow and fast rate percentages for oxygen dissociation resulting from isolation of the globin domain. This indicated that the heme pocket opening is affected by the additional two domains, but not the physical rates of dissociation. Additionally,

the stability of the isolated globin domain was determined to be greater than the full-length counter parts. This allowed for the isolated globins to remain soluble for days in the Fe(II)-O_2 conformation directly resulting in the crystallization of *Bpe*Globin to exist in the high activity state.

3. Proposed Future Work

Although certain aspects of the signaling pathway and mechanism for GCS are now understood, many aspects of GCS signaling would benefit from further study. For the signaling pathway, identifying additional amino acids involved in the pathway, especially residues outside of the heme pocket, is key. Other heme enzymes contain aromatic residues throughout the protein that help shuttle electrons and thus signal.[8-10] Additionally, multiple GCS contain aromatics at key locations around the heme, i.e. the distal Tyr, heme edge Trp, His, Tyr, and the proximal His.[7] Is there a potential pathway though these aromatics?

Even now the community is working on obtaining the first full length crystal structure of a GCS, the Weinert lab included. Since there are no full-length X-ray crystal or cryo-EM structures, it is difficult to understand the complete role of the middle domain, as well as the distance between the sensory and the output domains. The Weinert lab has used chemical cross-linking and mass spectrometric analysis to support the premise of certain connections or distances between domains,[11] and other labs have crystalized individual domains to allow docking programs to identify the quaternary structure,[12] but it does not provide the resolution of other structural techniques.

The crystal structure of *Bpe*Globin was highly impactful for determining additional residues. However, *Pcc*Globin and *Pd*Globin also need to be crystalized if their full lengths cannot be. Additional crystal structures, even of the globin domain, will elucidate the locations around the heme pocket that a single crystal structure cannot, and it may be important to identify trends that associate a type of heme pocket to its respective output domain.

Further understanding of the signaling pathway is necessary for the next step in the GCS community, to attempt regulation and alteration of biofilm formation or virulence factors. We have shown that the beginning of the pathway is more complex that initially thought but have also provided insight into the additional possibilities of heme deformation and electronic changes being involved in the pathway. There is potential of future therapeutics against bacterial infections by gaining this knowledge.

4. Chapter References

1. Ahmad, I., Lamprokostopoulou, A., Le Guyon, S., Streck, E., Barthel, M., Peters, V., Hardt, W. D., and Romling, U. (2011) Complex c-di-GMP signaling networks mediate transition between virulence properties and biofilm formation in *Salmonella enterica serovar Typhimurium*, *PLoS One 6*, e28351.

2. Armbruster, C. R., Forster, T. S., Donlan, R. M., O'Connell, H. A., Shams, A. M., and Williams, M. M. (2012) A biofilm model developed to investigate survival and disinfection of *Mycobacterium mucogenicum* in potable water, *Biofouling 28*, 1129-1139.

3. Bjarnsholt, T., Jensen, P. O., Fiandaca, M. J., Pedersen, J., Hansen, C. R., Andersen, C. B., Pressler, T., Givskov, M., and Hoiby, N. (2009) *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients, *Pediatr. Pulmonol. 44*, 547-558.

4. Burns, J. L., Deer, D. D., and Weinert, E. E. (2014) Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors, *Mol. Biosyst. 10*, 2823-2826.

5. Carlson, H. K., Vance, R. E., and Marletta, M. A. (2010) H-NOX regulation of c-di-GMP metabolism and biofilm formation in *Legionella pneumophila*, *Mol. Microbiol. 77*, 930-942.

6. Rivera, S., Burns, J. L., Vansuch, G. E., Chica, B., and Weinert, E. E. (2016) Globin domain interactions control heme pocket conformation and oligomerization of globin coupled sensors, *J. Inorg. Biochem. 164*, 70-76.

7. Walker, J. A., Rivera, S., and Weinert, E. E. (2017) Mechanism and Role of Globin Coupled Sensor Signaling, *Adv Microb Physiol 71*, 133-169.

8. Mayo, S. L., Ellis, W. R., Crutchley, R. J., and Gray, H. B. (1986) Long-range electron transfer in heme proteins, *Science 233*, 948-952.

9. Immoos, C. E., Di Bilio, A. J., Cohen, M. S., Van der Veer, W., Gray, H. B., and Farmer, P. J. (2004) Electron-Transfer Chemistry of Ru− Linker−(Heme)-Modified Myoglobin: Rapid Intraprotein Reduction of a Photogenerated Porphyrin Cation Radical, *Inorg. Chem. 43*, 3593-3596.

10. Wuttke, D. S., Bjerrum, M. J., Winkler, J. R., and Gray, H. B. (1992) Electron-tunneling pathways in cytochrome c, *Science 256*, 1007-1009.

11. Burns, J. L., Jariwala, P. B., Briggs, L., Fontaine, B. M., and Weinert, E. E. (2017) Oxygen-Dependent Globin Coupled Sensor Signaling Modulates Motility and Virulence of the Plant Pathogen *Pectobacterium carotovorum*, *ACS Chem. Biol. 12*, 2070-2077.

12. Tarnawski, M., Barends, T. R., and Schlichting, I. (2015) Structural analysis of an oxygen-regulated diguanylate cyclase, *Acta Crystallogr D Biol Crystallogr 71*, 2158-2177.