Georgia State University

ScholarWorks @ Georgia State University

Characterization of Heme Binding to PEFR, A Marr Transcriptional Regulator

Rosemary Andrews
randrews15@student.gsu.edu

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_theses

CHARACTERIZATION OF HEME BINDING TO PEFR, A MARR TRANSCRIPTIONAL

REGULATOR

by

ROSEMARY ANDREWS

Under the Direction of Dabney K. W. Dixon, PhD

ABSTRACT

PefR is a heme-binding transcriptional regulator found in *Streptococcus pyogenes,* a β-hemolytic pathogen that can be fatal. Heme binding results in the release of the protein from DNA and allows for the transcription of a proposed heme efflux system. The overall goals of this work were to devise an effective protocol to purify DNA-free PefR, verify the stoichiometry of heme binding, and identify the axial ligand. Strep-tagged PefR was purified by a variety of protocols, the most effective of which utilized high salt concentrations in the buffers and an ion-exchange column. Based on UV/visible studies of PefR, we suggest that heme binds via Cys109. Homology models place Cys109 in a region that several MarR proteins utilize to bind their ligands, supporting our hypothesis.

INDEX WORDS: PefR, Heme-binding, Thiolate-binding, Group-A Streptococcus, Heme sensor, Transcriptional regulator, Protein purification, *Streptococcus pyogenes,* DNA transcription regulator

CHARACTERIZATION OF HEME BINDING TO PEFR, A MARR TRANSCRIPTIONAL

REGULATOR

by

ROSEMARY ANDREWS

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2020

CHARACTERIZATION OF HEME BINDING TO PEFR, A MARR TRANSCRIPTIONAL

REGULATOR

by

ROSEMARY ANDREWS

Committee Chair: Dabney Dixon

Committee: Kathy Grant

Gregory Poom

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2020

DEDICATION

This thesis is dedicated to my family and friends. To my parents, I could not have done this without you. To all my siblings, this degree means that even though I'm the youngest, I'm now officially the smartest. To all my friends who have supported me through this process, sorry this thesis was all I talked about for two years, I could not be more thankful for you and all you put up with. Also to Lucy, I know you can't read because you're a cat, but I love you a lot.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CD, circular dichroism; CO, carbon monoxide; DMSO, dimethyl sulfoxide; HRM, Heme-Regulatory Motif; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, Luria-Bertani; PMSF, phenylmethylsulfonyl fluoride; ROS, reactive oxygen species; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; PMe_3, trimethylphosphine; WT, wild-type.

1   INTRODUCTION

Heme is widely known for its use in oxygen transport and electron transfer. More recently, its role in signaling has been highlighted [15, 38, 76, 107, 114]. Heme-based sensors can be used to sense gas molecules, redox conditions, or intracellular heme levels. These sensors control many important systems, including circadian rhymes, immune responses, the activity of ion channels, and interactions with reactive oxygen species (ROS). Many of these signaling proteins directly regulate transcription, including the transcription of systems directly related to iron synthesis, sequestration, regulation, and efflux. The concentration of heme and free iron in cells is an important and delicate balance. Free iron can produce ROS, but a high concentration of heme is toxic and will kill cells. Many pathogens used heme as their primary iron source [7, 121].

The study of pathogenic cells and their methods for survival is growing ever more important as strains of antibiotic-resistant diseases emerge [47]. *Streptococcus pyogenes* is a gram positive, β-hemolytic pathogen that one estimate suggests results in ~ 600,000 deaths per year occurring from the approximately 600 million infections [90, 131]. Cases of antibiotic-resistant strains are increasing. *S. pyogenes* requires iron for growth and can use heme as an iron source [30]. The intracellular heme concentration of in *S. pyogenes* is partially controlled by PefR, a heme-binding transcriptional regulator [95].

PefR belongs to the MarR (multiple antibiotic resistance regulator) family, and has been studied in *S. pyogenes* [95] as well as in *Streptococcus agalactiae* [32]. In *S. agalactiae,* PefR controls, in a heme-dependent manner, the expressions of two separate gene clusters (*pefAB* and *pefCD*) that both encode for proposed transport proteins [32]. The *S. pyogenes* PefR regulates the transcription of the *pefRCD* genes, where *pefC* and *pefD* each encode for one of the two subunits

that make up a heterooligomeric ATP-dependent exporter [94, 95]. This exporter is a multidrug resistance efflux system that prevents the buildup of intracellular heme, protects cells from DNA damage by ROS, and decreases sensitivity to certain drugs. This gene cluster is highly conserved among *S. pyogenes* strains, with a 100% identity for PefR in the 20 strains evaluated.

The overarching goal of this work is to characterize heme binding to PefR. Initially, this required the purification of DNA-free PefR. Titration of PefR with hemin was used to reveal that PefR binds one heme molecule per monomer. The UV/visible spectra of ferric holo-PefR suggests heme ligation via a Cys. The spectrum of ferrous CO-bound holo-PefR points to a neutral axial ligand. Sequence alignment and homology models suggest that that heme binds to PefR via Cys109.

Among transcriptional regulators, coordination of heme-binding in PefR is somewhat unusual, as most thiolate-based transcriptional regulators bind heme via a <u>C</u>ys-<u>P</u>ro (CP motif) [149]. PefR contains only one Cys, and it is not next to a Pro. Thorough characterization and study of PefR enhances our understanding of heme binding and signaling, which continues to show its involvement in a diverse assortment of research foci.

1.1 PefR

Transcriptome studies in *S. agalactiae* under different respiration conditions discovered changes in expression levels of two <u>p</u>orphyrin <u>ef</u>flux (pef) pump systems, as well as a protein that regulates the expression of this system [32]. These two porphyrin efflux systems are respectively composed of gene clusters *pefAB* and *pefRCD,* which are separated in the genome. The expression levels of both clusters are controlled by PefR.

An 18 bp inverted repeat (within a 23 bp consensus sequence) is found upstream of both *pefAB* and *pefRCD* regions but seen nowhere else in the genome [32]. This 23 bp region also

contains the -10 region of a punitive promotor. PefR binds the inverted repeat motif upstream of both gene clusters *pefAB* and *pefRCD*. Mobility shift assays with these promotor regions showed PefR concentration-dependent shifts. A northern blot study with *wt* and *ΔpefR* using *pefAB* and *pefRCD* as probes showed an increase in expression of both with *ΔpefR* compared to *wt*. PefR binds this operon directly to repress *pefAB* and *pefRCD* expression. This was demonstrated by studying a construct with the *lacz* gene placed upstream of the *pefA* promotor region. A β-galactosidase assay with this region and *wt* and mutant Δ*pefR* showed that downstream expression of *pefA* with the *ΔpefR* was 12 times higher than with the *wt*. A β-galactosidase assay of the *lacz-pefA* promotor regions with the addition of heme (0.1 to 10 µM) in the presence of PefR showed increased expression of *pefA* by up to nine-fold. This is presumably due to the heme-bound form of PefR being released from the DNA, thus allowing the expression to take place. In line with this, mobility shift assays showed that increasing concentrations of heme prevented binding of PefR to *pefAB* and *pefRCD*. Northern blot assays showed that both heme and PPIX induced expression of *pefAB* and *pefRCD*. Expression was also seen with protoporphyrin (PPIX), zinc mesoporphyrin, and gallium PPIX in a β-galactosidase assay. Free iron did not affect induction.

PefR in *S. pyogenes* was initially described by Eichenbaum and co-workers [95]. The apo-protein has a molecular mass of 17.35 kDa, a theoretical PI of 9.1, and an extinction coefficient at 280 nm of 7,450 M^{-1} cm^{-1} [36]. In *S. pyogenes,* transcription levels of the three-gene cluster *pefRCD* increased in response to increased heme levels [95]. These three genes have sequence similarities of 84%, 76%, and 76% to *pefRCD* in *S. agalactiae*. These three-gene clusters are highly conserved among *S. pyogenes* strains, showing 100%, 98%, and 99% identity for the PefR, PefC, and PefD proteins, respectively [94].

A 17 bp repeat region with 76% shared identity with the PefR binding site in the promoter region was also seen [95]. Microarray analysis showed that PefR and PefC expression was comparable over time, with a 2.5-fold increase in expression 30 min after heme exposure, and a 4.5-fold increase after 60 min.

Eichenbaum and colleges prepared a His-tagged PefR. However, in view of the potential binding of heme to the His-tag itself [83, 128], our laboratory has prepared a second construct with a Strep-tag. Previous members of our laboratory transformed cells to allow for Strep-tagged PefR expression [127]. The optical spectra of the His-tag and Strep-tag protein were different. The His-tagged PefR has a large Soret at 407 nm, with a shoulder peak at 350 nm and α/β bands at 540 and 560, and 665 nm. The Strep-tagged protein showed a Soret band at 382 nm, with a visible band at 510 nm and charge transfer band at 615 nm. As-isolated His-tagged PefR showed a significant level of bound heme, while as-isolated Strep-tagged PefR showed almost no heme loading.

1.2 MarR

MarR transcriptional regulators are a large family, with over 12,000 MarR-like proteins identified; this family has been subject to a number of reviews in the past ten years [27, 40, 41, 44, 63, 85, 136]. The Pfam database currently shows 3366 sequences, 5392 species, and 70 structures [31].

MarR proteins regulate key processes such as response to oxidative stress, catabolism of aromatic compounds, and production of virulence factors. MarR regulators bind a variety of ligands, most of which are small molecules. Several MarR proteins are well-characterized oxidative stress sensors, some of which utilize oxidative cysteine chemistry as a mechanism for transcriptional regulation. The MarR family shows low sequence homology, but high structural

homology. The low sequence homology within this family is thought to contribute to the specificity of the proteins, as the regions that bind DNA and ligands vary widely. MarR dimers often bind palindromic sequences adjacent to the gene they regulate. Most MarR regulators are repressors; however, there are a small number that act as activators. A common repression strategy involves the MarR regulator binding DNA and preventing transcription until ligand binds and the regulator dissociates.

MarR regulators tend to be triangularly shaped and are commonly found as homodimers. Monomers are typically made up of six α helices and at least one β sheet. Some urate-binding MarR homologs have seven helices, with an extra helix at the N-terminus. A fundamental structural aspect of MarR family proteins is the winged helix-turn-helix (wHTH) domain that binds DNA. Helices α2, α3, and α4, along with β1 (the loop connecting them) and β2, form the wHTH domain [12, 19, 122]. The wings bind to DNA via the minor groove and recognition helices bind the major groove. Helices α1 and α5 are proposed to connect the dimerization and DNA binding domains. In many cases, the ends of the N- and C-terminal helices are credited with the dimerization regions.

1.3 Heme Binding

1.3.1 Characterization of Heme Binding via a Thiolate

There are two classifications that are pertinent to understanding heme-binding transcriptional regulators that utilize thiolates and the broader field of heme-binding via a thiolate. These classifications are type-1 binding, type-2 binding, and the utilization of a CP motif.

A number of proteins, such as kinases and ion channels, have the ability to regulate transcription indirectly. There are reviews on heme-binding via a thiolate and transcription

regulation by heme that include such proteins [15, 38, 61, 76, 105, 107]. For the sake of brevity, only proteins that bind both heme and nucleic acids as their primary function are defined as transcriptional regulators in this work.

Select model systems are also included in this work. Exogenous thiols have been used to imitate Cys, most notably in the H93G myoglobin mutant [29]. β-Mercaptoethanol is often used as the exogenous thiolate [29, 89, 125], while ethanethiol [91] and cyclopentanethiol has also been used to imitate neutral thiols [29, 86, 119]. Exogenous thiolate ligands have a tendency to fall out of the binding cavity upon reduction [86, 91].

1.3.1.1   Type-1 and Type-2 Cys Ligated Heme-Binding Proteins

Smith et al. have sorted thiolate heme-binding proteins into two categories [114]. Type-1 heme-thiolate proteins are involved in the activation of small molecules; they maintain their thiolate ligand upon reduction and are catalytically competent in the five-coordinate, high-spin state. These include cytochrome P450$_{cam}$, chloroperoxidase in *Caldariomyces fumago*, nitric oxide synthase in *Bacillus subtilis*, and nitric oxide reductase in *Fusarium oxysporum*. One hypothesis for the lack of ligand substitution in type-1 proteins is that many utilize heme as part of their catalytic redox function [114]. This requires not only that the heme remain bound in both redox states, but also that the protein does not lose functionality as a result of reduction.

Type-2 heme thiolates are involved in the sensing or transport of small molecules. These proteins typically bind Fe(III) via both a Cys and another amino acid, often His. Almost all lose Cys upon reduction and may switch to a different axial ligand (ligand switch), resulting in a divergence in spectral characteristics compared to the type-1 proteins. A number of heme-binding transcriptional regulators are gas or redox sensors. It is logical that these sensor proteins

would utilize a more flexible binding environment to take advantage of conformation changes upon ligand binding or reduction.

Table 1 lists a number of cysteine-ligated proteins and their function, as well as their axial ligands in the ferric, ferrous, and CO bound states. This table includes transcriptional regulators that bind heme with Cys as the axial ligand alongside type-1 thiolate proteins and select model systems. The type-1 proteins and model systems are included because their spectroscopic features have been thoroughly studied.

1.3.1.2   Cysteine-Proline Binding Motif in Transcriptional Regulators

The most common heme-binding motif found in transcriptional regulators is the CP motif or heme regulatory motif (HRM), which has been widely reviewed in the past decade [15, 34, 38, 57, 76, 102, 107, 149]. CP motifs, as the name implies, require a cysteine and proline next to each other. CP motifs were first described by Zhang and Guarente, after they saw that the sequence Lys/Arg-Cys-Pro-Val/Ile-Asp-His repeated more than six times in Hap1, a sequence also found in several other heme-binding proteins [149]. Heme binds to the Cys, while the Pro is responsible for introducing a "kink" in the protein [61]. Models have shown this structural change affects the orientation of the heme which allows for improved association [67, 101, 105]. Table 1 notes transcriptional regulators that contain CP motifs.

1.3.2   *Effect of Reduction on Ligation of Heme-Based Transcriptional Regulators that Utilize Cys*

Table 1 lists a number of proteins and their function, as well as their axial ligands in the ferric, ferrous, and CO bound states. In general, transcriptional regulators have been shown to lose their Cys ligation upon reduction. Most switch to nearby ligands and some lose functionality [107, 114]. Several specific examples are described below.

1.3.2.1   Switching Ligands

Rcom-2, a CO sensing transcriptional regulator from *Burkholderia xenovorans*, coordinates heme via Cys94 and His74 [113]. Upon reduction, Cys is replaced with Met104 [14, 71]. This Cys does not belong to a CP motif and utilizes a Per-Arnt-Sim (PAS) domain instead. MCD spectra of ferric Rcom-2 are consistent with heme bound to a cysteine thiolate opposite a neutral donor. MCD and Raman evidence both indicate that CO and NO bind *trans* to a His.

CooA is a CO sensor from *Rhodospirillum rubrum* that binds heme ferric heme via Cys75 and the N-terminal Pro2 [22]. Upon reduction, Cys75 is replaced by His77. CooA does not bind DNA until CO has replaced Pro2, causing a conformational shift in the DNA binding domain [65]. This shift is redox-dependent and reversible.

Rev-erβ, a nuclear receptor and transcriptional repressor that regulates circadian rhythm pathways, binds ferrous heme via Cys384 and His568 [84]. Upon reduction, Cys418 is replaced by a neutral ligand, and forms a disulfide bond with Cys374 [45]. The reduction of heme leads to a 100-fold decrease in heme affinity, allowing Rev-erβ to act as a redox sensor. Reduced Rev-erβ exists as a mix of both a high-spin, five-coordinate species, and a low-spin, six-coordinate species [72, 84]. The reduced five-coordinate heme acts as a gas sensor for NO and CO.

1.3.2.2   Two Binding Sites

Irr is a transcriptional regulator from *Bradyrhizobium japonicum* that binds two molecules of heme in two different manners. One heme is in the five-coordinate high spin state with a Cys29 as the axial ligand, and the second heme is in a six-coordinate and predominately low spin state with two histidine ligands [56, 143]. The C29A mutant of *bj*Irr showed an identical spectrum to the wt, suggesting that Cys29 is not responsible for binding ferrous heme [143]. It has been suggested that binding at this site is not entirely lost, and heme is ligated to a

different axial ligand [56, 143]. Heme binding to *bj*Irr causes the protein to degrade, which leads to its release from the DNA and the de-repression of transcription. Heme binding to both sites is essential for degradation.

Mammalian Bach1 also has two heme-binding sites [49, 80]. Raman and mutational studies have shown that one becomes dominant at higher heme concentrations and binds a five-coordinate heme via a Cys; the other binds a six-coordinate heme via both Cys and His. The transcriptional regulation of Bach1 is co-mediated by binding to Maf.

IRP1 and IRP2 are mammalian regulators of iron metabolism that bind mRNA [81]. In the Fe(III) state, both a five-coordinate heme bound via a Cys and a six-coordinate heme bound to both a Cys and His are seen. Upon reduction to the Fe(II) state, the spectra show only a six-coordinate heme bound via His and either a neutral Cys or a His.

1.3.3 *Optical Spectra of Cys Transcriptional Regulators*

In the UV/visible spectra, five-coordinate ferric cysteine-bound heme proteins tend to show a broad, symmetrical Soret band at wavelengths below 400 nm, α and β bands in the 500-580 nm range, and a charge transfer band above 600 nm (Table 2). Dps [35] and heme-binding site 1 in mammalian Bach1 [49] exemplify this trend. Ferric Dps shows a Soret at 385 nm, with a shoulder at 365 nm [35]. Visible bands are also seen at 514, 544 and 650 nm. Dps binds a high spin, five-coordinate heme via Cys101. Wildtype Bach1 has Soret peaks at both 371 and 423 nm [49]. The isolated CP region that binds heme via Cys shows a peak at 371 nm.

Some transcriptional regulators have two heme binding sites. The transcriptional regulators Bach1, IRP1, IRP2, and *bj*Irr (Tables 1 and 2) all bind a five-coordinate heme with Cys as the axial ligand and show a Soret at approximately 370 nm. They also have a second, six-coordinate heme site with a Soret between 414 and 420 nm.

Several thiolates bind strictly six-coordinate heme. Mammalian RNA-binding protein DiGeorge Critical Region 8 (DGCR8) binds six-coordinate heme via two cysteines and shows a hyperporphyrin spectrum due to its bis-Cys ligation [11]. When a six-coordinate heme is bound via a Cys and a second neutral ligand, as in the cases of CooA, DHR51, Rev-erβ, and *bx*Rcom-2 (Tables 1 and 2), the UV/visible spectra often shows two bands, a Soret around 420 nm, along with a blue-shifted band around 360 nm. Going from Fe^{3+} to Fe^{2+} with no significant change in the position of the Soret, but shifts in the visible bands is characteristic of a redox-mediated ligand switch [71]. This phenomenon is seen with Rcom-2, CooA, DHR51 (Tables 1 and 3). Burstyn and colleagues have discussed the molecular origin of the blue-shifted Soret band [72].

1.3.4   CO Binding

Table 3 lists the axial ligands and locations of the Soret for the ferrous, CO-bound heme-binding proteins. The classic example of CO binding to a thiolate-ligated heme is cytochrome $P450_{cam}$, which has a thiolate as the axial ligand in ferrous state; the heme is also bound to water [88]. CO binds to give a Soret at ~450 nm [25, 46]. Nitric oxide reductase [78] and nitric oxide synthase [96, 118], both type-1 heme thiolates, bind CO to give peaks at 440 and 445 nm, respectively. Chloroperoxidase binds hemin via a Cys; CO binding to the reduced protein shifts the Soret to 446 nm [117]

The crystal structure of nitrophorin from *Cimex lectularius* shows that both the Cys and CO are bound simultaneously (1SI6). When bound to CO, the Soret shifts to 422 nm [130]. It has been suggested that the blue-shifted Soret is due to heme binding to the neutral thiol rather than the thiolate [114].

H93C mutants of myoglobin have also studied [2, 48, 91]. In human myoglobin, carbon monoxide binding to ferrous holo-H93C shifts the Soret to 420 nm [2]. In horse heart myoglobin,

CO binding causes the Soret of the H93C and H64V/H93C mutants to shift to 422 and 421 nm, respectively [48].

1.3.5   Phosphine Studies

Phosphines have been used for many years as heme ligand to probe heme active sites. Early work, largely with trimethylphosphine, was reviewed by Simonneaux [111]. Ferric heme proteins with a thiolate ligand have a hyperporphyrin spectrum, e.g., a split Soret [93, 116]. Ferric cytochrome $P450_{cam}$, when titrated with bis(hydroxymethyl)methylphosphine gave a low-spin complex with split Soret with peaks at 375 and 446 nm [117]. Ferric chloroperoxidase, when treated with the same phosphine, showed a split Soret at 376 and 450 nm [117]. The titration was isosbestic, indicating a clean conversion to single phosphine adducts. More recently, Dps has been treated with tris(2-carboxyethyl)phosphine and showed a split Soret with bands at 370 and 441 nm [35]. This spectral consistency with $P450_{cam}$ and chloroperoxidase [117] was used to support the hypothesis that Cys was the axial ligand in Dps.

Ferrous cytochrome $P450_{cam}$ bound to bis(hydroxymethyl)methylphosphine showed bands at 337, 365, and 457.5 nm [117]. Ferrous chloroperoxidase, when treated with the same phosphine, gave a band at 457.5 nm [117].

The studies of these Cys binding proteins can be compared to studies done with histidine-binding proteins, for example, myoglobin. Titration of ferrous myoglobin with trimethylphosphine results in the Soret shifting from 436 to 430 nm and the α/β region transforming from one broad peak at 555 nm to two sharper peaks at 535 and 568 nm [17]. For ferric sperm whale myoglobin, the addition of 3-4 equivalents of trimethylphosphine gave a complex with peaks at 370, 424, and 536 nm [112].

Binding constants for phosphines have been measured in select systems. For ferric proteins, the dissociation constant for chloroperoxidase with bis(hydroxymethyl)methylphosphine was 2.0 mM [117]. For ferrous proteins, the dissociation constant for chloroperoxidase is 7.7 mM. For ferrous sperm whale myoglobin, the dissociation constant with trimethylphosphine was 1.1 µM and the dissociation constant for horse heart myoglobin with trimethylphosphine is 1.2 µM, which is similar to that of O_2 and about 40-fold less than that of CO [17].

1.3.6   Crystal Structures of Heme-Binding Transcription Factors

There are a limited number of heme-binding transcriptional regulators crystallized in the presence of heme. CooA (1FT9) [65] and HrtR (3VP5) [99] have been crystallized in their entireties with heme bound. Both have DNA binding and heme-binding domains perpendicular to each other. Figure 1 shows both crystal structures, with axial ligands highlighted.

CooA, which utilizes a CP motif, binds ferric heme via Cys75 and the N-terminal Pro2. Upon reduction, Cys75 is replaced by His77. The heme binding pocket of CooA is hydrophobic, while the residues surrounding the pocket are comparatively hydrophilic. Pro2 is adjacent to another Pro; these are two of the only residues in this binding pocket that are not hydrophobic. Ser78 and Arg4 appear to be stabilizing the other propionates via electrostatic interactions.

HrtR is an example of a heme-binding transcriptional regulator that does not utilize a CP motif. This protein controls the heme efflux system in *Lactococcus lactis*. [99]. HrtR is a member of the TetR family of transcriptional regulators [66, 76]. HrtR is a dimer, and each monomer has nine conserved helices, as well as short helices at each terminus. Helices α4-9 and $α_c$ interact with heme [99]. A four-helix bundle made up the α8 and α9 helices from the two monomers form the dimer interface. Heme binds in a hydrophobic cavity produced by a four-

helix bundle consisting of α4, α5, α8, and α9 (Figure 1). Four helix bundles have been shown to be very effective heme-binding structures [21]. His72 and His149 bind the iron itself [99]. Similar to MarR regulators, HrtR also contains an HTH motif which interacts with the target DNA. Crystal structure and spectroscopic studies showed that heme binding causes a coil-to-helix transition that rotates α4 and shifts the recognition helices in the HTH too far apart to continue to stay bound to the DNA, causing depression of transcription [76, 99]. HrtR is also found in *S. pyogenes* and works alongside PefR to control intracellular heme levels [95].

The heme-binding pocket of HrtR is deeper than CooA. This may be due to a loop made up from residues 125-135. This loop is not seen in apo-HrtR, suggesting its flexibility [99]. The exterior of the cleft is relatively negatively charged. The two vinyl groups are pointed inwards to a hydrophobic slightly positively charged pocket, while the two propionates point outwards.

1.4 Methods Background

1.4.1 Circular Dichroism

Circular dichroism (CD) uses circularly polarized light to detect the chirality of a sample [23, 39, 59]. For proteins, α helices produce negative bands at 222 and 208 nm and a positive band at 193 nm, while β sheets show a negative band at 218 nm and a positive band at 195 nm [39, 70, 137]. Disordered structures are characterized by a strong negative band around 200 nm and possible weak bands around 218 and 238 nm [70, 137]. A variety of computer programs can be used to deconvolute the spectra and estimate the percentages of these secondary structures [39, 70, 137]. CD is used to assess whether mutations of specific residues have large-scale effects on the protein secondary structure [28, 43, 59, 70, 110]. This helps to make sense of an overall structure [92, 140], as well as monitor changes in the structure of the protein that occur as a result of ligand binding [110, 137]. For CD spectra, concentrations can typically vary between

0.005 to 5 mg/ml and samples must be at least 95% pure [39, 59, 70, 110]. Due to their absorption in the near UV range, solutions should avoid chloride, nitrate, certain buffers, and DMSO as well EDTA at concentrations higher than 0.25 mM, HEPES and similar biological buffers at concentrations above 25 mM, and reducing agents at a concentration above 1 mM [59, 70]. Greenfield has listed lower wavelength limits for common buffers [39].

1.4.1.1 CD Studies of Heme Binding Transcription Factors

CD has been used to study heme-binding transcription factors. For example, the transcription factor hPer2, associated with circadian rhythm, binds heme in two distinct regions. One of the heme-binding domains [hPer2(V4-VII)] has 24% α-helical, 18% β-strand and 19% β-turn secondary structure. These percentages did not change significantly upon heme binding [140].

HrtR shows an increase the magnitude of the negative band at 208 nm upon heme binding, characteristic of an increase in the helical content. X-ray crystallography showed this to be due to a coil-to-helix transition in which two shorter helical regions (α4a and α4b) become one helix as the coiled region separating them also becomes part of the helix [99]. CD spectra of HrtR mutants without key amino acids in this coiled region did not show the same band change at 208 nm, showing the importance of these residues in the helix to coil transition [99].

1.4.1.2 CD Studies of MarR Proteins

MarR transcription factors have also been studied by CD. For example, HosA, a 4-hydroxybenzoic acid transcription factor from *Enterobacteriaceae*, has > 50% alpha helix [92]. MftR, a urate binding transcription factor from *Burkholderia thailandensis* was predicted to have 57% α-helical and 8% β-sheet content [43]. Table 4 has a number of additional literature examples.

CD has been used to evaluate structural changes upon ligand binding. For example, far-UV CD was used to study binding of 3-hydroxybenzoate to MobR, a MarR family transcriptional regulator found in *Comamonas testosteroni* [50, 146]. Binding of the ligand causes shifts in negative bands in the 250-270 and 290-300 nm range, indicative of phenylalanine and tryptophan residues respectively, showing that binding occurred near these residues [146]. Dialysis to remove the ligand produced CD spectra identical to the initial readings, showing that 3-hydroxybenzoate binds reversibly. ST1710, a MarR homolog from *Sulfolobus tokodaii* binds salicylate and carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) [64, 147]. Salicylate binding resulted in shift from 275 to 280 nm and an increase in the positive signal at 300 nm [147]. CCCP binding produced a new negative peak at 325 nm, which increased with concentration of the ligand until the binding reached saturation.

1.4.1.3 CD of Hemin in Proteins

Free heme is CD inactive because it is achiral. In an asymmetrical environment, it can show a CD signal; however, this is often small [23, 58]. Myoglobin, hemoglobin, and cytochrome c all show molar ellipticities of approximately 10^5 deg cm^2 dmol^{-1} on a per heme basis [13, 124]. One origin of the small size of the band is the fact that many proteins with *b*-type hemes bind the heme in two orientations, flipped by 180° about the α, γ axis [8, 138]. This can reverse the sign of the CD. For example, one form of horse heart myoglobin has a strong positive CD band while the other has a weak negative band [8].

Changes in the heme CD band can be used to monitor heme binding. For example, when heme was bound to human serum albumin (HSA) a negative band at 397 nm appeared [58]. The magnitude of the band increased as the concentration of heme increased until the ratio of

heme:BSA reached 1, indicating a 1:1 heme:protein complex. Heme is known to bind to HSA at tyrosine 161 [132].

1.4.2   Purification background

Because the role of transcription factors is to bind DNA, it is often difficult to purify them without DNA. Various strategies utilized to isolate a protein free of DNA include: using the effector molecule to displace the DNA, cleaving the DNA into fragments too small to bind, and increasing the ionic strength of the solution to interfere with DNA binding. In general, successful purification involves high salt concentration. Selected MarR purification protocols include [4, 19, 24, 32, 33, 43, 51, 92, 95], all of which use high salt in their purification buffers.

2   EXPERIMENT

2.1   General

2.1.1   *Instrumentation*

UV/visible spectra were taken on a Varian 50 Bio UV/visible spectrophotometer in quartz cells with a path length of 1 cm unless otherwise specified. Spectra of solutions with absorbances significantly above 1.0 on standard spectrophotometer were taken on GE Nanodrop spectrophotometer. CD spectroscopy was performed on a Jasco J-1500 circular dichroism spectrophotometer. The spectra were collected using quartz cuvettes with a 0.1 cm path length. Measurements were collected in duplicate accumulations with 0.5 nm steps from 300 to 190 nm. Protein samples were recorded in 20 mM phosphate buffer. All water was 18 mΩ from a Barnstead water purifier.

Most centrifugation was done using an Eppendorf 5804 R with fixed rotor. Centrifugation of large volume solutions was done using a JLA 8.100 rotor Beckman Coulter Avanti-T26 XPI. Centrifugation above 8,000 rpm was performed using a JA-25.50 Fixed-Angle Rotor.

The HiTrap Desalting column with Sephadex G-25 resin (5 mL), Strep-Tactin Superflow column (5 mL), HiTrap SP HP Ion exchange column (1 mL), and PD-10 Desalting column (8.3 mL) were all obtained from GE Life Sciences. Column progress was monitored by UV/visible spectroscopy. The ion exchange column was run on a GE Healthcare ÄKTA fast protein liquid chromatography (FPLC) instrument (Amersham BioSciences). An Amicon© Ultra-15 10K centrifugal unit with 10 KDa MW cutoff was used to concentrate protein solutions or exchange buffers.

2.1.2   Buffers and Chemicals

The sodium phosphate dibasic and sodium phosphate monobasic used for phosphate buffer were both obtained from Fisher Scientific. Myoglobin from equine skeletal muscle was purchased from Sigma Aldrich as lyophilized powder. DNase and RNase A from bovine pancreas were both obtained from Roche. Streptomycin sulfate [30% (wt/ vol)] solution and benzonuclease were kindly provided by laboratory of Dr. Jenny Yang.

Luria-Bertani (LB) media was prepared using 10 g/L tryptone, 10 g/L NaCl, and 5 g/L yeast extract in water. The media was sterilized in an autoclave. Once cooled to room temperature, 500 µL of 100 mM kanamycin stock solution was added for each liter of media for a final concentration of 50 µM.

Buffer A contained 100 mM Tris-HCl and 150 mM NaCl in water. Buffer B contained 100 mM Tris-HCl, 150 mM NaCl, 2.5 mM desthiobiotin in water. Buffer C contained 100 mM Tris-base, 500 mM sodium chloride in water. The pH of all three buffers was adjusted to 8.0 using HCl. Phosphate buffer (20 mM, pH 7.5) was prepared through combination of 20 mM sodium phosphate dibasic in water, and 20 mM sodium phosphate monobasic in water.

2.2   Growth and Purification

2.2.1   Sequencing for Verification

Luria-Bertani (LB) agar was prepared using 1.0 g tryptone, 1.0 g NaCl, 0.5 g yeast extract, and 1.5 g agarose in 100 mL water. This media was autoclaved on 20 min cycle. After cooling to room temperature, 50 µL of a 100 mM kanamycin stock was added. Before the mixture solidified, in a sterile environment, the media was poured into plates and left to set. Plates were sealed with parafilm and stored at 4°C.

The PefR plasmid (designed by Rizvan Uluisik) was transformed into *E. coli* BL21 DE3 cells (New England BioLabs). The protein was expressed from glycerol stock prepared by Stephanie Thompson. The stock was streaked onto an agar plate using the sterile loop technique. The plate was incubated at 37 °C overnight (18 h). A single colony was selected and used to inoculate 10 mL of LB media, which was shaken overnight at 37 °C, 220 rpm. The culture was centrifuged at 6,500 rpm, 4 °C for 25 min and the cell pellet was collected. The plasmid was isolated from the cell pellet using QIAprep Spin Miniprep Kit (Qiagen) and sequenced by Georgia State University's DNA Services. The sequence obtained was compared with the sequence provided by Stephanie Thompson of the protein construct.

2.2.2   PefR Growth and Expression

Two 2 L flasks, each containing 1 L of LB media, were prepared. Small scale growths were made from approximately 10 mL of LB media, which were inoculated with glycerol stock using the sterile flame loop technique. The starters were left in the shaker for 18 h at 37 °C, 220 rpm. The small-scale growths were used to inoculate the flasks containing the remaining liter of LB media and placed in the shaker at 37 °C, 220 rpm until OD_{600} reached between 0.6 and 0.8. The media was induced with 1 ml of 100 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) per liter of LB media. The shaking was continued for 18 h at 27 °C, 220 rpm. After induction, the cells were harvested by centrifugation at 8000 rpm, 4 °C for 30 min.

2.2.3   Protein Purification Protocols

2.2.3.1   Purification Protocols from Cells

The cells were grown in in 1 L LB media as described above. The harvested cell pellet was divided into four falcon tubes, each containing between 1.5 and 2.2 g of cell pellet, and stored at -80 °C. The volume of lysis buffer used in each trial was approximately 10 mL of

buffer per 1 g of cell pellet. All steps in the following protocols were performed on ice unless otherwise noted. SDS-PAGE was used to assess the purity of the protein for all samples. The UV/visible absorption spectrum of all five protein samples was recorded using the Nanodrop spectrophotometer.

Protocol 1, the control, was performed on the same day the cells were collected. A 1.5 g cell pellet was lysed in 15 mL of lysis buffer [10 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.2 mg/ml lysozyme, 5 µg/mL each of DNase I and RNase A in buffer A] for 30 min on ice. The lysate was sonicated on ice at an amplitude of 19%, 10 s pulse on, 10 s pulse off, for 30 min. The lysate was centrifuged at 7000 rpm, 4 °C for 45 min and the supernatant was stored on ice. Purification of the supernatant was done using a Strep-Tactin column equilibrated with buffer A. The sample was loaded onto the column and the unbound material was washed out with buffer A. A 50:50 mixture of buffer A:buffer B was used to elute protein. After elution, a centrifugal filter unit was used to concentrate and exchange PefR into buffer A. The protein was stored at -20 °C.

Protocol 2 used a higher salt content through the addition of NaCl salt to the lysis buffer. The 2.2 g cell pellet was thawed, suspended in 20 ml of lysis buffer (500mM NaCl, 10 mM $MgCl_2$, 1 mM PMSF, 0.2 mg/ml lysozyme, 5 µg/mL each of DNase I and RNase A in buffer A) lysed for 30 min, and sonicated on ice using the same settings as Protocol 1. The sonicated cells were centrifuged similarly to Protocol 1. The supernatant was split into two fractions. One fraction (2b) was stored at -20 °C, the other (2a) was immediately purified by the Strep-Tactin column to wash out unbound material, and eluted with a 50:50 mixture of buffer C:buffer B. After elution, a centrifugal filter unit was used to concentrate and exchange PefR into buffer A. The protein was stored at -20 °C. The next day, fraction 2b was purified by the Strep-Tactin

column following Protocol 1 using buffer A. After elution, a centrifugal filter unit was used to concentrate and exchange PefR into buffer A. The protein was stored at -20 °C.

Protocol 3 used a doubled concentration of DNase in the lysis buffer. The 1.5 g cell pellet was thawed, and suspended in 15 ml of lysis buffer (500 mM NaCl, 10 mM MgCl$_2$, 1 mM PMSF, 0.2 mg/ml lysozyme, 10 µg/mL each of DNase I and 5 µg/mL RNase A in buffer A), lysed for 30 min, and sonicated on ice using the same settings as Protocol 1. The sonicated cells were centrifuged similarly to Protocol 1 and the supernatant was stored at -20 °C. The supernatant was thawed on ice the next day and purified on the Strep-Tactin column, using the procedure from Protocol 1. After elution, a centrifugal filter unit was used to concentrate and exchange PefR into buffer A. The protein was stored at -20 °C.

Protocol 4 started with a 2.2 g cell pellet which was mixed with 20 ml of lysis buffer (500mM NaCl, 10 mM MgCl$_2$, 1 mM PMSF, 0.2 mg/ml lysozyme, 1 µL benzonuclease). The cells were lysed, sonicated and centrifuged in a manner consistent with previous Protocols. The supernatant was transferred to a falcon tube, which was placed in boiling water for 3 min, and centrifuged for 30 min at 6500 rpm, 4 °C. The supernatant was centrifuged again for 30 min at 6500 rpm, 4 °C. The supernatant, 13 ml total, was collected and 1.3 mL of 30% (wt/vol) streptomycin sulfate was added for a final concentration of 3%; the sample stored overnight at 4 °C. The sample was centrifuged for 30 min at 6500 rpm and 4 °C. The supernatant was purified using the Strep-Tactin column and exchanged into buffer A following Protocol 1.

2.2.3.2   DNA Removal from Purified Proteins

UV/visible spectroscopy was used to monitor the effectiveness of each step of the procedures. All solutions were prepared in buffer A unless otherwise noted.

The first experiment attempted to displace DNA by heme binding to the PefR. A 1900 µl solution of WT PefR, (31 µm, calculated using the extinction coefficient of the construct calculated in ExPASy, $\varepsilon_{280 \text{ nm}} = 14{,}440 \text{ M}^{-1} \text{cm}^{-1}$) was prepared in the cuvette, along with a 4.4 mM solution of hemin chloride in DMSO. The concentration of hemin was calculated by 100-fold dilution of stock in DMSO ($\varepsilon_{624 \text{nm}} = 6{,}230 \text{ M}^{-1} \text{cm}^{-1}$) [77]. Titration was performed by comparing the spectra of two cuvettes with equal buffer volumes, one with buffer A and one with a solution of PefR in buffer A. Using a glass syringe, a 15 µl aliquot of the heme solution was added to each cuvette and solutions were allowed to mix on ice for 15 min before collecting the absorption spectra. A second aliquot of 5 µl heme in DMSO was added to each. The solutions were allowed to stand overnight at 4 °C. The hemin solution spectra were manually subtracted from PefR/hemin spectra using Excel. A PD-10 column was equilibrated with buffer A. The PefR solution with heme (1 mL) was run down the PD-10 column to remove excess heme as well as free DNA. The column was equilibrated again with buffer containing doubled salt (100 mM Tris-base, 300 mM sodium chloride, pH 8.0), and the remaining approximately 1 ml PefR was passed down the column using this buffer. The eluents were stored at -20 °C.

The second experiment utilized DNase in an attempt to cleave DNA in solution. Approximately 3 µL DNase (1 U/µL) was added to 1 mL of holo-PefR (32 µM, assuming $\varepsilon_{280 \text{ nm}} = 14{,}440 \text{ M}^{-1} \text{cm}^{-1}$). Buffer A with the equivalent amount of DNase was used as the baseline, and three absorption spectra were taken at room temperature at 5 min intervals. The sample was placed on ice, and spectral monitoring at 5 min increments continued. After 30 min, the sample was loaded onto the Strep-Tactin column, buffer A was used to wash away unbound material, and the protein was eluted using 50:50 buffer A:buffer B. The samples were stored at -20 °C.

The third experiment involved apo-PefR purification using a HiTrap SP HP Ion exchange column, run using ÄKTA. The ion exchange column was equilibrated with buffer A, PefR was loaded onto the column, and unbound material was washed out with buffer A. PefR was eluted using a stepwise gradient of 100 mM Tris buffer containing from 150 mM to 2 M NaCl. Buffer mix steps were 5% from 0 to 30%, then a 20% increase to 50%, then two 25% step increases until 100% 2 M NaCl buffer was reached. Both the wash and the elute were concentrated using centrifugal units, and the absorption spectrum of both was collected using the Nanodrop spectrophotometer.

2.2.3.3   *Final Purification Protocol*

In view of the results from the experiments above, a final purification protocol was established. PefR was purified from a cell pellet that was collected as described above. The cell pellet was lysed in 60 mL of a buffer A solution containing in addition 500 mM NaCl, 10 mM $MgCl_2$, 1 mM PMSF, 0.2 mg/ml lysozyme, and 5 µg/mL each of DNase I and RNase A in buffer A, on ice for a minimum of 30 min. The lysate was sonicated on ice using the same settings as Protocol 1, and the cell debris was removed by centrifugation at 17,000 rpm at 4 °C for 20 min. The supernatant was purified using the Strep-Tactin column rinsed with buffer A, and PefR was eluted using 50:50 buffer A:buffer B. Using the ÄKTA, this protein was loaded onto the HiTrap SP HP ion exchange column, rinsed with buffer A and eluted using 100 mM Tris buffer containing 0.5 M NaCl. A centrifugal filter unit was used to remove the excess salt and exchange the protein into buffer A.

2.3   UV/Visible Spectroscopy

2.3.1   Heme Titration of PefR

A 2000 µl solution of PefR (14 µM in buffer A) was prepared in a cuvette, along with a 2.5 mM solution of hemin chloride in DMSO. The concentration of hemin was calculated by 1,000-fold dilution of stock into DMSO ($\varepsilon_{404\ nm}$ = 174,000 $M^{-1}\ cm^{-1}$) [16]. The titration was done using a two-cuvette system, one cuvette containing 2000 µl of buffer A, the other containing the 2000 µl of PefR solution. Using a glass syringe, 3 µl of hemin solution was added to each cuvette which was allowed to stir on ice for 15 min before collecting the absorption spectra of both. This was done until 2.0 molar eq of hemin was added. Both solutions were incubated overnight at 4 °C, and the spectra were collected again the next morning. The buffer solution spectra were manually subtracted from PefR/hemin spectra using Excel. Excess hemin in solution was removed using the HiTrap Desalting column, run with buffer A.

2.3.2   Carbon Monoxide Binding Study

Carbon monoxide was gently bubbled into holo-PefR in buffer A in a cuvette with a septum top for several min. After the spectrum was collected, holo-PefR was reduced using 3 µl of a saturated sodium dithionite solution in buffer A, and the spectrum was collected once again.

2.3.3   Myoglobin Phosphine Binding Studies

2.3.3.1   Ferric Myoglobin with TCEP

An 18.1 mM solution of tris(2-carboxyethyl)phosphine (TCEP) in buffer A was prepared by dissolving 0.0052 g TCEP in 1000 ul buffer A. Using a glass syringe, 50 µl of the TCEP solution was added to a cuvette containing 900 µl of 1.7 µM myoglobin solution ($\varepsilon_{502\ nm}$ = 3900 $M^{-1}\ cm^{-1}$) [6] in buffer A. The absorption spectra were collected over 10 min. A second 100 µl

aliquot of TCEP solution was added, to reach a total of 2.6 mM TCEP. The spectra were

collected again over 15 min, with a third spectrum collected 90 min later.

2.3.3.2   Ferric Myoglobin with PMe$_3$

Using a glass syringe, 1 μl of 1 M PMe$_3$ in THF was added to 1000 μl of a 22 μM

myoglobin solution ($\varepsilon_{409\ nm}$ = 188000 M^{-1} cm^{-1}) [6] in a cuvette, and the absorption spectrum was

collected. This was repeated nine times until a total of over 400 eq of PMe$_3$ had been added.

2.3.3.3   Ferrous Myoglobin with PMe$_3$

A 40 mM PMe$_3$ solution was prepared by diluting 40 μl of 1 M PMe$_3$ in THF into 1960

μl of buffer A. A cuvette containing 900 μl of a 22 μM myoglobin solution was reduced using 3

μl of a saturated solution of sodium dithionite in buffer A, and the spectrum was collected. Using

a glass syringe, a 50 μl aliquot of the 40 mM PMe$_3$ solution was added to the reduced

myoglobin; the spectrum was collected immediately after addition, as well as several min later.

Two 10 μl aliquots of the 1 M PMe$_3$ in THF were added, collecting spectra after each addition as

well as several minutes later.

2.4   Stability Studies Using CD

Freshly synthesized PefR was exchanged into 20 mM sodium phosphate buffer, using

centrifugal filter units. All protein was stored in 1.5 mL microcentrifuge tubes. Any time the

protein samples were not at 4 °C or -20 °C, they were kept on ice. For all samples, the

background spectrum of the buffer and cuvette was subtracted from the spectrum of buffer and

protein, using Excel.

Table 5 timeline gives the timeline of various storage conditions for samples A-H.

Sample G, in buffer A, was stored at 4 °C for 6 d after purification, -20 °C for 10 d, and 3 d at 4

°C. The sample was heme loaded as described above. Approximately 14 days later, a 1 ml

aliquot of the sample was exchanged into phosphate buffer using a centrifugal filtration unit.

Secondary structure composition was calculated using the analysis K2d Program from the

website DichroWeb [134, 135].

2.5   Bioinformatics and Modeling

Structural figures were produced using UCSF Chimera 1.14 [87]. The I-TASSER suite

[144, 148] was used to create a homology model based on the sequence provided by the

Eichenbaum lab. The sequence of PefR was submitted to the SPRING server [42] for both

monomers to create a model homodimer. A model of the predicted homodimer was created from

two copies of the monomer produced by the I-TASSER suite, using the Matchmaker function of

Chimera, with the SPRING model used as the reference structure. The distance between atoms

was measured using Chimera.

Sequence alignment of PefR was done using Clustal Omega and Mview. PefR sequence

was compared with Sco3205 from *Streptomyces coelicolor* (3ZPL) and MepR from

Staphylococcus aureus (4LLN).

The secondary structure composition was also calculated from CD sample E using the

analysis program K2D from the website DichroWeb [134, 135].

3   RESULTS AND DISCUSSION

3.1   Purification and Sequencing

The first milestone in this work was to purify PefR, from previous stocks in this laboratory [127]. Sequencing was performed to confirm that the desired sequence had been maintained and that no mutations had occurred during storage. Figure 2 shows that the sequence obtained matches the desired sequence.

The purified protein construct migrated as a single band with an apparent molecular mass of 20 kDa (Figure 3), which is in good agreement with the calculated molecular mass of the construct of 20.2 kDa.

A number of attempts were made to obtain PefR solutions that did not contain DNA. Four protocols were tested directly from cells, and three experiments were run on purified protein. Nucleic acids absorb at 260 nm, and proteins absorb at 280 nm. The ratio of absorbance at these two wavelengths is typically used to determine the purity of a sample. The 260/280 nm ratio seen for all these samples are shown in Table 6. The spectra of the final samples in each instance are shown in Figures 4 and 5, respectively.

3.1.1   Purification from Cells

A control (PefR purified by Protocol 1) showed a broad peak at 260 nm that overshadowed the peak at 280 nm. As the protein had been purified using a Strep-Tactin column, this peak was assumed to be DNA that remained bound to the protein. The 260/280 nm ratio was 1.5.

As discussed in the Introduction, a high salt content in the lysis buffer is often used to purify MarR transcription factors. Protocol 2a lysis buffer contained an additional 500 mM NaCl, and PefR was eluted using a buffer containing 500 mM NaCl rather than 150 mM NaCl.

This sample showed the sharpest 280 nm peak, and the lowest 260/280 nm ratio (0.6) of the protocols used to purify from cells. PefR purified by Protocol 2b, which had the additional 500 mM NaCl in the lysis buffer, but was eluted using buffer A, showed a peak between 260 and 280 nm, which was narrower than the one seen from Protocol 1, with the second-lowest 260/280 nm ratio of 0.9.

If the DNA in solution were cleaved into small enough fragments, PefR would be unable to bind them. Protocol 3, which had the same salt concentration as Protocol 2, as well as a doubled concentration of DNase in the lysis buffer, produced PefR that showed a peak at 280 nm which was slightly broader than the one seen from Protocol 2a and similar to that from Protocol 2b. Protocol 3 had a 260/280 nm ratio of 0.9. The observation that Protocols 2b and 3 have the same 260/280 nm ratio may be due to the salt content of the lysis buffer, reducing the activity of the enzyme. DNase activity has been shown to be inversely correlated with NaCl concentration [103]. Increasing concentration of NaCl or KCl from 0 to 30 mM decreased the activity of DNase I more than two-fold [126], meaning that the activity of the enzyme in the presence of 500 mM NaCl, would have been significantly decreased.

The fourth Protocol utilized the dual rationale that streptomycin sulfate can precipitate DNA [68] and that benzonuclease is more effective than DNase [74]. This protocol was based on a procedure reported by Salarian [97]. After centrifugation, the sample was placed in boiling water in an attempt to denature the protein and nucleic acids, resulting in a loss of binding. A DNA pellet was seen at the bottom of the tube after overnight incubation with streptomycin sulfate, showing that some DNA had precipitated. Unfortunately, loss of sample due to a cracked centrifuge tube made seeing the protein peak more difficult. Protocol 4 showed the highest 260/280 nm ratio of 1.4.

3.1.2   Purification from Protein

Several techniques were attempted to remove DNA from PefR that had been purified

initially using Protocol 1. Spectra of each of these steps are shown in Figure 5.

The first step was to load PefR with heme in an attempt to displace the DNA with its

effector molecule. The concentration of PefR was assumed to be 31 µM, calculated using the

extinction coefficient of the construct at 280 nm. This concentration is an overestimation, as

nucleic acids absorb significantly at 280 nm. As the actual concentration of PefR was lower than

that calculated, more than 1.4 molar equivalents of heme were added. The spectrum of the

reconstituted protein showed a peak at 390 nm, consistent with the previous spectra of holo-

PefR. However, a broad peak at 260 nm was still seen, with a 260/280 nm ratio of 1.3.

It was unclear if the absorbance was due to DNA bound to PefR or if the heme binding

had displaced the DNA, which was now free in solution. A PD-10 desalting column was used in

an attempt to remove DNA from the solution. Half of the sample was run in buffer A (Figure 5)

and the other half in 100 mM Tris 300 mM NaCl (data not shown). There was no change in the

260/280 nm ratio in either case. The Soret narrowed after the PefR had been through the PD-10

column, perhaps due to loss of unbound heme in solution.

For the second experiment, a catalytic amount of DNase was added to holo-PefR solution

on ice in an attempt to cleave any unbound DNA in the solution into smaller pieces to allow for

easier removal. Very little change was seen spectroscopically upon DNase addition, and there

was no change in the 260/280 nm ratio. Several spectra were taken over a 30 min period, with

minimal discernable differences between the spectra. The reaction continued at room

temperature, as many enzyme activity levels are temperature dependent. The 260/280 nm ratio

remained at 1.3. The relatively high sodium content of the buffer (150 mM NaCl) may have

decreased the activity of the DNase. Running this sample down the Strep-Tactin column resulted in a complete loss of heme and an increase in 260/280 nm ratio from 1.3 to 1.5, showing it was ineffective in removing the DNA. This increase in 260/280 nm ratio upon heme loss is attributed to the absorbance of heme at 280 nm. Further studies could be done with benzonuclease, which cleaves both DNA and RNA and shows no proteolytic activity [74]. Benzonuclease activity is inversely correlated with Na^+/K^+ concentration, similar to DNase; however, benzonuclease still shows a fairly high level of activity at 150 mM NaCl [73]

The third experiment, running apo-PefR through an ion-exchange column, proved the most successful method. When eluting using a stepwise gradient of 100 mM Tris buffer containing from 150 mM to 2 M NaCl, PefR eluted with 100 mM Tris buffer containing 0.5 M NaCl. The protein peak narrowed and shifted from 260 to 280 nm. The 260/280 nm ratio decreased to 0.5, less than half the ratio seen with other steps taken to remove the DNA from purified PefR. The absorbance of the flow-through from the column had a 260/280 nm ratio of 1.9.

In light of the experiments above, the most effective ways to remove DNA from PefR were to increase the ionic strength of the buffers and to utilize an ion-exchange column. The final protocol that was utilized to purify PefR after completion of these trials involves the addition of 500 mM NaCl to the lysis buffer, elution from the Strep-Tactin column with buffer A, then the utilization of an ion-exchange column. The addition of higher salt to the lysis buffer, as well as the elution buffer used with the Strep-Tactin column, produced the cleanest protein of all the protocols used to purify PefR from cells. However, the protein purified using the ion exchange column showed the lowest 260/280 nm ratio of all the techniques used to purify PefR. (Table 6) PefR elutes from the ion exchange column at 0.5 M NaCl. Therefore, in order to utilize

the ion exchange column and elute from the Strep-Tactin column with high salt buffer, PefR would have to be exchanged from its elution buffer back into buffer A (150 mM NaCl), and then run down the ion exchange column, exchanging it into 0.5 M NaCl, which would then need to be exchanged into buffer A or phosphate buffer for further experiments. We felt this much fluctuation in the environment of PefR posed too great of a risk of denaturing the protein.

3.2   UV/visible Spectroscopic Studies

3.2.1   Heme Titration of PefR

The extinction coefficient of PefR protein at 280 nm (7,450 M^{-1} cm^{-1}) is fairly low, due to the lack of intrinsic Trp. The Strep-tag construct ($\varepsilon_{280\ nm}$ = 14,440 M^{-1} cm^{-1}), does contain one Trp in the tag, however. The Soret of ferric PefR was seen at 382 nm with a visible band around 510 nm with a charge transfer bands at 615 nm. This spectrum is consistent with the absorption spectra seen with other of five-coordinate thiolate-binding proteins (Table 2).

The titration of PefR with hemin in DMSO can be seen in Figure 6. This titration was done using a two-cuvette system, with equal amounts of hemin being added to both the protein and buffer solutions.

Without subtracting the absorbance of the hemin solution, the absorbance at the Soret increases linearly as a function of the ratio of heme:PefR. This is to be expected, as the PefR Soret overlaps with the absorbance peaks of free heme. When the spectrum of the hemin solution is subtracted, the absorbance at the Soret of the PefR solution increases until a break point can be seen at a heme:PefR ratio of one (Figure 7). After this point, the absorbance at the Soret decreases. For a 1:1 complex, one would expect no decrease, after adding equimolar heme. The small decrease may be due to the spectra having absorbances greater than 1.

3.2.2 Carbon Monoxide Binding Study of PefR

A solution of the holoprotein was saturated with carbon monoxide. The reduction of hemin via the addition of an aliquot of saturated sodium dithionite solution caused a red-shift of the Soret from 383 nm to 420 nm (Figure 8). The α and β bands sharpened to 569 and 539 nm, respectively. A red-shift of the Soret upon reduction to 450 nm, similar to the one seen with cytochrome P450$_{cam}$, is characteristic of a low spin cysteine-bound heme [25, 29, 89, 96] (Table 3). A shift to 420 nm is associated with neutral ligand opposite CO. Typically assumed to be His, it been shown that a Soret at 420 nm can also be indicative of neutral thiol.

Early examples of neutral thiols bound to ferrous CO heme were found in model systems [86]. Ferrous H93G titrated with cyclopentanethiol showed several isosbestic points, indicating a clean titration. The location of the Soret (428 nm), in comparison to the apo-protein, and known thiolate proteins, along with the lack of change seen in the absorption and MCD spectra between pH 7 and 10.5, were all indicators that the thiol was protonated. The MCD of this complex was similar to other ferrous five-coordinate heme bound to neutral ligands, with less intense peaks in the 500-700 nm range, and the addition of symmetric band in 300-500 nm range. MCD studies confirmed that the CO bound to ferrous heme *trans* to a neutral ligand. This evidence shows that neutral thiols are plausible axial ligands, and the Soret that results from the addition of CO to ferrous heme bound to a neutral thiol appears at ~420 nm [86]. Neutral thiol is also thought to be the axial ligand the ferrous H175C/D235L mutant of cytochrome c peroxidase [109].

Nitrophorin is another example of a protein that may have a neutral cysteine ligand. As described above, the crystal structure of nitrophorin shows that upon reduction, the heme appears to be ligated by a Cys [133], and CO binding to nitrophorin causes the Soret to shift to 422 nm [130]. This has been interpreted as a neutral thiol opposite the CO [114].

Ferrous mammalian DGCR8 binds heme via two Cys [10, 11]. It was initially assumed ferrous heme dissociated from DGCR8 or switched ligation to another residue. However, Girvan et al. suggests that DGCR8 maintains its bis-Cys ligation upon reduction [37]. The absorption spectra of reduced DGCR8 spectra was consistent with a bis-thiol spectrum seen by Perera [37, 86]. This, along with Raman, MCD, and ENDOR evidence has been taken to indicate that the ferrous heme is coordinated to two identical neutral ligands, that are not water or His. This evidence, in addition to previous work showing that the sixth axial ligand is not a methionine [11], lead Girvan et al. to conclude that the ferrous heme is ligated by two neutral Cys. Girvan suggests based on CO-back bonding correlation that CO is bound to a neutral thiol. The Soret of CO bound DGCR8 is seen at 421 nm [37].

These studies all show that ferrous heme is able to stably bind a neutral thiol and that the Soret produced by CO binding heme ligated to a neutral thiol appears at 420 nm. Therefore, in the case of PefR we can conclude that ferrous heme does not bind a thiolate but rather a neutral ligand. Further studies are needed to confirm if heme binds neutral Cys, His, or another neutral amino acid.

3.2.3 Myoglobin Phosphine Studies

As noted above, Gao et al. used TCEP in the titration of Dps (1000:1 eq of TCEP:Dps) and saw clear spectral shifts [35]. In our laboratory, addition of TCEP to ferric horse heart myoglobin, even in a ratio of over 1500:1 eq, did not result in any spectral changes, as shown in figure 9. This may be due to the relatively large size of TCEP. PMe_3 [17] and bis(hydroxymethyl)methylphosphine [117] both can bind to myoglobin. TCEP is a larger molecule than these two more commonly utilized phosphines. The heme binding site in myoglobin may not be large enough for the TCEP to interact with the bound heme.

A scouting titration of 1.0 M PMe$_3$ in THF into ferric myoglobin caused the Soret to shift from 409 to 425 nm and split, with a band at 370 nm appearing (Figure 10). The visible bands increased in amplitude and shifted to 535 nm (Figure 11). We note there was an abrupt change in the spectrum upon the third addition of PMe$_3$ into solution. We attribute this to the fact that the PMe$_3$ in THF solution is not fully miscible with buffer A. Further studies should employ PMe$_3$ in buffer. Our data on horse heart myoglobin are very similar to those seen with ferric sperm whale myoglobin, for which the PMe$_3$ complex showed peaks at 370, 424 and 536 nm [112].

Addition of PMe$_3$ (40 mM PMe$_3$ in 1:49 THF:buffer A, followed by 1.0 M PMe$_3$ in THF) caused the Soret of ferrous myoglobin (Figure 12) to narrow, increase and shift from 432 to 437 nm (accuracy of peak positions perhaps diminished by high concentration of myoglobin). In the visible region, the single broad band converted to two sharper bands at 534 and 568 nm, both of which have slight shoulders (Figure 13). This is consistent with findings by Brunel et al., who saw peaks at 430 and 436 nm, as well as 535 and 568 nm with PMe$_3$ titration into ferrous myoglobin [17]. We note that as in the previous experiment, the PMe$_3$ in THF is not miscible with the buffer; the reaction should be rerun with neat PMe$_3$ or PMe$_3$ in buffer.

3.3   PefR Stability Under Different Storage Conditions.

A preliminary attempt to reconstitute PefR with heme after storage for several months at -20 °C showed minimal binding. Therefore, extensive studies were done to see the effect of storage time and temperature on the stability of PefR. The stability of PefR was studied by monitoring the secondary structure using CD. Spectra were collected over a six-week period of PefR stored at either 4 °C or -20 °C (Table 5).

No differences are seen in the shape of these spectra for all of these samples (Figure 14). Therefore, we can conclude that PefR retains its structure for at least three weeks at 4 °C.

Storage for approximately two weeks at -20 °C also did not significantly affect the structure. Further studies over longer time periods are needed to determine a more accurate timeline for PefR stability.

The spectra collected for holo-PefR does appear to be taller than all the other samples at ~190 nm, a peak associated with the α-helical content of the protein, which could be taken to indicate conformational shifts upon heme binding.

Some discrepancies are seen between the samples below 200 nm, even after normalization. This is most likely due to high tension (HT) values, which lead to high signal/noise ratios. HT values in the far UV-range tend to be high as less light is available at these lower wavelengths. Further studies with lower concentrations of PefR would be needed to make any definitive statements, however.

3.4   Structural Studies

3.4.1   Sequence Alignment

I-TASSER, using the COACH server [141, 142], reports the predicted function/ligand binding sites, along with the proteins used to make those predictions. To investigate the structure of the protein bound to DNA, the sequence alignment of PefR was compared with two reported protein/DNA complexes with the highest confidence scores (C-score), MepR from *Staphylococcus aureus* (4LLN, C-score of 0.41) [12] and Sco3205 from *Streptomyces coelicolor* (3ZPL, C-score of 0.25) [122] (Figure 15).

3.4.1.1   Aligned Residues Involved with DNA Interactions

Overall, 13 amino acids are conserved between all three proteins. Of those, several were explicitly noted in the literature. Three conserved residues are found in the wing region of the wHTH, which interacts with the minor groove of DNA. For MepR these are Ser76, Asp96,

Arg98; in Sco3205 these are Ser65, Asp85, and Arg87 [12, 122]. All three residues are noted for their role in interacting with specific thymines and adenines in the TA-rich promotor regions. This Arg (98/88/87) (MepR/PefR/Sco3205) is also conserved in several other MarR homologs [122].

Some residues that were conserved between all three were only specifically noted in the literature for either MepR or Sco3205. This includes a second Arg, (Arg10 in MepR) which is also conserved between all three proteins. This residue has been said to contribute to protein/protein contacts that play a role in the high DNA binding affinity of MepR [12]. Its mutation resulted in an ~10-fold decrease in DNA-binding affinity.

3.4.1.2 *Residues Aligning with the Proposed Axial Ligand*

We have proposed that heme binds PefR via a Cys. PefR contains only one Cys per monomer, Cys109. PefR Cys109 aligns with Ile107 in MepR and Gly118 in Sco3205 (Figure 16). Ile107 in MepR and Gly118 in Sco3205 are each one residue away from amino acids specifically noted for their role in ligand interactions, Phe108 in MepR [12] and His119 in Sco3205 [122]. Consistent with PefR, all of these residues are in α5 helix, facing towards the α1' helix.

In Sco3205, His119 is proposed to hydrogen bond to the carboxylate group of its salicylate ligand [122]. It is conserved in several salicylic acid-binding MarR regulators [122], including MarR from *Methanobacterium thermoautotrophicum* (3BPX), which I-TASSER listed as the structural analog with the highest similarity to PefR.

For MepR, Phe108, His35, and Arg10' (which belong in the α5, α2, and α1' helices, respectively) stack in the holo form of the protein, presumably facilitating DNA binding [12]. MepR typically binds cationic lipophilic substrates with delocalized π-electron systems; this

substrate binding disrupts the interaction of these three residues. This region also undergoes conformational changes upon DNA binding that involve the formation of a kink in the α5, and shift in the position of the α1 helix. These changes point to the possible importance of structural changes in this region, which, in PefR, would/could occur if heme bound to Cys109. PefR has been shown to release DNA upon heme binding [32, 95]. The DNA binding ability of several MarR proteins has been shown to be disturbed, and often lost as a result of conformational shifts in α5.

3.4.2   Homology Modeling

3.4.2.1   Overall Structure

The I-TASSER program predicted five possible structures for PefR; the 3D model with the highest C-score (0.24) was used for all further studies (Figure 17). This model was 67% α-helix, 7% β-sheet and 25% coil. This is fairly consistent with the secondary structure content calculated using K2D algorithm of Dichroweb [134] which was predicted to be 66% α-helix, 5% β-sheet and 29% coil.

MarR proteins are typically homodimers. The SPRING algorithm, used for protein-protein docking predictions, was used to predict how PefR dimerizes. This produced ten possible model homodimers. The top-ranking spaghetti model dimer was used as a reference structure to create a model of the PefR dimer using the I-TASSER monomer.

This predicted structure of PefR is shown in Figure 17. Each subunit contains six α-helices, and two β-strands (Figure 18). PefR is triangularly shaped, and a channel can also be seen running through the center of the dimer.

3.4.2.2   wHTH Motif

A wHTH motif is located at the base of the protein, characteristic of MarR proteins (Figure 18). The wHTH motif is typically composed of helices α2-α4, with a β-hairpin containing β1 and β2 [27, 44, 85, 136]. PefR also contains a wHTH region found at the bottom of the protein (Figure 19). This wHTH is made up of α2, α3 α4, β1, and β2 and is strongly electropositive, which is consistent with its presumptive task of DNA binding.

3.4.2.3   Proposed Heme Binding Cleft

Cys109, the presumed axial ligand, is found in the α5 helix, with the residue pointed toward the α1 helix. The α2 helix sits perpendicular to the α1 and α5 helices. A cleft formed from theses helices can be seen in Figures 19 and 20. We propose the heme binds in this cleft, which results in a loss of DNA binding, as PefR has been shown to release DNA upon heme binding [32, 95]. In line with this assumption, several MarR proteins have been shown to bind their respective ligands in a cleft made up from α1, α2, and α5 [12, 20, 24, 43, 62, 82, 98, 122, 150].

This region is particularly well suited to conveying information throughout the protein, as the α5 helix connects the dimerization region to the DNA binding region (wHTH) in many MarR proteins [27]. The flexibility of this region has shown to be critical in the ability of MarR protein to bind DNA. In the case of OhrR from *Xanthomonas campestris*, oxidation of a key Cys in α5 helix causes a conformational change, which results in a disulfide bond between cysteine in α1 and α5, forcing OhrR into a rigid structure that prohibits DNA binding [79]. Ligand binding to, or oxidation of, this region has been shown to cause conformational changes in several proteins that adversely affect the DNA binding abilities of the protein [12, 20, 24, 27, 62, 79, 98, 139, 150].

This cleft itself contains several hydrophobic, positively-charged residues (Figure 21). These may play a role in interacting with the heme propionates. Bioinformatic studies of heme-binding proteins have highlighted the role of hydrophobic, positively charged amino acids such as Arg, Lys, or His in anchoring the heme propionates [67, 101, 115]. In a study of peptides synthesized to represent the binding regions of proteins that use Cys as an axial ligand, a majority contained at least one positively charged amino acid on either side of the Cys. Of these, half also contained an aromatic, hydrophobic amino acid two amino acids away, thought to stabilize the porphyrin ring through van der Waals interactions [61]. Cys109 of PefR has three positively charged residues, Arg107, Arg110 and Lys111, within this region. These residues can be seen around the exterior of the cleft, creating an environment that would be beneficial in stabilizing heme.

The presence of this cleft near Cys109, which has an environment theoretically conducive to heme binding; and this clefts location in region that has been seen as a common ligand-binding site, which is key in structural rearrangement that results in dissociation from DNA, is further evidence to support the hypothesis that Cys109 binds heme, especially when paired with the fact that DNA binding to PefR results in PefR dissociating from DNA.

3.4.2.4   Other Possible Axial Ligands

CO bound to ferrous PefR shows a Soret at 420 nm. While this may indicate a neutral thiol, as described above, it can also indicate the heme binding to a His or other neutral ligand. PefR contains five His residues (His13, His30, His45, His68, and His115). The distances between the nitrogen of His and the sulfur of Cys, are 19.8, 16.3, 23.2, 28.2, 14.6 Å, respectively, according to the homology model (Figure 22). Looking at the homology model, it is clear that none of them are in a position to bind to the heme without a large conformational shift.

However, ferrous HO-2, Rev-erβ, and HRI all bind heme hexacoordinately, utilizing a His outside of the binding pocket [34], so no definitive statement can be made.

We note that His13 of PefR aligns with His14 of MepR and His68 aligns with His77 of Sco3205 both of which have been noted for their role in DNA binding to the minor groove. This decreases the likelihood that these two residues are involved in heme binding.

4   CONCLUSIONS

The goals of this study were to obtain pure PefR and to determine the axial ligand that PefR uses to bind heme.

In purification, PefR requires extra consideration, like many transcriptional regulators, in order to obtain DNA free protein. The presence of high levels of NaCl in both the lysis and the elution buffers proved to be the most effective method when purifying PefR from cells. Running purified protein through an ion-exchange column showed to be the most effective method for removing DNA that bound to purified PefR.

Previous work showed that *S. pyogenes* PefR binds heme [95]. That study used the His-tagged protein. We have used the Strep-tagged protein, for which the holoprotein spectrum is different [127]. While Strep-tags are in general viewed as not interacting with the heme, we note that there are exceptions [35]. It may be necessary to move to tagless proteins for further studies.

The UV/visible spectra of holo-PefR, which shows a Soret at 382 nm, is characteristic of a five-coordinate heme bound via a thiolate. This thiolate would be is Cys109, as there is only one Cys found in PefR. A CO-binding study showed that ferrous heme is ligated to a neutral ligand; further studies are needed to confirm whether this neutral ligand is protonated Cys or another residue.

CD studies suggest that heme is stable for three weeks at 4 °C and that additional storage for approximately two weeks at -20 °C did not significantly affect the structure.

Bioinformatics studies support Cys109 acting as the axial ligand. Cys109 aligns with residues that are next to residues involved with DNA binding in structurally similar MarR proteins. A homology model of PefR also suggests that Cys109 is in a ligand-binding cleft with

an environment conducive to heme-binding. This cleft is in a region that several MarR proteins also use to bind their respective axial ligands.

5   FURTHER STUDIES

Future studies are needed to confirm the heme-binding axial ligand. Phosphine binding studies could be used to show that the heme in PefR binds via Cys109. As described above, ferric thiolate heme proteins show a hyperporphyrin spectrum when titrated with phosphines, with peaks around 375 nm and 445 nm [35, 69, 93, 111, 117]. Ferrous thiolate heme proteins show bands at 457 nm [117]. Based on the results of studies done in this work with myoglobin, this titration should utilize PMe_3 with no additives.

Mutational studies of PefR could provide valuable information about the role of Cys in heme binding. For example, the C29A mutant of *bj*Irr [56] lost the capacity to bind heme at its heme-binding site. This, along with UV/visible, EPR, and Raman evidence was used to support the idea that *bj*Irr binds heme via Cys29. In the case of PefR, if Cys109 were mutated and the ability of PefR to bind heme were lost, this would also provide evidence that Cys109 is the axial ligand.

Raman spectroscopy can provide significant information about the environment of heme bound to a protein. The high-frequency regions give information about oxidation (v_4) and spin state ($v_2,$ $v_3,$ v_{10}) [18, 120, 123]. Type-1 thiolate proteins show v_4 typically around 1341-1348 cm^{-1}, at a lower energy than other high-spin, Fe^{2+} heme-centers [114]. v_4 is seen in analogous centers with Fe^{2+} heme coordinated via a single His between 1355-1360 cm^{-1}. v_3 appears at 1466, 1470-1475, or 1485-1490 cm^{-1} respectively, for reduced type-1 thiolates, His-coordinated high-spin, and low-spin heme proteins with neutral ligands [114].

The low-frequency region is particularly useful for ligand identification as it is where the iron-sulfur stretching vibration appears. First identified in P450$_{cam}$ [18], the Fe-S band appears in Fe^{3+} thiolates around 340 cm^{-1} for the high-spin type-1 proteins, while the low-spin type-2 proteins show this band around 310 cm^{-1} [114]. Isotopic studies are often done to confirm that this is an Fe-S vibration.

As an example of this type of approach, *bj*Irr showed a Raman band at 333 cm^{-1}, evidence that this protein binds heme via a Cys [56]. The ^{54}Fe-substituted porphyrin showed a reproducible isotopic shift of this band to 336 cm^{-1}.

In nitric oxide synthase, the ν_4 band is seen at 1374 cm^{-1}, with ν_3 seen at 1489/1503 cm^{-1} [100]. A ν_8 band at 338 cm^{-1} was assigned as an Fe-S vibration; this was verified with isotopic substitution.

The human DGCR8 protein has been proposed to bind both ferrous and ferric heme via bis-Cys ligation [37]. A band is seen at 349 cm^{-1} in the ferric form, and at 347 cm^{-1} in the ferrous form. Remaining studies done in by Girvan et al. to suggest Cys ligation were discussed above in section 3.2.2 [37].

In conclusion, some of the next steps to identifying the axial ligand that binds heme could include titration of PefR with PMe$_3$, mutation of Cys109, and collection of the Raman spectra of the ferric, ferrous and CO-bound PefR. If Cys109 is the axial ligand, PMe$_3$ titration would produce a hyperporphyrin spectrum with distinct bands. Cys109 mutation would result in a decrease or loss of heme-binding abilities if Cys109 is the axial ligand. The presence of an Fe-S stretching band between 310-350 cm^{-1} would support our hypothesis that PefR binds heme via a Cys. Further inferences could be made about the spin state, and coordination of ferrous heme in PefR based on the shifts of bands in the high-frequency region upon reduction.

6 FIGURES

Table 1. Select thiolate heme-binding proteins and their function, as well as their axial ligands in the ferric, ferrous, and CO bound states.

Protein		Function	Ligand		
			Fe^{3+}	Fe^{2+}	Fe^{2+}-CO
Five-coordinate					
Chloroperoxidase [a]		Peroxidation	Cys	Cys	Cys/CO
CP motif [a, b]		Model System	Cys		
Dps		DNA Protective Factor	Cys		
HRI		Transcriptional Regulator	Cys	His	His/CO
Mb H93C		Model System	Cys	Cys	Cys/CO
NO Reductase		NO Activation	Cys/H_2O	Cys	Cys/CO
NO Synthase		Oxygen Activation	Cys	Cys	Cys/CO
Nitrophorin		NO Storage/Transport	Cys	CysH [c]	CysH/CO
PpsR		Transcriptional Regulator	Cys	His	His/CO
$P450_{cam}$		Oxygen Activation	Cys/H_2O	Cys	Cys/CO
Six-coordinate					
CooA [a]		CO Sensor/Transcription Regulator	Cys/Pro	His/Pro	His/CO
Cystathionine β-synthase [a]		CO Sensor/Enzyme	Cys/His	Cys/His	
DGCR8 [a]		mRNA Processing	Cys/Cys	CysH/CysH [c]	CysH/CO
DHR51 [b]		Hormone Receptor	Cys/His	Z [d]	Z/CO [d]
E75 [b]		Nuclear Receptor	Cys/His	His	His/CO
Rcom-2 [a]		CO Sensing Transcriptional Regulator	Cys/His	Met/His	His/CO
Rev-erβ [a]		Nuclear Hormone Receptor	Cys/His	His/ Z [d]	His
Two binding sites					
Bach 1 [a]	Site 1	Transcriptional Regulator	Cys		
	Site 2		Cys/His		
Irr [a,b]	Site 1	Transcriptional Regulator	Cys		
	Site 2		His/His	His/His	
IRP1 [a]	Site 1	Post Transcriptional Regulator	Cys		His
	Site 2		His/(Cys/His) [e]	Cys/His	
IRP2 [a]	Site 1	Post Transcriptional Regulator	Cys		His/CO
	Site 2		His/(Cys/His) [e]	Cys/His	

[a] Cys binding proteins that contain/utilize a CP motif.
[b] Indicates truncated heme binding region of protein, not full-length protein.
[c] CysH indicates a neutral thiol.
[d] Unidentified axial ligands denoted as Z.

[e] (x/x) indicates one of the two, ligation hasn't been confirmed
Data compiled from reviews including [15, 34, 53, 76, 106, 107, 114]

Table 2. Axial ligands and UV/visible absorption spectra of ferric cysteine-bound heme proteins.

Protein		Axial Ligand	n Band	Soret Band	Visible Bands	Charge Transfer	Source
Five-coordinate							
Chloroperoxidase		Cys		390	514, 542	650	[116, 117]
CP motif [a]		Cys		362			[149]
Dps		Cys		385	544	650	[35]
HRI		Cys		418	538		[54]
Mb H93C		Cys		391	509	629	[1]
Mb H93G		BME [b]		390	510	644	[29]
NO Reductase		Cys		406	540		[108]
NO Synthase		Cys		393		643	[96]
Nitrophorin		Cys		389			[130]
PpsR		Cys		372	515, 550	~640	[145]
P450$_{cam}$		Cys/H$_2$O		391	509	618	[118]
Six-coordinate							
CooA		Cys/Pro	360	424	538, 570	760	[9]
Cystathionine β-synthase		Cys/His	363	428	550		[60]
DGCR8		Cys/Cys	365	450	556	660	[10, 11]
DHR51		Cys/His	362	424	542, 575	645	[26]
E75 [a]		Cys/His	360	423	541, 569	650, 750	[3, 72]
Rcom-2		Cys/His	354	423	541, 565	649, 750	[75]
Rev-erβ		Cys/His	358	423	543, 577		[71]
Two binding sites							
Bach 1	Site 1	Cys		371	521, 541	650	[49]
	Site 2	Cys/His		423	540, 580		
Irr [a]	Site 1	Cys	372	414		650	[56, 143]
	Site 2	His/His					
IRP1	Site 1	Cys	372	415	530	650	[81]
	Site 2	His/(Cys/His) [c]					
IRP2	Site 1	Cys	373	415	530		[81]
	Site 2	His/(Cys/His) [c]					

[a] Indicates truncated heme binding region of protein, not full-length protein.
[b] BME is abbreviation for β-mercaptoethanol
[c] (x/x) indicates one of the two, ligation hasn't been confirmed

Table 3. Axial ligands and UV/visible absorption spectra of ferrous cysteine-bound heme proteins.

Protein	Fe^{2+}			Fe^{2+}-CO		
	Axial ligand	Soret	Visible bands	Axial ligand	Soret	Source
Five-coordinate						
Chloroperoxidase	Cys	508	552	Cys/CO	446	[116, 117]
Dps	nr	535	555	---	---	[35]
HRI	His	426	531,560	His/CO	421	[54]
Mb H93C	Cys	428	558	Cys/CO	420	[1]
Mb H93G	CPSH [b]	426	559	CPSH/CO [b]	422	[86]
NO Reductase	Cys	414	533	Cys/CO	447	[108, 129]
NO Synthase	Cys	409	533	Cys/CO	445	[96]
Nitrophorin	CysH [c]	400		CysH/CO	422	[130]
PpsR	His/Cys	426	530, 560	His/CO	418	[145]
P450$_{cam}$	Cys	412	542	Cys/CO	446	[25, 29]
Six-coordinate						
CooA	His/Pro	424	528, 557	His/CO	422	[9, 22, 104]
DGCR8	CysH/CysH [c]	425	530, 557	CysH/CO	422	[10] [37]
DHR51	Z [d]	424	531, 560	Z/CO [d]	420	[26]
E75 [a]	His	425	530, 559	His/CO [d]	420	[3, 72]
Rcom-2	Met/His	425	532, 562	His/CO	423	[71]
Rev-erβ	His/ Z [d]	426	530, 559	His/CO	420	[84]
Two binding sites						
IRP1	Cys/His	372, 420		His/CO	421	[81] [55]
IRP2	Cys/His	373, 420		His/CO	421	[55, 81]

[a] Indicates truncated heme binding region of protein, not full-length protein.
[b] Cyclopentanethiol abbreviated as CPSH.
[c] CysH indicates a neutral thiol.
[d] Unidentified axial ligands denoted as Z.

Table 4. Secondary structure content of select MarR proteins, predicted using CD.

Protein	Species	Ligand	α-Helix	β-Sheet	Turn	Disordered	Source
HucR	*D. radiodurans*	Urate	64%	12%	9%	15%	[28]
PecS	*S. coelicolor*	Urate	52%	6%	17%	25%	[52]
BldR	*S. solfataricus*	Benzaldehyde, salicylate	66%	9%	-	-	[33]
MexR	*P. aeruginosa*	β-Lactamin	61%	5%	12%	21%	[5]
TamR	*S. coelicolor*	*trans-*Aconitate, citrate, *cis-*aconitate, isocitrate	57%	10%	33% random coil		[51]

Table 5. The timeline of storage conditions for samples studied using CD.

Sample	Time of first condition	Time of second condition	Time of third condition
A	Day of purification		
B	6 d at 4 °C		
C	7 d at -20 °C		
D	13 d at 4 °C		
E	6 d at 4 °C	10 d at -20 °C	3 d at 4 °C
F	19 d at 4 °C		
G	22 d at 4 °C		

Table 6. Ratio of absorbance at 260 and 280 nm for PefR as a function of purification technique.

Purification from Cells

Protocol 1	1.3
Protocol 2a	0.6
Protocol 2b	0.9
Protocol 3	0.9
Protocol 4	1.5

Purification from Protein

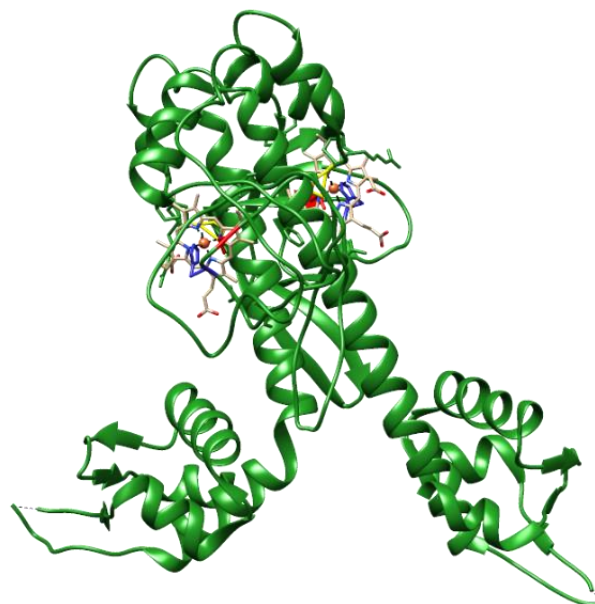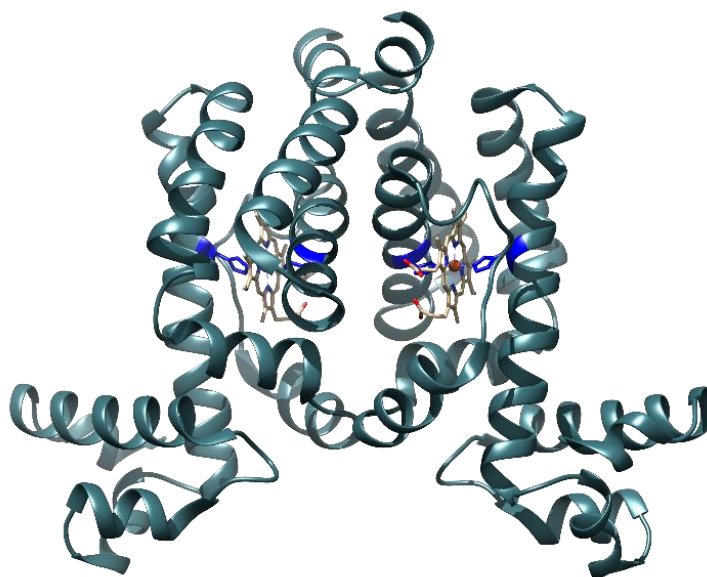|	
Protein as isolated	1.5
Heme loaded	1.3
PD-10 eluent	1.3
Holo-PefR with DNase	1.3
Strep-Tactin Eluent	1.5
Ion Exchange Eluent	0.5

A

B

Figure 1. Full-sequence heme-binding transcriptional regulators that have been crystalized bound to heme. A. CooA (1FT9) Cys75 (Red), Pro2 (Yellow) and His77 (Blue) highlighted. B. HrtR (3VP5) His-72 and His-149 (Blue)

```
Lab        GAGACCGGGGGTTCGGGGGTGGGCGaATTCCTCTAGAATAATTTTGTTTAACTTTAAGAAG 60
Confirmed  ------------------------------------------------------------ 0


Lab        GAGATATACATATGGCCAGCTGGAGCCACCCGCAGTTCGAAAAGAGCGGTGGTGGTGGTG 120
Confirmed  ------------------------------------------------------------ 0


Lab        GTGAAAATCTGTATTTCCAGGGTATGTCACAAGTGATAGGTGATTTACGTGAATTGATAC 180
Confirmed  ---------------------Atgtcacaagtgataggtgatttacgtgaattgatac 37
                                ********************************

Lab        ATCAAATCGAACAAATTAGTGATGAGATTGCAAAAAAATATGATGTAGAGCATCTAGCAG 240
Confirmed  atcaaatcgaacaaattagtgatgagattgcaaaaaaatatgatgtagagcatctagcag 97
           ************************************************************

Lab        GTCCTCAAGGTTATGTTCTTGTTTTTTTAGCTAAACACCAAAATCAAGAAATATTTGTCA 300
Confirmed  gtcctcaaggttatgttcttgttttttttagctaaacaccaaaatcaagaaatatttgtca 157
           ************************************************************

Lab        AAGATATTGAAAAACAACTTCGTATCTCAAAGTCAGTTGCTAGTCATTTAGTGAAACGTA 360
Confirmed  aagatattgaaaaacaacttcgtatctcaaagtcagttgctagtcatttagtgaaacgta 217
           ************************************************************

Lab        TGGTCAAAAATGGGTTTATCAATGTGATGCCTTCCCAAGTGGATAAGCGTTATAAGCAAG 420
Confirmed  tggtcaaaaatgggtttatcaatgtgatgccttcccaagtggataagcgttataagcaag 277
           ************************************************************

Lab        TAGTGTTAGCGCAGGTTGGTAGAGATAAATTGCCTTTGTTGCGGGAGTGTCGTAAGGATA 480
Confirmed  tagtgttagcgcaggttggtagagataaattgcctttgttgcgggagtgtcgtaaggata 337
           ************************************************************

Lab        TCGAGCACTATTTTTTAAAAGAAATTACAAAAGAAGAGTTGCTGACAGCGAAAAAAGTAA 540
Confirmed  tcgagcactattttttaaaagaaattacaaaagaagagttgctgacagcgaaaaaagtaa 397
           ************************************************************

Lab        TTGAACAGCTCAAGCAAAATATGCTAACTTATAAAGGAGACAACGATGCTTAAAAGCTTG 600
Confirmed  ttgaacagctcaagcaaaatatgctaacttataaaggagacaacgatgcttaa------- 450
           ***********************************************************

Lab        CGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAA 660
Confirmed  ------------------------------------------------------------ 450


Lab        AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT 720
Confirmed  ------------------------------------------------------------ 450


Lab        CTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATG 780
Confirmed  ------------------------------------------------------------ 450
```

Figure 2. Sequence alignment of the plasmid isolated from cells used to grow PefR for this work (Lab), and the sequence provided by the Eichenbaum lab (Confirmed).
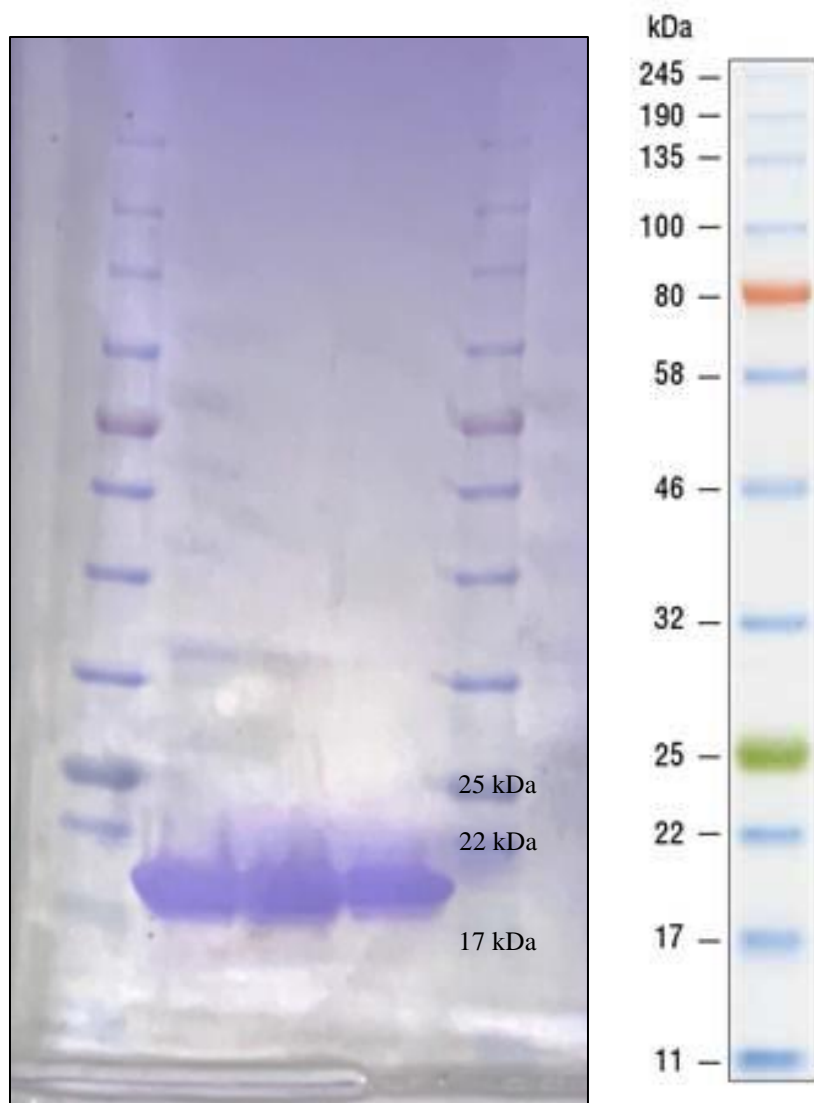
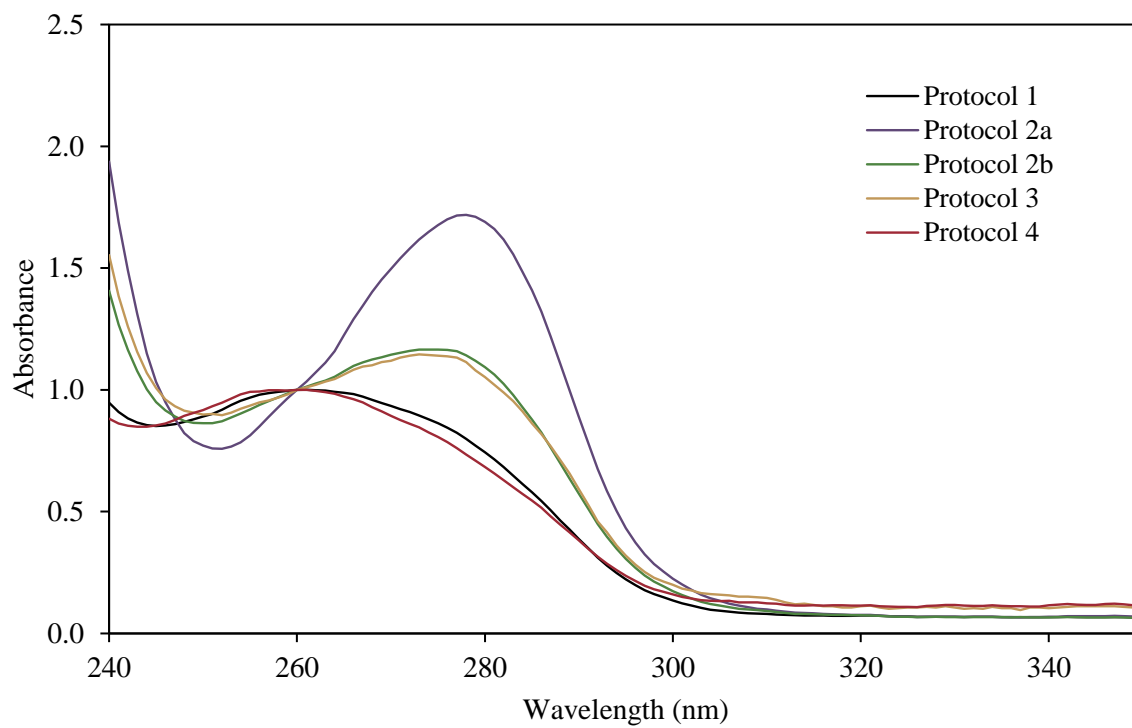*Figure 3. SDS-PAGE gel of PefR purified Lane 1: Protein ladder. Lanes 2-4: PefR, Lane 5: Protein ladder.*

Figure 4. UV/visible spectra of different samples from protocols utilized to purify DNA-free PefR from cells. The spectra were normalized to 1.0 at 260 nm.
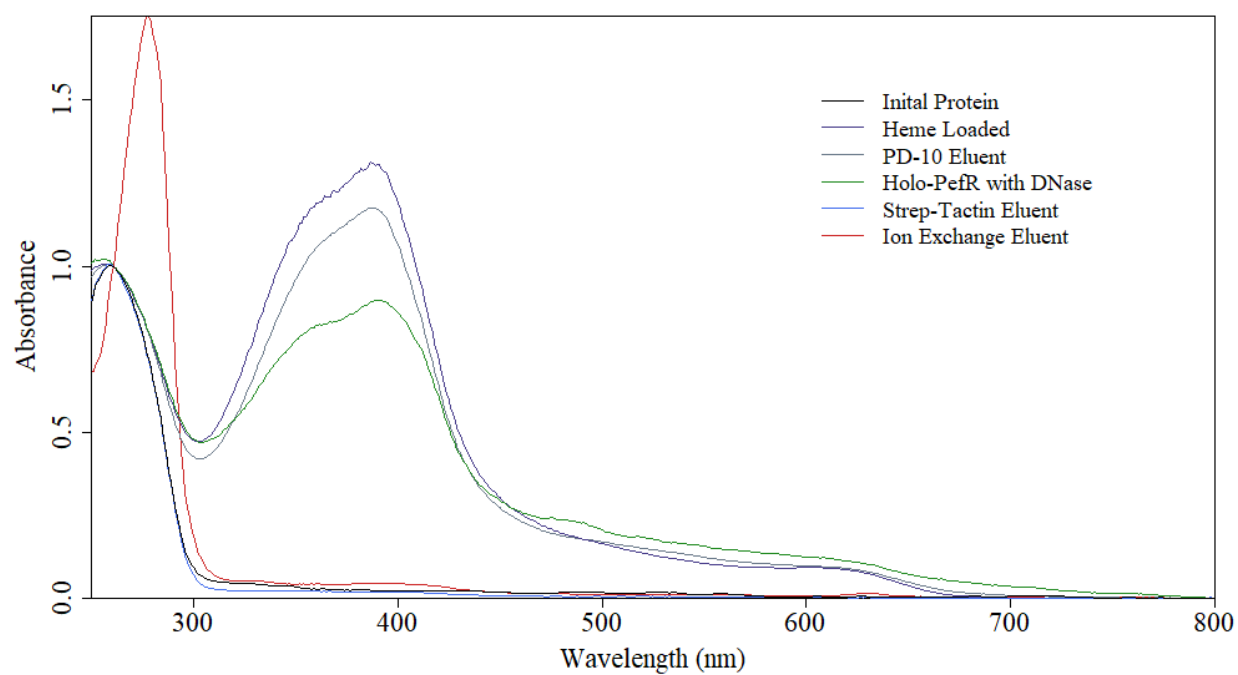
Figure 5. UV/visible spectra of samples from different steps taken to remove DNA bound to purified PefR. The spectra were normalized to 1.0 at 260 nm.
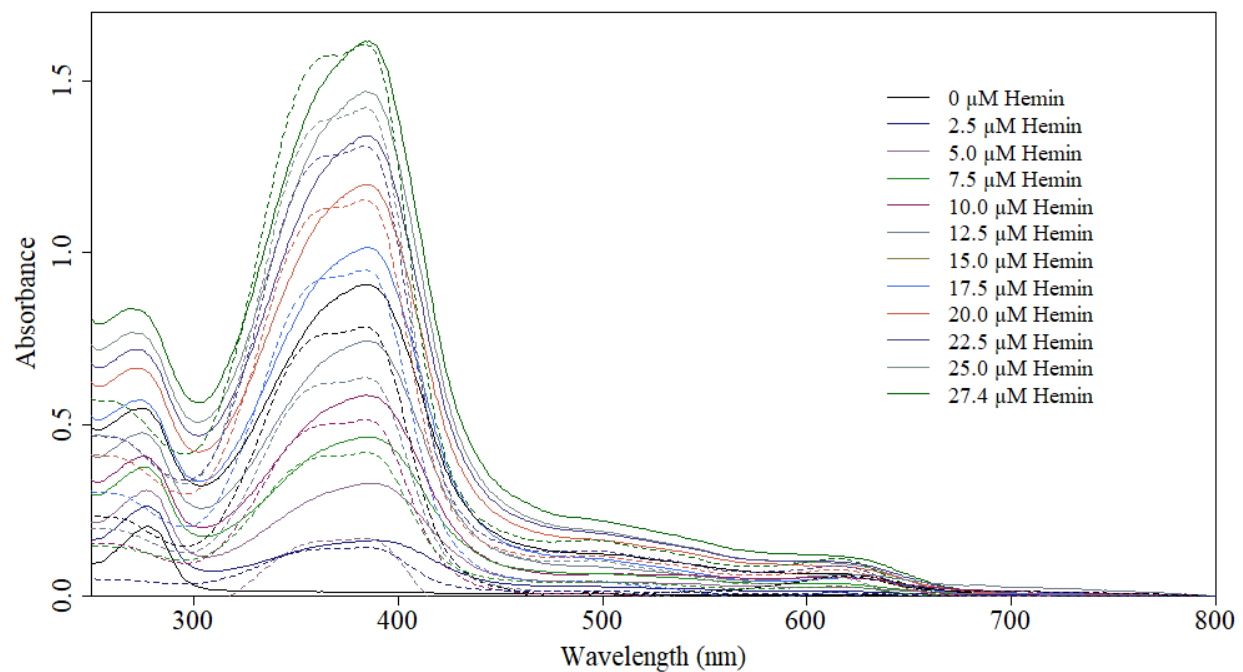
Figure 6. Heme titration of PefR in buffer A. Absorbance spectra of 14 μm PefR solutions (solid line) as a function of amount of added hemin in DMSO. Absorbance spectra of buffer solution (---) containing equal amounts of hemin shown with dashed lines.
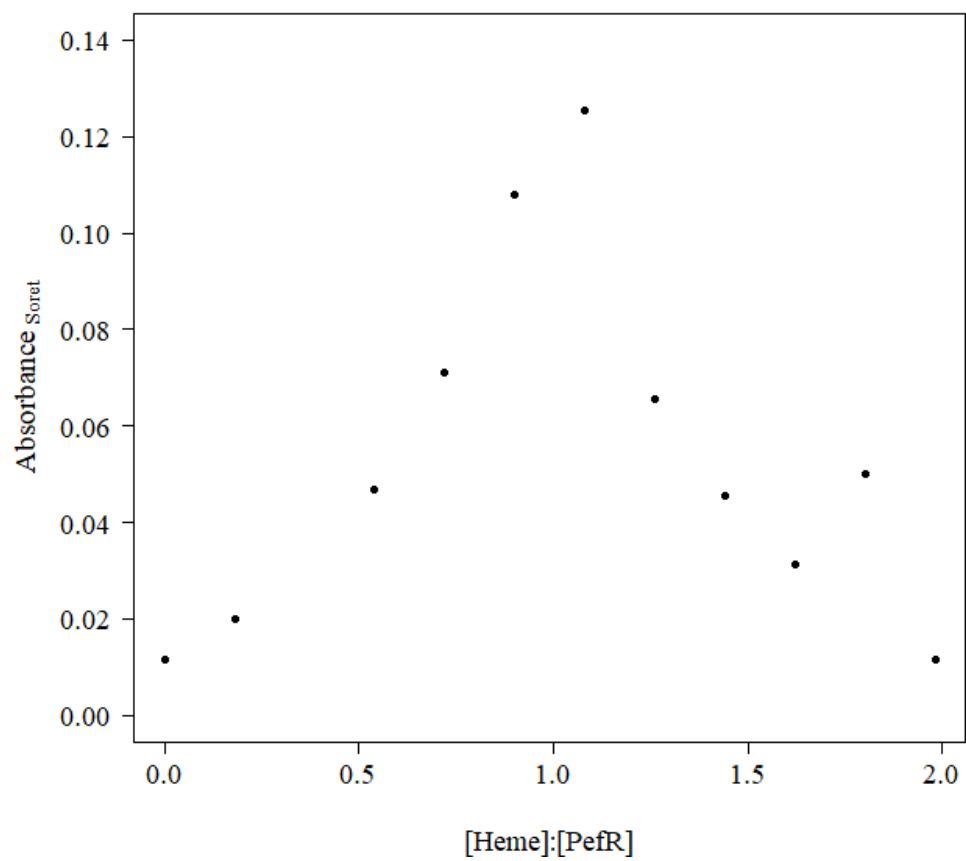
Figure 7. Absorbance at the Soret of PefR after subtracting the absorbance of the hemin solution, as a function of the ratio of heme:PefR.
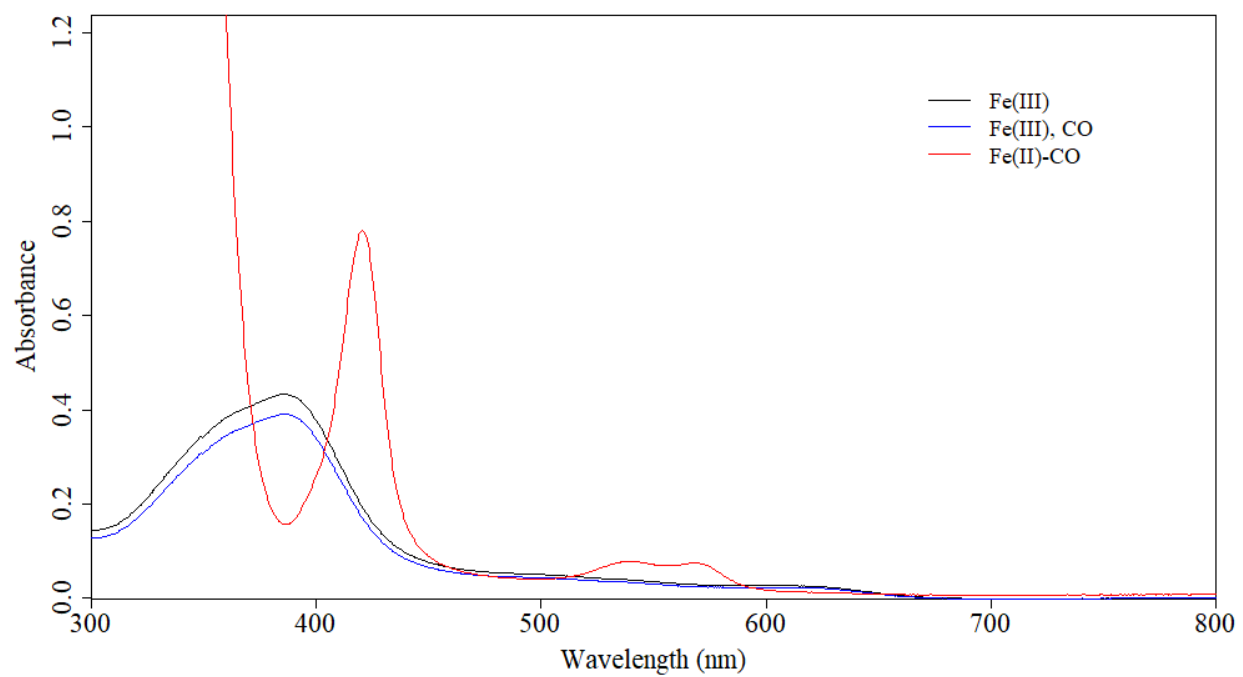
Figure 8. UV/visible spectra of ferric PefR, ferric PefR with CO in solution, and ferrous PefR with CO. All solutions in buffer A.
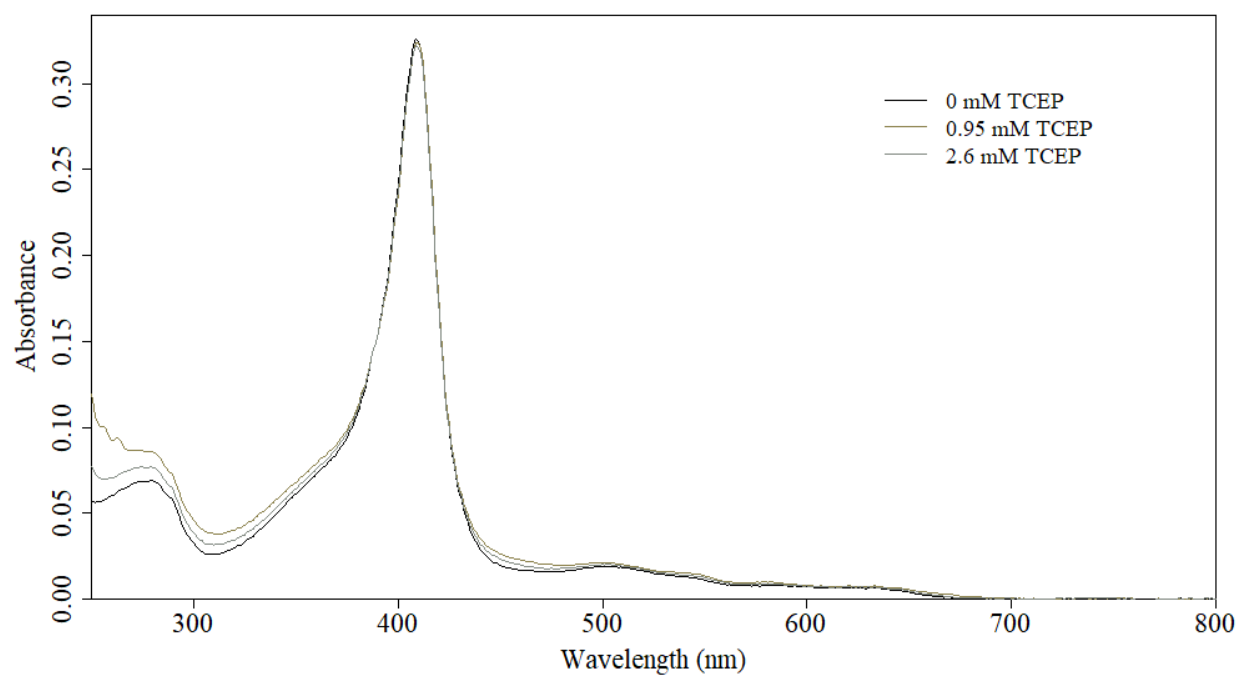
Figure 9. The UV/visible absorption spectra of 1.4 µM myoglobin solution in buffer A as a function of the addition of 18.1 mM TCEP in buffer A.
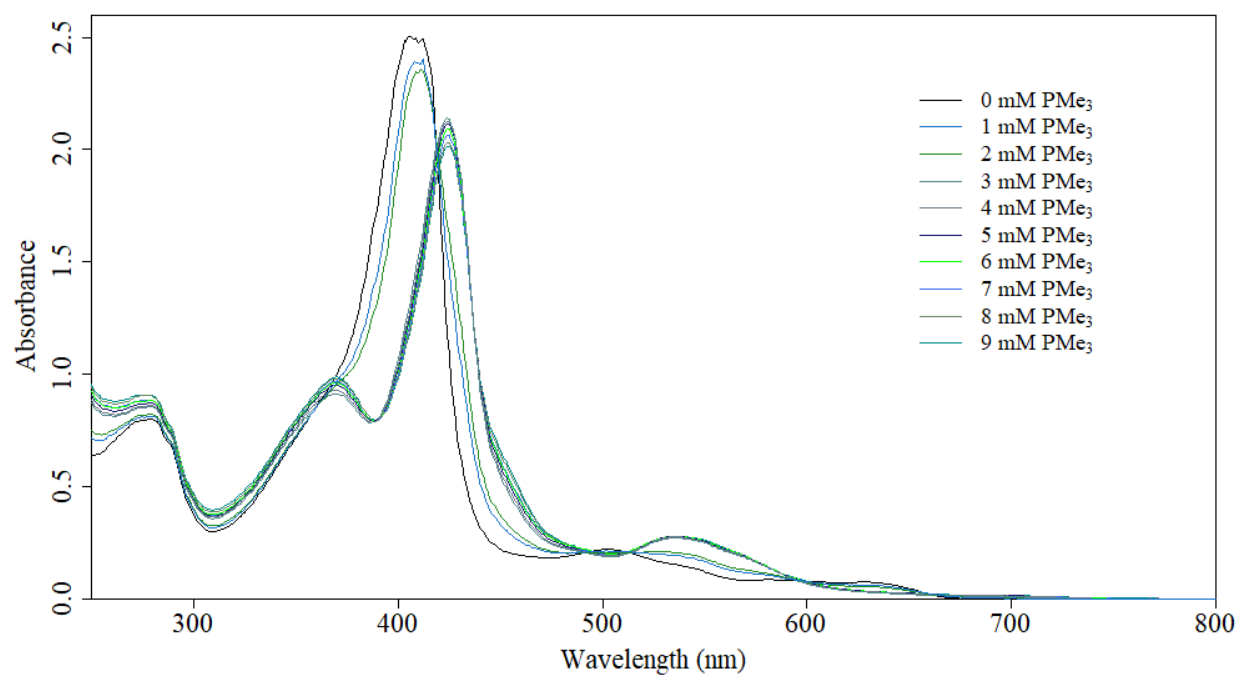
Figure 10. UV/visible absorption spectra of 22 μM myoglobin solution in buffer A as a function of the addition 1 M Me3P in THF. Over 400 eq (9 mM) of PMe3 was titrated into the myoglobin solution
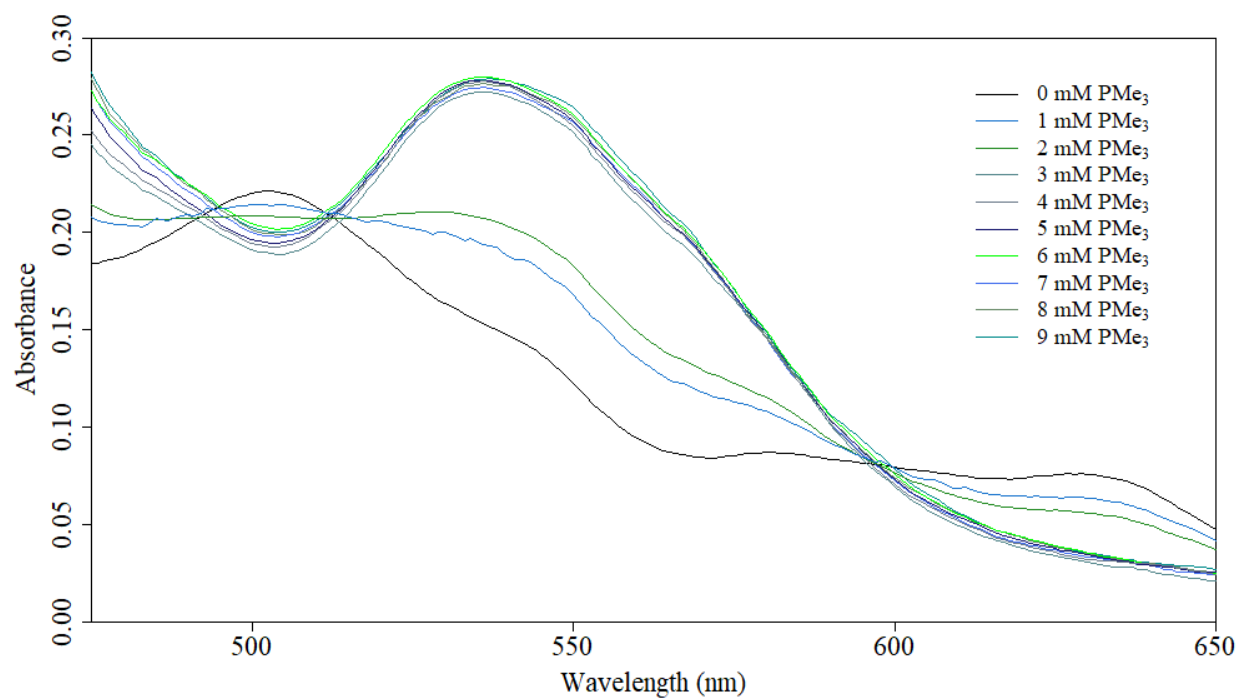
Figure 11. The UV/visible spectra of visible region of 22 µM myoglobin solution in buffer A as a function of the addition of aliquots of 1 M Me3P in THF. Over 400 eq (9 mM) of Me3P was titrated into the myoglobin solution.

Figure 12. UV/visible spectra of reduced 22 μM myoglobin solution in buffer A as a function of Me3P addition (40 mM Me3P in 1:49 THF:buffer A, followed by 1.0 M Me3P in THF).

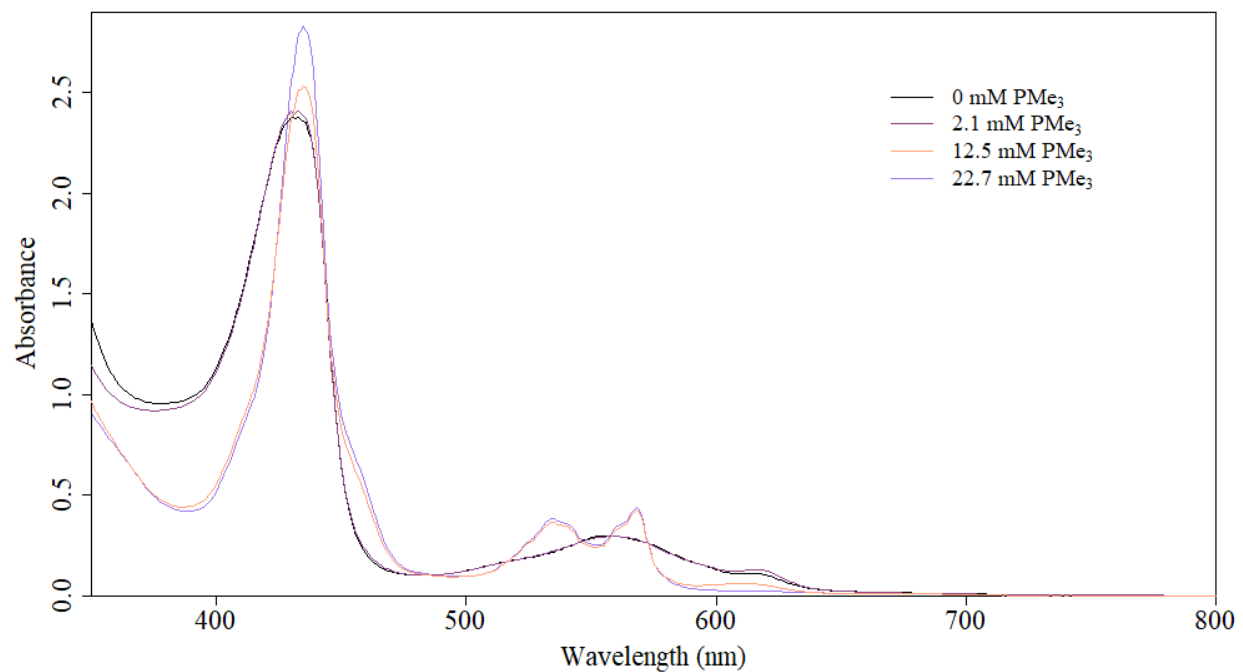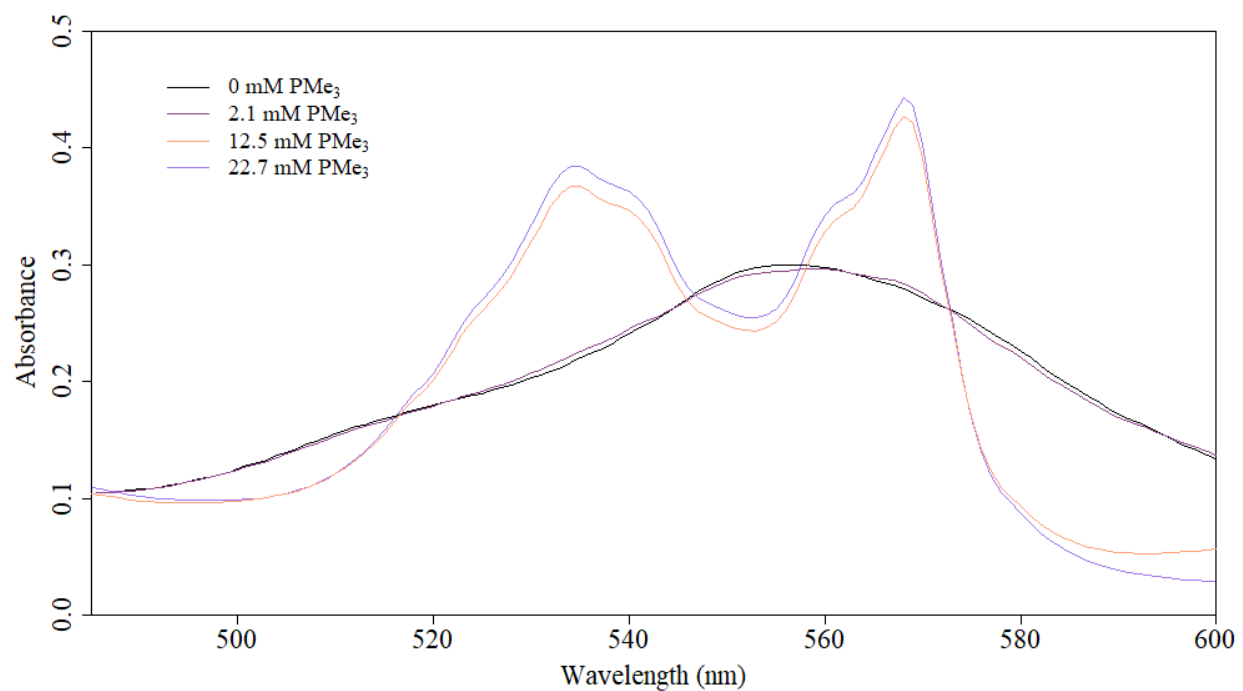*Figure 13. UV/visible spectra in visible region of reduced 22 µM myoglobin solution in buffer A as a function of Me3P addition (40 mM Me3P in 1:49 THF:buffer A, followed by 1.0 M Me3P in THF).*
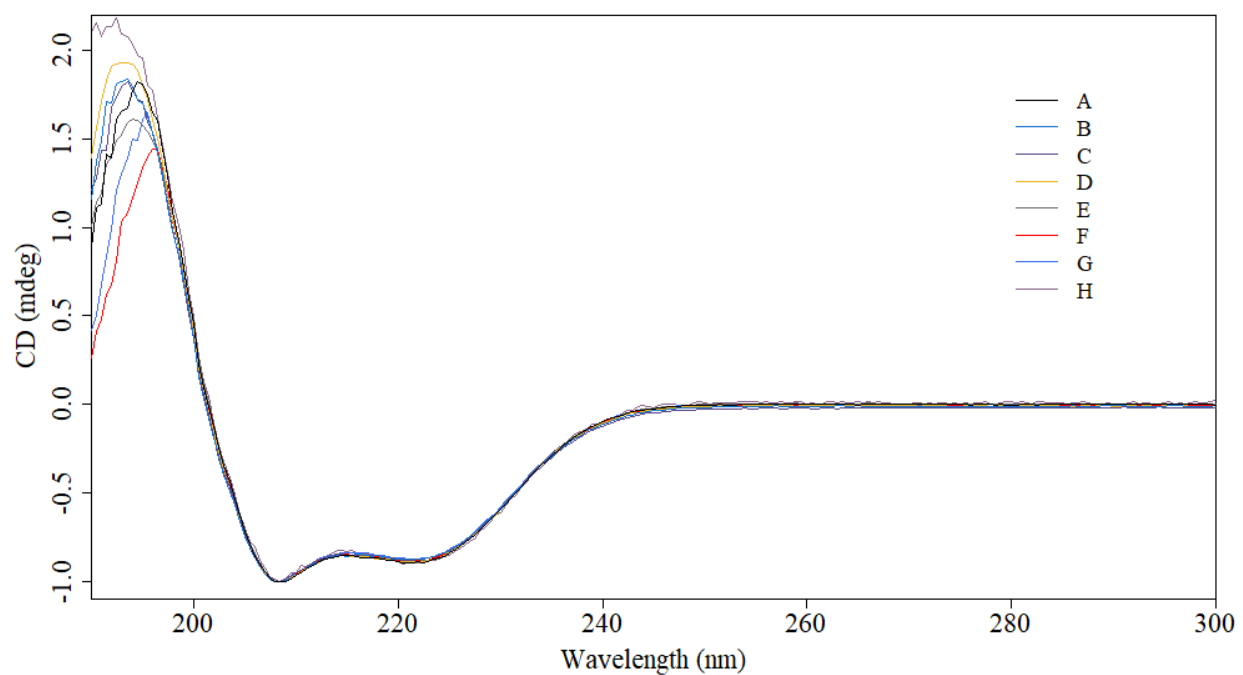
Figure 14. CD spectra of PefR in phosphate buffer as a function of time and storage temperature. PefR storage conditions are given in table 5. Spectra are all normalized to -1.0 at 208 nm.

```
Sco3205    MHHHHHHENLYFQGMNDEPRWLTAEEQLVWRSYIEAATLLEDHLDRQ-----LQRDAGMP      55
PefR       ----------------MSQVI-----GDLRELIHQIEQISDEIAKKYDVEHLAGPQGY-      37
MepR       ---------------SNEFTYS-----YLFRMISHEMKQKADQKLEQFD---ITNEQGH-      36
                                         *    .       *.  .:      :      *

Sco3205    HVYYGLLVKLAESPRRRLRMTELAKYAKITRSRLSHAVARLEKNGWVRREDCPSDKRGQF     115
PefR       -----VLVFLAKHQNQEIFVKDIEKQLRISKSVASHLVKRMVKNGFINVMPSQVDKRYKQ      92
MepR       -----TLGYLYAHQQDGLTQNDIAKALQRTGPTVSNLLRNLERKKLIYRYVDAQDTRRKN      91
                *   *      . :  .:: *  : :     *: : .: ::  :       *.* :

Sco3205    AILTDEGYEVLRRTAPGHVDAVRQAVFDRLTPEQQKSLGEIMRIVAEGLQP----SEAGA     171
PefR       VVLAQVGRDKLPLLREC-RKDIEHYFLKEITKEELLTAK----KVIEQLKQNMLTYKGDN     147
MepR       IGLTTSGIKLVEAFTSI-FDEMEQTLVSQLSEEENEQMKANLTKMLSSLQ----------     140
               *:  * . :         . :.: ....:: *:            : . *:

Sco3205    DLPWLR      177
PefR       DA----      149
MepR       ------      140
```

Figure 15. Sequence Alignment of PefR with Sco3205 from S. coelicolor (3ZPL) and MepR from S. aureus (4LLN) done using Clustal Omega.
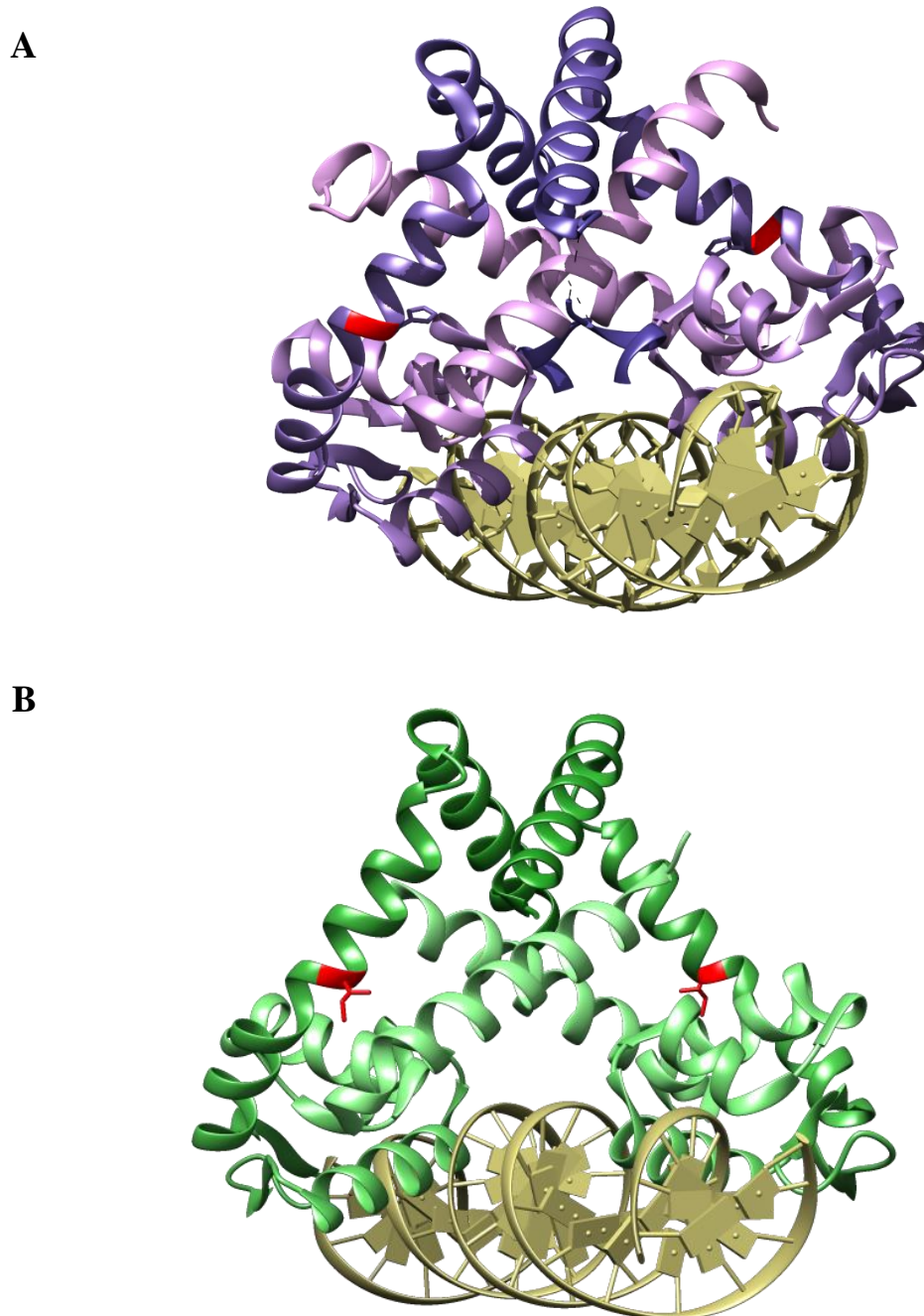
*Figure 16. **A**. MepR from Staphylococcus aureus (4LLN) bound to DNA with the Ile107 highlighted. **B**. Sco3205 from Streptomyces coelicolor (3ZPL) bound to DNA with Gly118 highlighted. Ile107 and Gly118 both align with Cys109 in PefR.*
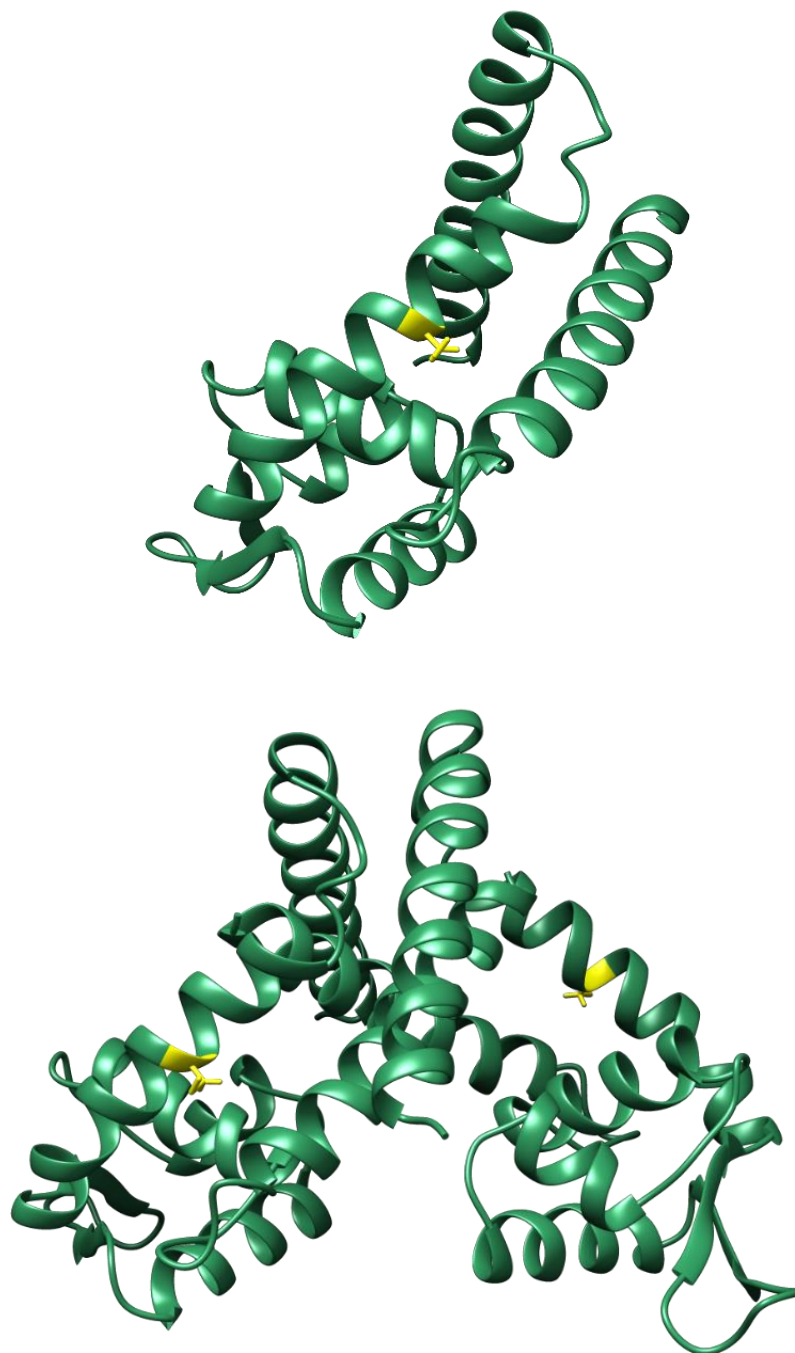
Figure 17. Homology model of PefR produced by I-TASSER. Cys109, the expected axial ligand, is highlighted in yellow. (Top) Monomer produced from I-TASSER. (Bottom) Dimer created using the Matchmaker function of Chimera and a model generated by the SPRING server.
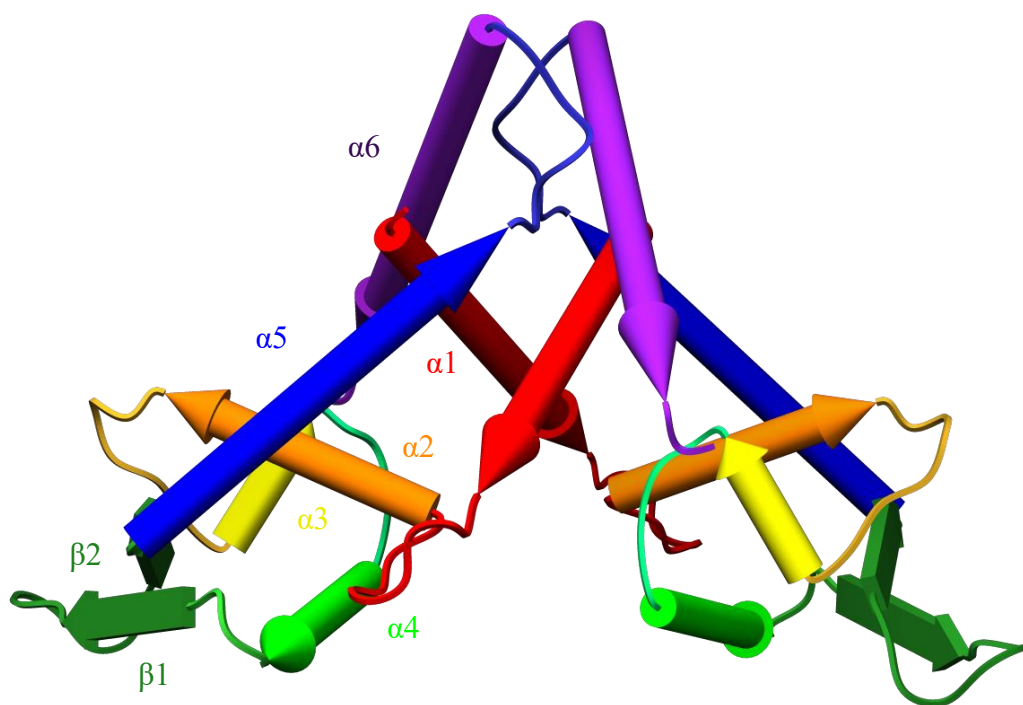
Figure 18. Cartoon representation of PefR. Helices are shown in rainbow order α1 (red) to α6 (purple).
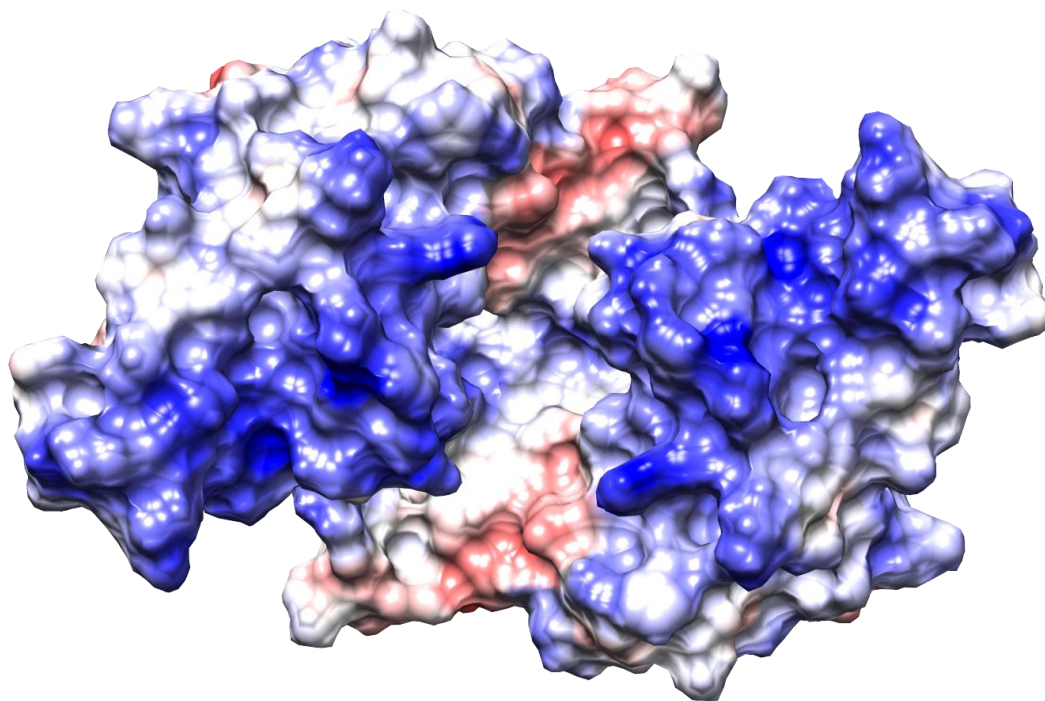
Figure 19. Electrostatic surface representation of predicted DNA binding region of PefR dimer. Negative residues depicted in red, positive residues depicted in blue.
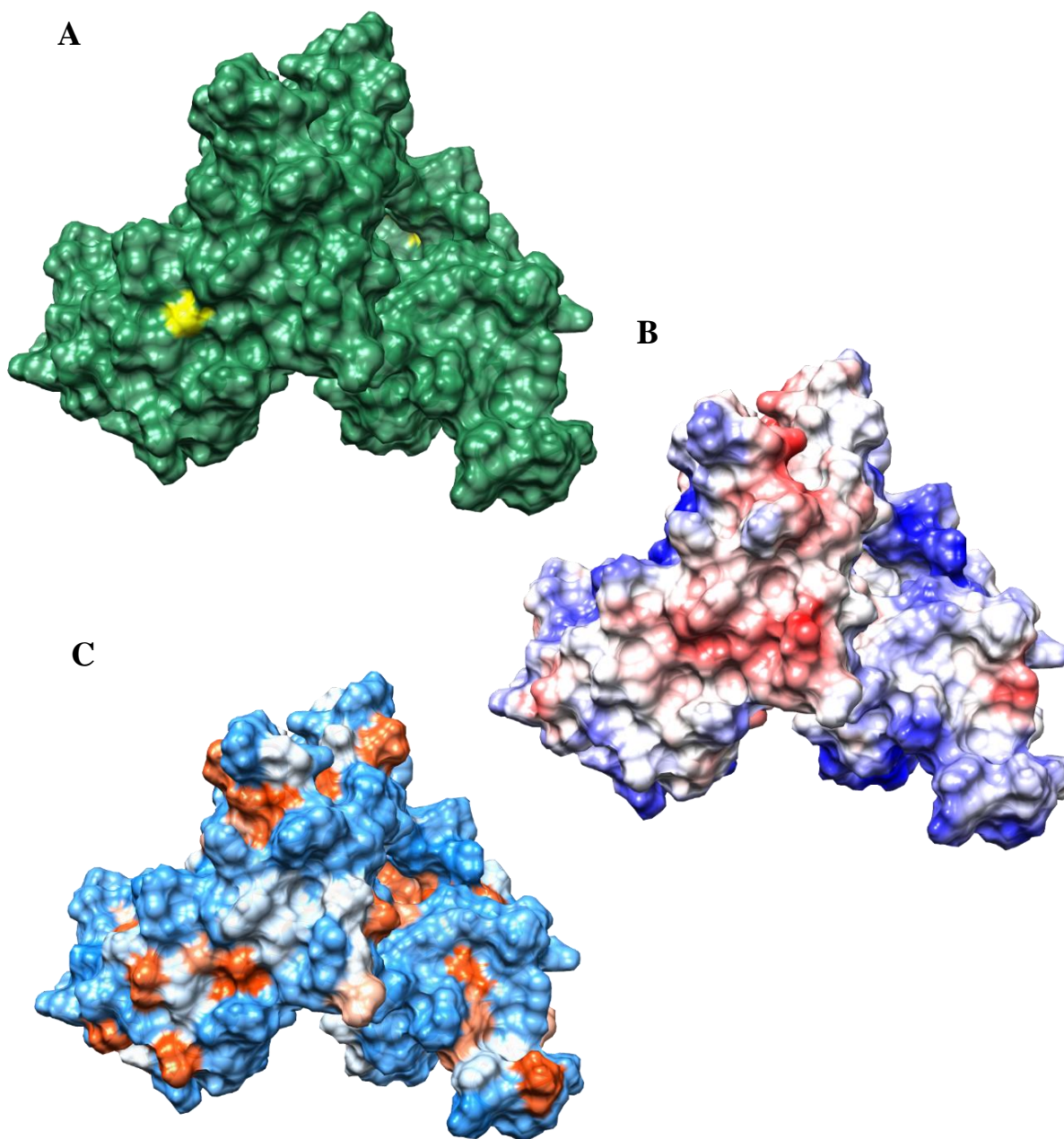
Figure 20. ***A****. Surface representation of PefR. Cys109 depicted in yellow.* ***B****. Electrostatic surface representation of PefR dimer. Negative residues depicted in red, positive residues depicted in blue.* ***C****. Hydrophilic residues depicted in blue, hydrophobic residues depicted in orange.*
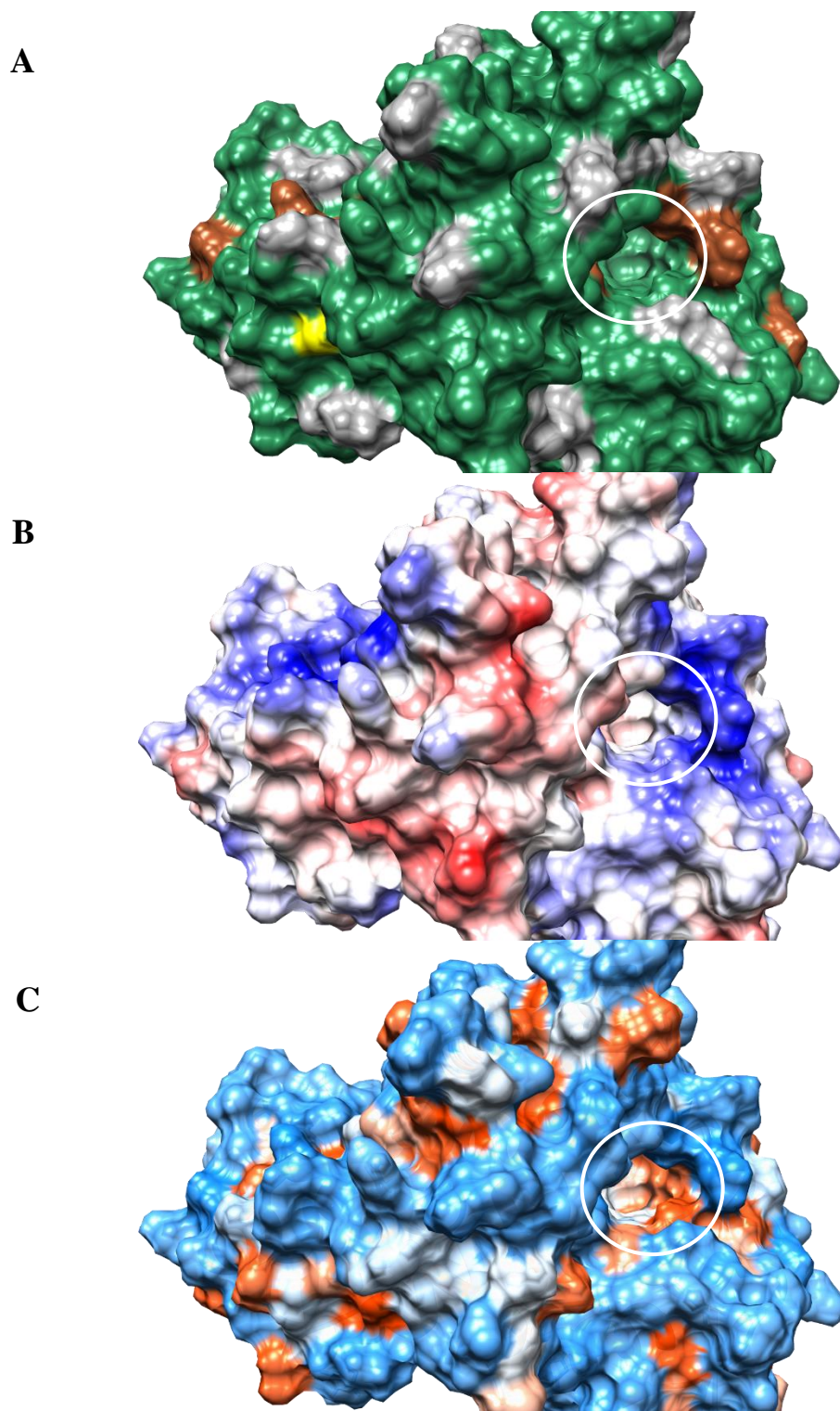
*Figure 21. Proposed heme binding cleft marked with white circle. **A.** Surface representation of PefR dimer. Cys109 depicted in yellow, Arg depicted in brown, Lys depicted in grey. **B.** Electrostatic surface representation of PefR dimer. Negative residues depicted in red, positive residues depicted in blue. **C.** Hydrophobicity of surface residues. Hydrophilic residues depicted in blue, hydrophobic residues depicted in orange.*

Figure 22. The distances between the sulfur of Cys109 (yellow) and the nitrogen of all of the His (blue) in the same monomer. Distances measured using Chimera.

REFERENCES

Uncategorized References

1. Adachi, S., et al., *Roles of proximal ligand in heme proteins: Replacement of proximal histidine of human myoglobin with cysteine and tyrosine by site-directed mutagenesis as models for P-450, chloroperoxidase, and catalase.* Biochemistry, 1993. **32**(1): p. 241-252.

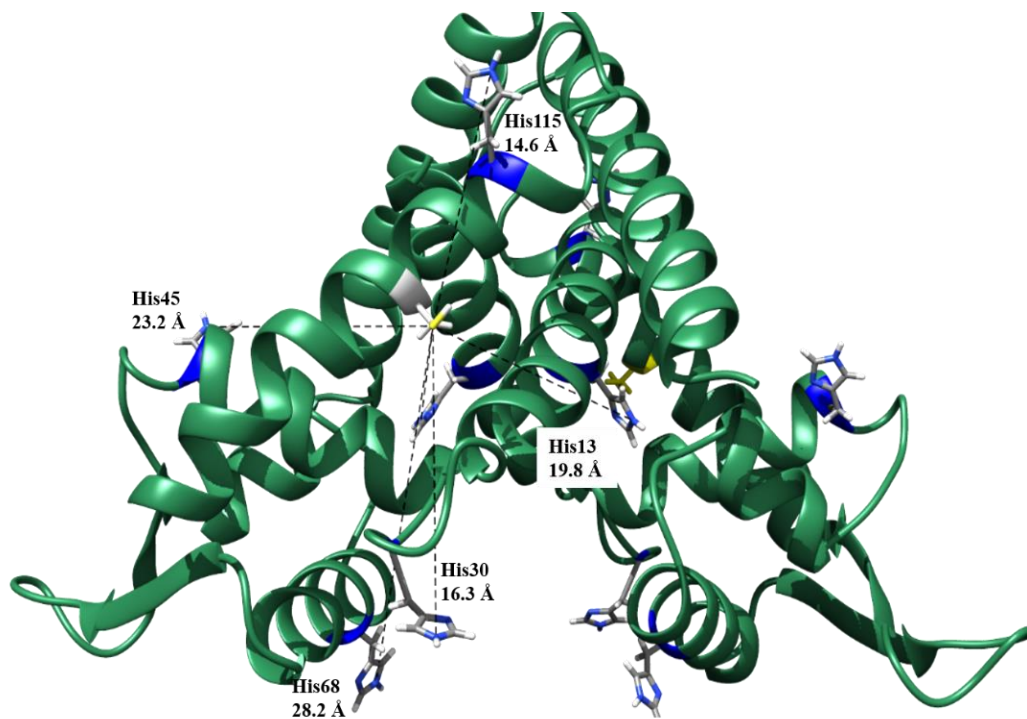2. Adachi, S., et al., *Alteration of human myoglobin proximal histidine to cysteine or tyrosine by site-directed mutagenesis: characterization and their catalytic activities.* Biochem. Biophys. Res. Commun., 1991. **180**(1): p. 138-144.

3. Aicart-Ramos, C., et al., *Covalent attachment of heme to the protein moiety in an insect E75 nitric oxide sensor.* Biochemistry, 2012. **51**(37): p. 7403-7416.

4. Alekshun, M.N., et al., *The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution.* Nat. Struct. Biol., 2001. **8**(8): p. 710-714.

5. Andresen, C., et al., *Critical biophysical properties in the Pseudomonas aeruginosa efflux gene regulator MexR are targeted by mutations conferring multidrug resistance.* Protein Sci., 2010. **19**(4): p. 680-692.

6. Antonini, E. and M. Brunori, *Hemoglobin and Myoglobin in their Reactions with Ligands*. Front. Biol. Vol. 21. 1971, Amsterdam: North-Holland Publishing Company.

7. Anzaldi, L.L. and E.P. Skaar, *Overcoming the heme paradox: Heme toxicity and tolerance in bacterial pathogens.* Infect. Immun., 2010. **78**(12): p. 4977-4989.

8. Aojula, H.S., et al., *H-1-NMR and CD studies of heme orientational disorder in sperm-whale myoglobin and human-hemoglobin.* Biochem. J., 1988. **250**(3): p. 853-858.

9. Aono, S., et al., *Redox-controlled ligand exchange of the heme in the CO-sensing transcriptional activator CooA.* J. Biol. Chem., 1998. **273**(40): p. 25757-25764.

10. Barr, I., et al., *Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing.* Proc. Natl. Acad. Sci. U. S. A., 2012. **109**(6): p. 1919-1924.

11. Barr, I., et al., *DiGeorge critical region 8 (DGCR8) is a double-cysteine-ligated heme protein.* J. Biol. Chem., 2011. **286**(19): p. 16716-25.

12. Birukou, I., et al., *Structural mechanism of transcription regulation of the Staphylococcus aureus multidrug efflux operon mepRA by the MarR family repressor MepR.* Nucleic Acids Res., 2014. **42**(4): p. 2774-2788.

13. Blauer, G., N. Sreerama, and R.W. Woody, *Optical activity of hemoproteins in the Soret region. Circular dichroism of the heme undecapeptide of cytochrome c in aqueous solution.* Biochemistry, 1993. **32**: p. 6674-6679.

14. Bowman, H.E., M.R. Dent, and J.N. Burstyn, *Met(104) is the CO-replaceable ligand at Fe(II) heme in the CO-sensing transcription factor BxRcoM-1.* J. Biol. Inorg. Chem., 2016. **21**(4): p. 559-569.

15. Brewitz, H.H., G. Hagelueken, and D. Imhof, *Structural and functional diversity of transient heme binding to bacterial proteins.* Biochim. Biophys. Acta, 2017. **1861**(3): p. 683-697.

16. Brown, S.B. and I.R. Lantzke, *Solution structures of ferrihaem in some dipolar aprotic solvents and their binary aqueous mixtures.* Biochem. J., 1969. **115**(2): p. 279-285.

17. Brunel, C., A. Bondon, and G. Simonneaux, *Trimethylphosphine binding to horse-heart and sperm-whale myoglobins. Kinetics, proton magnetic resonance assignment and nuclear Overhauser effect investigation of the heme pocket.* Eur. J. Biochem., 1993. **214**(2): p. 405-414.

18. Champion, P.M., et al., *Resonance Raman detection of an iron-sulfur bond in cytochrome P 450cam.* J. Am. Chem. Soc., 1982. **104**(20): p. 5469-5472.

19. Chang, Y.M., et al., *Functional studies of ssDNA binding ability of MarR family protein TcaR from Staphylococcus epidermidis.* PLoS One, 2012. **7**(9): p. 1-10.

20. Chang, Y.M., et al., *Structural study of TcaR and its complexes with multiple antibiotics from Staphylococcus epidermidis.* Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(19): p. 8617-8622.

21. Choma, C.T., et al., *Design of a heme-binding 4-helix bundle.* J. Am. Chem. Soc., 1994. **116**(3): p. 856-865.

22. Clark, R.W., et al., *Investigation of the role of the N-terminal proline, the distal heme ligand in the CO sensor CooA.* Biochemistry, 2004. **43**(44): p. 14149-14160.

23. Comer, J.M. and L. Zhang, *Experimental methods for studying cellular heme signaling.* Cells, 2018. **7**(6): p. 12.

24. Davis, J.R., et al., *Study of PcaV from Streptomyces coelicolor yields new insights into ligand-responsive MarR family transcription factors.* Nucleic Acids Res., 2013. **41**(6): p. 3888-3900.

25. Dawson, J.H., L.A. Andersson, and M. Sono, *The diverse spectroscopic properties of ferrous cytochrome P-450-CAM ligand complexes.* J. Biol. Chem., 1983. **258**(22): p. 13637-13645.

26.   de Rosny, E., et al., *DHR51, the Drosophila melanogaster homologue of the human photoreceptor cell-specific nuclear receptor, is a thiolate heme-binding protein.* Biochemistry, 2008. **47**(50): p. 13252-13260.

27.   Deochand, D.K. and A. Grove, *MarR family transcription factors: Dynamic variations on a common scaffold.* Crit. Rev. Biochem. Mol. Biol., 2017. **52**(6): p. 595-613.

28.   Deochand, D.K., et al., *Histidine switch controlling pH-dependent protein folding and DNA binding in a transcription factor at the core of synthetic network devices.* Mol. Biosyst., 2016. **12**(8): p. 2417-2426.

29.   Du, J., M. Sono, and J.H. Dawson, *The H93G myoglobin cavity mutant as a versatile scaffold for modeling heme iron coordination structures in protein active sites and their characterization with magnetic circular dichroism spectroscopy.* Coord. Chem. Rev., 2011. **255**(7-8): p. 700-716.

30.   Eichenbaum, Z., et al., *Acquisition of iron from host proteins by the group A Streptococcus.* Infect. Immun., 1996. **64**(12): p. 5428-5429.

31.   El-Gebali, S., et al., *The Pfam protein families database in 2019.* Nucleic Acids Res., 2019. **47**(D1): p. D427-D432.

32.   Fernández, A., et al., *Two coregulated efflux transporters modulate intracellular heme and protoporphyrin IX availability in Streptococcus agalactiae.* PLoS Pathog., 2010. **6**(4): p. 1-14.

33.   Fiorentino, G., et al., *Identification and physicochemical characterization of BldR2 from Sulfolobus solfataricus, a novel archaeal member of the MarR transcription factor family.* Biochemistry, 2011. **50**(31): p. 6607-6621.

34.   Fleischhacker, A.S., E.L. Carter, and S.W. Ragsdale, *Redox regulation of heme oxygenase-2 and the transcription factor, Rev-Erb, through heme regulatory motifs.* Antioxid. Redox Signal., 2018. **29**(18): p. 1841-1857.

35.   Gao, J.L., et al., *The role of heme binding by DNA-protective protein from starved cells (Dps) in the Tolerance of Porphyromonas gingivalis to heme toxicity.* J. Biol. Chem., 2012. **287**(50): p. 42243-58.

36.   Gasteiger, E., et al., *Protein identification and analysis tools on the ExPASy server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press: Totowa, N.J. p. 571-607.

37.   Girvan, H.M., et al., *Analysis of heme iron coordination in DGCR8: The heme-binding component of the microprocessor complex.* Biochemistry, 2016. **55**(36): p. 5073-5083.

38.   Girvan, H.M. and A.W. Munro, *Heme sensor proteins.* J. Biol. Chem., 2013. **288**(19): p. 13194-13203.

39.   Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nat. Protoc., 2006. **1**(6): p. 2876-2890.

40.   Grove, A., *MarR family transcription factors.* Curr. Biol., 2013. **23**(4): p. R142-R143.

41.   Grove, A., *Regulation of metabolic pathways by MarR family transcription factors.* Comput. Struct. Biotechnol. J., 2017. **15**: p. 366-371.

42.   Guerler, A., B. Govindarajoo, and Y. Zhang, *Mapping monomeric threading to protein-protein structure prediction.* J. Chem. Inf. Model., 2013. **53**(3): p. 717-725.

43.   Gupta, A. and A. Grove, *Ligand-binding pocket bridges DNA-binding and dimerization domains of the urate-responsive MarR homologue MftR from Burkholderia thailandensis.* Biochemistry, 2014. **53**(27): p. 4368-4380.

44.   Gupta, A., et al., *MarR family transcription factors from Burkholderia species: hidden clues to control of virulence-associated genes.* Microbiol. Mol. Biol. Rev., 2019. **83**(1): p. 1-19.

45.   Gupta, N. and S.W. Ragsdale, *Thiol-disulfide redox dependence of heme binding and heme ligand switching in nuclear hormone receptor rev-erbß.* J. Biol. Chem., 2011. **286**(6): p. 4392-4403.

46.   Hanson, L.K., et al., *Letter: Origin of the anomalous Soret spectra of carboxycytochrome P-450.* J. Am. Chem. Soc., 1976. **98**(9): p. 2672-2674.

47.   Hemlata, A.T. Jan, and A. Tiwari, *The ever-changing face of antibiotic resistance: Prevailing problems and preventive measures.* Curr. Drug Metab., 2017. **18**(1): p. 69-77.

48.   Hildebrand, D.P., et al., *Trans effects on cysteine ligation in the proximal His93Cys variant of horse heart myoglobin.* Biochemistry, 1995. **34**(36): p. 11598-11605.

49.   Hira, S., et al., *Bach1, a heme-dependent transcription factor, reveals presence of multiple heme binding sites with distinct coordination structure.* IUBMB Life, 2007. **59**(8-9): p. 542-551.

50.   Hiromoto, T., et al., *Characterization of MobR, the 3-hydroxybenzoate-responsive transcriptional regulator for the 3-hydroxybenzoate hydroxylase gene of Comamonas testosteroni KH122-3s.* J. Mol. Biol., 2006. **364**(5): p. 863-877.

51.   Huang, H. and A. Grove, *The transcriptional regulator TamR from Streptomyces coelicolor controls a key step in central metabolism during oxidative stress.* Mol. Microbiol., 2013. **87**(6): p. 1151-1166.

52.   Huang, H., B.J. Mackel, and A. Grove, *Streptomyces coelicolor encodes aurate-responsive transcriptional regulator with homology to PecS from plant pathogens.* J. Bacteriol., 2013. **195**(21): p. 4954-4965.

53. Igarashi, J., et al., *The roles of thiolate-heme proteins, other than the P450 cytochromes, in the regulation of heme-sensor proteins.* Acta. Chim. Slov., 2008. **55**(1): p. 67-74.

54. Igarashi, J., et al., *Activation of heme-regulated eukaryotic initiation factor 2α kinase by nitric oxide is induced by the formation of a five-coordinate NO-heme complex: Optical absorption, electron spin resonance, and resonance raman spectral studies.* J. Biol. Chem., 2004. **279**(16): p. 15752-15762.

55. Ishikawa, H., et al., *Involvement of heme regulatory motif in heme-mediated ubiquitination and degradation of IRP2.* Mol. Cell, 2005. **19**(2): p. 171-181.

56. Ishikawa, H., et al., *Unusual heme binding in the bacterial iron response regulator protein: Spectral characterization of heme binding to the heme regulatory motif.* Biochemistry, 2011. **50**(6): p. 1016-1022.

57. Ishimori, K. and Y. Watanabe, *Unique heme environmental structures in heme-regulated proteins using heme as the signaling molecule.* Chem. Lett., 2014. **43**(11): p. 1680-1689.

58. Kamal, J.K.A. and D.V. Behere, *Binding of heme to human serum albumin: Steady-state fluorescence, circular dichroism and optical difference spectroscopic studies.* Indian J. Biochem. Biophys., 2005. **42**(1): p. 7-12.

59. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism.* Biochim. Biophys. Acta, 2005. **1751**(2): p. 119-139.

60. Kery, V., G. Bukovska, and J.P. Kraus, *Transsulfuration depends on heme in addition to pyridoxal 5'-phosphate. Cystathionine beta-synthase is a heme protein.* J. Biol. Chem., 1994. **269**(41): p. 25283-25288.

61. Kühl, T., et al., *Analysis of Fe(III) heme binding to cysteine-containing heme-regulatory motifs in proteins.* ACS Chem. Biol., 2013. **8**(8): p. 1785-1793.

62. Kumaraswami, M., et al., *Structural and biochemical characterization of MepR, a multidrug binding transcription regulator of the Staphylococcus aureus multidrug efflux pump MepA.* Nucleic Acids Res., 2009. **37**(4): p. 1211-1224.

63. Kumarevel, T., *The MarR family of transcriptional regulators - A structural perspective*, in *Antibiotic Resistant Bacteria - a Continuous Challenge in the New Millennium*, M. Pana, Editor. 2012, Intech Europe: Rijeka. p. 403-418.

64. Kumarevel, T., et al., *ST1710-DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators.* Nucleic Acids Res., 2009. **37**(14): p. 4723-35.

65. Lanzilotta, W.N., et al., *Structure of the CO sensing transcription activator CooA.* Nat. Struct. Biol., 2000. **7**(10): p. 876-880.

66.   Lechardeur, D., et al., *Discovery of intracellular heme-binding protein HrtR, which controls heme efflux by the conserved HrtB-HrtA transporter in Lactococcus lactis.* J. Biol. Chem., 2012. **287**(7): p. 4752-4758.

67.   Li, T., H.L. Bonkovsky, and J.T. Guo, *Structural analysis of heme proteins: implications for design and prediction.* BMC Struct. Biol., 2011. **11**: p. 1-13.

68.   Liang, J., et al., *Effective elimination of nucleic acids from bacterial protein samples for optimized blue native polyacrylamide gel electrophoresis.* Electrophoresis, 2009. **30**(14): p. 2454-2459.

69.   Mansuy, D., et al., *Phosphines as ligands to microsomal cytochrome P-450.* Hoppe Seylers Z. Physiol. Chem., 1974. **355**(11): p. 1341-1349.

70.   Martin, S.R. and M.J. Schilstra, *Circular dichroism and its application to the study of biomolecules*, in *Biophysical Tools for Biologists: Vol 1 In Vitro Techniques*, J.J. Correia and H.W. Detrich, Editors. 2008, Elsevier Academic Press Inc: San Diego. p. 263-293.

71.   Marvin, K.A., et al., *The transcription regulator RcoM-2 from Burkholderia xenovorans is a cysteine-ligated hemoprotein that undergoes a redox-mediated ligand switch.* Biochemistry, 2008. **47**(34): p. 9016-9028.

72.   Marvin, K.A., et al., *Nuclear receptors Homo sapiens Rev-erbβ and Drosophila melanogaster E75 are thiolate-ligated heme proteins which undergo redox-mediated ligand switching and bind CO and NO.* Biochemistry, 2009. **48**(29): p. 7056-7071.

73.   MilliporeSigma, *Benzonase® endonuclease*, in *https://www.emdmillipore.com/Web-US-Site/en_CA/-/USD/ShowDocument-Pronet?id=201312.078*. 2020, Merck.

74.   MilliporeSigma. *Benzonase® Nuclease & Enhancers*. https://www.sigmaaldrich.com/life-science/protein-sample-preparation/protein-extraction/benzonase-endonuclease.html 2020 [cited 2020; Available from: https://www.sigmaaldrich.com/life-science/protein-sample-preparation/protein-extraction/benzonase-endonuclease.html.

75.   Mukaiyama, Y., et al., *Spectroscopic and DNA-binding characterization of the isolated heme-bound basic helix-loop-helix-PAS-A domain of neuronal PAS protein 2 (NPAS2), a transcription activator protein associated with circadian rhythms.* Febs. J., 2006. **273**(11): p. 2528-2539.

76.   Muraki, N., C. Kitatsuji, and S. Aono, *A new biological function of heme as a signaling molecule.* J. Porph. Phthalo., 2015. **19**(1-3): p. 9-20.

77.   Muryoi, N., et al., *Demonstration of the iron-regulated surface determinant (Isd) heme transfer pathway in Staphylococcus aureus.* J. Biol. Chem., 2008. **283**(42): p. 28125-28136.

78.   Nakahara, K., et al., *Cytochrome P-450 55A1 (P-450dNIR) acts as nitric-oxide reductase employing NADH as the direct electron-donor.* J. Biol. Chem., 1993. **268**(11): p. 8350-8355.

79.   Newberry, K.J., et al., *Structural mechanism of organic hydroperoxide induction of the transcription regulator OhrR.* Mol. Cell, 2007. **28**(4): p. 652-664.

80.   Ogawa, K., et al., *Heme mediates derepression of Maf recognition element through direct binding to transcription repressor Bach1.* EMBO J., 2001. **20**(11): p. 2835-2843.

81.   Ogura, M., et al., *Redox-dependent axial ligand replacement and its functional significance in heme-bound iron regulatory proteins.* J. Inorg. Biochem., 2018. **182**: p. 238-248.

82.   Otani, H., et al., *The activity of CouR, a MarR family transcriptional regulator, is modulated through a novel molecular mechanism.* Nucleic Acids Res., 2016. **44**(2): p. 595-607.

83.   Owens, C.P., et al., *Characterization of heme ligation properties of Rv0203, a secreted heme binding protein involved in Mycobacterium tuberculosis heme uptake.* Biochemistry, 2012. **51**(7): p. 1518-1531.

84.   Pardee, K.I., et al., *The structural basis of gas-responsive transcription by the human nuclear hormone receptor REV-ERBβ.* PLoS Biol., 2009. **7**(2): p. e43.

85.   Perera, I.C. and A. Grove, *Molecular mechanisms of ligand-mediated attenuation of DNA binding by MarR family transcriptional regulators.* J. Mol. Cell. Biol., 2010. **2**(5): p. 243-254.

86.   Perera, R., et al., *Neutral thiol as a proximal ligand to ferrous heme iron: Implications for heme proteins that lose cysteine thiolate ligation on reduction.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(7): p. 3641-3646.

87.   Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J. Comput. Chem., 2004. **25**(13): p. 1605-1612.

88.   Poulos, T.L., B.C. Finzel, and A.J. Howard, *High-resolution crystal structure of cytochrome P450cam.* J. Mol. Biol., 1987. **195**(3): p. 687-700.

89.   Qin, J., et al., *Structures of thiolate- and carboxylate-ligated ferric H93G myoglobin: models for cytochrome P450 and for oxyanion-bound heme proteins.* Biochemistry, 2006. **45**(10): p. 3170-3177.

90.   Ralph, A.P. and J.R. Carapetis, *Group A streptococcal diseases and their global burden.* Curr. Top. Microbiol. Immunol., 2013. **368**: p. 1-27.

91.   Roach, M.P., et al., *The role of the distal and proximal protein environments in controlling the ferric spin state and in stabilizing thiolate ligation in heme systems:*

Thiolate adducts of the myoglobin H93G cavity mutant. J. Am. Chem. Soc., 1999. **121**(51): p. 12088-12093.

92. Roy, A., et al., *Expression, functional characterization and x-ray analysis of HosA, a member of MarR family of transcription regulator from uropathogenic Escherichia coli.* Protein J., 2016. **35**(4): p. 269-282.

93. Ruf, H.H. and P. Wende, *Hyperporphyrin spectra of ferric dimercaptide-hemin complexes. Models for ferric cytochrome P450-thiol complexes.* J. Am. Chem. Soc., 1977. **99**(16): p. 5499-5500.

94. Sachla, A.J. and Z. Eichenbaum, *The GAS PefCD exporter is a MDR system that confers resistance to heme and structurally diverse compounds.* BMC Microbiol., 2016. **16**(1): p. 1-15.

95. Sachla, A.J., et al., *The crimson conundrum: Heme toxicity and tolerance in GAS.* Front. Cell. Infect. Microbiol., 2014. **4**(159): p. 1-14.

96. Salard-Arnaud, I., et al., *Spectroscopic, catalytic and binding properties of Bacillus subtilis NO synthase-like protein: Comparison with other bacterial and mammalian NO synthases.* J. Inorg. Biochem., 2012. **106**(1): p. 164-171.

97. Salarian, M., et al., *Designing calcium-binding proteins for molecular MR imaging.* Methods Mol. Biol., 2019. **1929**: p. 111-125.

98. Saridakis, V., et al., *Structural insight on the mechanism of regulation of the MarR family of proteins: high-resolution crystal structure of a transcriptional repressor from Methanobacterium thermoautotrophicum.* J. Mol. Biol., 2008. **377**(3): p. 655-667.

99. Sawai, H., et al., *Structural basis for the transcriptional regulation of heme homeostasis in Lactococcus lactis.* J. Biol. Chem., 2012. **287**(36): p. 30755-30768.

100. Schelvis, J.P., et al., *Resonance Raman detection of the Fe-S bond in endothelial nitric oxide synthase.* Biochemistry, 2002. **41**(18): p. 5695-701.

101. Schneider, S., et al., *Diversity and conservation of interactions for binding heme in b-type heme proteins.* Nat. Prod. Rep., 2007. **24**(3): p. 621-630.

102. Schubert, E., et al., *Spectroscopic studies on peptides and proteins with cysteine-containing heme regulatory motifs (HRM).* J. Inorg. Biochem., 2015. **148**: p. 49-56.

103. Shack, J., *The influence of sodium and magnesium ions on the action of deoxyribonuclease II.* J. Biol. Chem., 1959. **234**: p. 3003-3006.

104. Shelver, D., et al., *Identification of two important heme site residues (cysteine 75 and histidine 77) in CooA, the CO-sensing transcription factor of Rhodospirillum rubrum.* Biochemistry, 1999. **38**(9): p. 2669-2678.

105. Shimizu, T., *Binding of cysteine thiolate to the Fe(III) heme complex is critical for the function of heme sensor proteins.* J. Inorg. Biochem., 2012. **108**: p. 171-177.

106. Shimizu, T., et al., *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): p. 6491-6533.

107. Shimizu, T., et al., *Heme: Emergent roles of heme in signal transduction, functional regulation and as catalytic centres.* Chem. Soc. Rev., 2019. **48**(24): p. 5624-5657.

108. Shiro, Y., et al., *Spectroscopic and kinetic studies on reaction of cytochrome P450nor with nitric oxide. Implication for its nitric oxide reduction mechanism.* The Journal of Biological Chemistry, 1995. **270**(4): p. 1617-1623.

109. Sigman, J.A., et al., *Engineering cytochrome c peroxidase into cytochrome P450: A proximal effect on heme-thiolate ligation.* Biochemistry, 1999. **38**(34): p. 11122-11129.

110. Siligardi, G., C.S. Hughes, and R. Hussain, *Characterisation of sensor kinase by CD spectroscopy: Golden rules and tips.* Biochem. Soc. Trans., 2018. **46**: p. 1627-1642.

111. Simonneaux, G., *Phosphines as structural and functional probes of hemoproteins.* Coord. Chem. Rev., 1997. **165**: p. 447-474.

112. Simonneaux, G., A. Bondon, and P. Sodano, *Hydrogen-1 NMR study of iron-bound trimethylphosphine in low-spin ferric complexes of various hemoglobins and myoglobins.* Inorg. Chem., 1987. **26**(21): p. 3636-3638.

113. Smith, A.T., et al., *Identification of Cys94 as the distal ligand to the Fe(III) heme in the transcriptional regulator RcoM-2 from Burkholderia xenovorans.* J. Biol. Inorg. Chem., 2012. **17**(7): p. 1071-1082.

114. Smith, A.T., et al., *Functional divergence of heme-thiolate proteins: A classification based on spectroscopic attributes.* Chem. Rev., 2015. **115**(7): p. 2532-2558.

115. Smith, L.J., A. Kahraman, and J.M. Thornton, *Heme proteins-diversity in structural characteristics, function, and folding.* Proteins, 2010. **78**(10): p. 2349-2368.

116. Sono, M., J.H. Dawson, and L.P. Hager, *The generation of a hyperporphyrin spectrum upon thiol binding to ferric chloroperoxidase. Further evidence of endogenous thiolate ligation to the ferric enzyme.* J. Biol. Chem., 1984. **259**(21): p. 13209-13216.

117. Sono, M., J.H. Dawson, and L.P. Hager, *Phosphine binding as a structural probe of the chloroperoxidase active site: Spectroscopic evidence for endogenous thiolate ligation to the heme iron.* Inorg. Chem., 1985. **24**(25): p. 4339-4343.

118. Sono, M., et al., *Identification of nitric oxide synthase as a thiolate-ligated heme protein using magnetic circular dichroism spectroscopy. Comparison with cytochrome P-450-CAM and chloroperoxidase.* J. Biol. Chem., 1995. **270**(34): p. 19943-19948.

119.   Sono, M., et al., *Spectroscopic evidence supporting neutral thiol ligation to ferrous heme iron.* J. Biol. Inorg. Chem., 2018. **23**(7): p. 1085-1092.

120.   Spiro, T.G. and T.C. Strekas, *Resonance Raman spectra of heme proteins - Effects of oxidation and spin state.* J. Am. Chem. Soc., 1974. **96**(2): p. 338-345.

121.   Stauff, D.L. and E.P. Skaar, *Bacillus anthracis HssRS signalling to HrtAB regulates haem resistance during infection.* Mol. Microbiol., 2009. **72**(3): p. 763-778.

122.   Stevenson, C.E.M., et al., *Investigation of DNA sequence recognition by a streptomycete MarR family transcriptional regulator through surface plasmon resonance and X-ray crystallography.* Nucleic Acids Res., 2013. **41**(14): p. 7009-7022.

123.   Strekas, T.C. and T.G. Spiro, *Resonance Raman evidence for anomalous heme structures in cytochrome c' from Rhodopseudomonas palustris.* Biochim. Biophys. Acta, 1974. **351**(2): p. 237-245.

124.   Sugita, Y. and K. Yoneyama, *Oxygen equilibrium of hemoglobins containing unnatural hemes.* J. Biol. Chem., 1971. **246**: p. 389-394.

125.   Sun, S.F., M. Sono, and J.H. Dawson, *Mono- and bis-phosphine-ligated H93G myoglobin: Spectral models for ferrous-phosphine and ferrous-CO cytochrome P450.* J. Inorg. Biochem., 2013. **127**: p. 238-245.

126.   Thermo Fisher Scientific *DNase I Demystified*. DNase, a powerful research tool for DNA manipulations, 2020. DOI: https://www.thermofisher.com/us/en/home/references/ambion-tech-support/nuclease-enzymes/general-articles/dnase-i-demystified.html.

127.   Thompson, S., *Biophysical heme binding studies of Corynebacterium diphtheriae and Streptococcus pyogenes*, in *Chemistry (Easton)*. 2017, Georgia State University: Atlanta, GA. p. 80.

128.   Tullius, M.V., et al., *Discovery and characterization of a unique mycobacterial heme acquisition system.* Proc. Natl. Acad. Sci. U. S. A., 2011. **108**(12): p. 5051-5056.

129.   Usuda, K., et al., *Denitrification by the fungus cylindrocarpon tonkinense: Anaerobic cell-growth and two isozyme forms of cytochrome P-450nor.* Appl. Environ. Microbiol., 1995. **61**(3): p. 883-889.

130.   Valenzuela, J.G., F.A. Walker, and J.M. Ribeiro, *A salivary nitrophorin (nitric-oxide-carrying hemoprotein) in the bedbug Cimex lectularius.* J. Exp. Biol., 1995. **198**(Pt 7): p. 1519-1526.

131.   Walker, M.J., et al., *Disease manifestations and pathogenic mechanisms of Group A Streptococcus* Clin. Microbiol. Rev., 2014. **27**(2): p. 264-301.

132.	Wardell, M., et al., *The atomic structure of human methemalbumin at 1.9 Å.* Biochem. Biophys. Res. Commun., 2002. **291**(4): p. 813-819.

133.	Weichsel, A., et al., *Heme-assisted S-nitrosation of a proximal thiolate in a nitric oxide transport protein.* Proc. Natl. Acad. Sci. U. S. A., 2005. **102**(3): p. 594-599.

134.	Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data.* Nucleic Acids Res., 2004. **32**(Web Server issue): p. W668-73.

135.	Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases.* Biopolymers, 2008. **89**(5): p. 392-400.

136.	Wilkinson, S.P. and A. Grove, *Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins.* Curr. Issues Mol. Biol., 2006. **8**: p. 51-62.

137.	Woody, R.W., *The development and current state of protein circular dichroism.* Biomedical Spectroscopy and Imaging, 2015. **4**(1): p. 5-34.

138.	Woody, R.W. and G. Pescitelli, *The role of heme chirality in the circular dichroism of heme proteins.* Z. Naturforsch. A, 2014. **69**(7): p. 313-325.

139.	Wu, R.Y., et al., *Crystal structure of Enterococcus faecalis SlyA-like transcriptional factor.* J. Biol. Chem., 2003. **278**(22): p. 20240-20244.

140.	Yang, J., et al., *A novel heme-regulatory motif mediates heme-dependent degradation of the circadian factor period 2.* Mol. Cell. Biol., 2008. **28**(15): p. 4697-4711.

141.	Yang, J., A. Roy, and Y. Zhang, *BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions.* Nucleic Acids Res., 2013. **41**(Database issue): p. D1096-1103.

142.	Yang, J., A. Roy, and Y. Zhang, *Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment.* Bioinformatics, 2013. **29**(20): p. 2588-2595.

143.	Yang, J.H., K. Ishimori, and M.R. O'Brian, *Two heme binding sites are involved in the regulated degradation of the bacterial iron response regulator (Irr) protein.* J. Biol. Chem., 2005. **280**(9): p. 7671-7676.

144.	Yang, J.Y., et al., *The I-TASSER Suite: Protein structure and function prediction.* Nature Methods, 2015. **12**(1): p. 7-8.

145.	Yin, L., V. Dragnea, and C.E. Bauer, *PpsR, a regulator of heme and bacteriochlorophyll biosynthesis, is a heme-sensing protein.* J. Biol. Chem., 2012. **287**(17): p. 13850-13858.

146.   Yoshida, M., et al., *Ligand specificity of MobR, a transcriptional regulator for the 3-hydroxybenzoate hydroxylase gene of Comamonas testosteroni KH122-3s.* Biochem. Biophys. Res. Commun., 2007. **362**(2): p. 275-280.

147.   Yu, L.L., J. Fang, and Y.N. Wei, *Characterization of the ligand and DNA binding properties of a putative archaeal regulator ST1710.* Biochemistry, 2009. **48**(10): p. 2099-2108.

148.   Zhang, C., P.L. Freddolino, and Y. Zhang, *COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information.* Nucleic Acids Res., 2017. **45**(W1): p. W291-W299.

149.   Zhang, L. and L. Guarente, *Heme binds to a short sequence that serves a regulatory function in diverse proteins.* EMBO J., 1995. **14**(2): p. 313-320.

150.   Zhu, R.F., et al., *Structural characterization of the DNA-binding mechanism underlying the copper(II)-sensing MarR transcriptional regulator.* J. Biol. Inorg. Chem., 2017. **22**(5): p. 685-693.