Structural and Mechanistic Insights into Iron Acquisition from Human Hemoglobin by Gram-positive Pathogens with a Focus on Two Organisms: *Streptococcus pyogenes* and *Staphylococcus aureus*.

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biochemistry and Molecular Biology

by

Ramsay Aristotle Macdonald

2018
© Copyright by
Ramsay Aristotle Macdonald
2018
ABSTRACT OF THE DISSERTATION

Structural and Mechanistic Insights into Iron Acquisition from Human Hemoglobin by Gram-positive Pathogens with a Focus on Two Organisms: *Streptococcus pyogenes* and *Staphylococcus aureus*.

by

Ramsay Aristotle Macdonald

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2018

Professor Robert Thompson Clubb, Chair

*Staphylococcus aureus* and other clinically important species of Gram-positive bacteria have acquired resistance to commonly used antibiotics. This underscores the need for new therapeutics and a greater understanding of the molecular mechanisms that enable bacterial pathogens to cause disease. Iron acquisition systems are an attractive target for new antibiotics as many microbial pathogens must acquire this metal from their human host in order to establish an infection. The majority of the body’s iron supply (~60-80%) is found in the oxygen transporter hemoglobin (Hb) in the form of heme (iron + protoporphyrin IX). Many Gram-positive pathogens have therefore evolved mechanisms to acquire heme-iron form Hb. In Chapter 1, I survey our current understanding of the heme-acquisition strategies that are employed by different species of Gram-positive bacteria, especially those found in *Staphylococcus aureus* and *Streptococcus pyogenes* as they are the focus of my research. I further concentrate on the first step in the heme acquisition process, the extraction of heme from human Hb by receptors displayed on the surface of each microbe: Shr in *S. pyogenes* and IsdH in *S. aureus*. In chapters 2 and 3, structural and
biochemical studies of the N-terminal region within the Shr receptor are presented. We show that
the N-terminal region in Shr contains three autonomously folded domains. Two of these domains
engage Hb and are renamed Hemoglobin Interacting Domains (HIDs). We present the crystal
structure of HID2 which reveals a structurally unique Hb-binding domain. In Chapter 4, the
kinetics and thermodynamics of heme extraction from Hb by the S. aureus IsdH are presented.
Using this data and molecular dynamics simulations, a step-by-step reaction mechanism of the
heme extraction process is proposed. In Chapter 5, the NMR-derived solution structure and
backbone dynamics of the C-terminal domain of PhoP are reported. PhoP is part of the PhoPR
two-component system in Mycobacterium tuberculosis, the causative agent of tuberculosis. This
work contributes to an understanding of the mechanism by which PhoP regulates gene expression.
The dissertation of Ramsay Aristotle Macdonald is approved.

David S Eisenberg

Feng Guo

Joseph Ambrose Loo

Robert Thompson Clubb, Committee Chair

University of California, Los Angeles

2018
DEDICATION

I would like to dedicate this work to my mother, father, and sister who all supported me over the years and to my beautiful wife whom I met here in the program.
# TABLE OF CONTENTS

Abstract of the Dissertation...........................................................................................................iii
Committee Page.............................................................................................................................v
Dedication Page.............................................................................................................................vi
Table of Contents.......................................................................................................................vii
List of Figures.............................................................................................................................xii
List of Tables...............................................................................................................................xiv
Acknowledgements..................................................................................................................xv
Vita................................................................................................................................................xviii
Publications....................................................................................................................................xix

## CHAPTER 1. Introduction........................................................................................................1

1.1 Overview..............................................................................................................................2
1.2 Introduction to Iron Acquisition Strategies by Gram-positive Pathogens.........................3
   1.2.1 Iron Acquisition by *Staphylococcus aureus*..............................................................5
   1.2.2 Iron Acquisition by *Streptococcus pyogenes*...........................................................9
   1.2.3 Iron Acquisition by *Bacillus anthracis*.................................................................10
   1.2.4 Iron Acquisition by *Corynebacterium diphtheriae*.................................................13
   1.2.5 Iron Acquisition Systems found in other Gram-positive Pathogens........................14
1.3 Scope of Dissertation........................................................................................................17
1.4 Figures...............................................................................................................................19
1.5 References........................................................................................................................25
CHAPTER 2. The *Streptococcus pyogenes* Shr Protein Captures Human Hemoglobin Using Two Structurally Unique Binding Domains………40

2.1 Overview……………………………………………………………………………………………42

2.2 Introduction……………………………………………………………………………………43

2.3 Results…………………………………………………………………………………………46

2.3.1 The Hb interacting N-terminal region within Shr contains three autonomously folded domains…………………………………………………………………………..46

2.3.2 DUF1533 domains bind Hb……………………………………………………...47

2.3.3 Two Shr$^{\text{HID1-2}}$ proteins bind to tetrameric Hb…………………………………49

2.3.4 The HID is a structurally novel Hb-binding domain………………………………50

2.3.5 The HIDs within Shr$^{\text{HID1}}$ and Shr$^{\text{HID2}}$ engage Hb using a conserved tyrosine residue……………………………………………………………………………………………52

2.4 Discussion………………………………………………………………………………………53

2.5 Experimental Procedures……………………………………………………………………58

2.5.1 Cloning, Protein Expression, and Purification…………………………………………58

2.5.2 NMR Spectroscopy…………………………………………………………………………59

2.5.3 Isothermal Titration Calorimetry…………………………………………………………60

2.5.4 Size Exclusion Chromatography with In-line Multi-Angle Light Scattering……61

2.5.5 Analytical Ultracentrifugation……………………………………………………………61

2.5.6 X-ray Crystallography……………………………………………………………………62

2.6 Figures……………………………………………………………………………………………64

2.7 Supplemental Material………………………………………………………………………75

2.8 References……………………………………………………………………………………78
CHAPTER 3. The *Streptococcus pyogenes* Shr$^{\text{HID2}}$ Domain in Complex with Hemoglobin Exhibits a Novel Mechanism of Capture within the Field of Gram-positive Bacterial Pathogens..........................................................91

3.1 Overview...........................................................................................................93
3.2 Introduction.......................................................................................................94
3.3 Results...............................................................................................................97
   3.3.1 The Hb:Shr$^{\text{AHID2}}$ Complex Exhibits an Unusual Interface..................97
   3.3.2 Identification of Important Residues in the Hb:Shr$^{\text{AHID2}}$ Interaction.........100
   3.3.3 Shr$^{\text{AHID2}}$ slows hemin release from Hb.............................................101
3.4 Discussion.........................................................................................................101
3.5 Future Directions............................................................................................104
3.6 Experimental Procedures.................................................................................105
   3.6.1 Cloning, Protein Expression, and Purification.............................................105
   3.6.2 X-ray Crystallography..............................................................................106
   3.6.3 NMR Spectroscopy..................................................................................107
   3.6.4 Heme Transfer Experiments.................................................................107
3.7 Figures...............................................................................................................108
3.8 References.........................................................................................................114

CHAPTER 4. Energetics Underlying Hemin Extraction from Human

Hemoglobin by *Staphylococcus aureus*..............................................................122

4.1 Overview.........................................................................................................124
4.2 Introduction.......................................................................................................................125
4.3 Results.............................................................................................................................128
  4.3.1 Development of a quantitative assay to study hemin extraction.........................128
  4.3.2 Hb tetramerization slows hemin removal from the α-subunit.........................130
  4.3.3 IsdH extracts hemin from both the α- and β-subunits in an effectively
       irreversible process...........................................................................................................132
  4.3.4 Determination of the micro-rate constants describing hemin removal from the
       α-subunit..........................................................................................................................134
  4.3.5 An enthalpic barrier must be surmounted to capture hemin............................136
  4.3.6 Two energetically distinct receptor-Hb interfaces are used to bind Hb and
       distort its hemin pocket..................................................................................................137
  4.3.7 Molecular dynamics simulations reveal the receptor promotes hemin
       Solvation..........................................................................................................................138
4.4 Discussion.......................................................................................................................140
4.5 Experimental Procedures...............................................................................................148
  4.5.1 Cloning and protein preparation...............................................................................149
  4.5.2 Hemin transfer experiments.....................................................................................151
  4.5.3 Analytical Ultracentrifugation................................................................................152
  4.5.4 Isothermal Titration Calorimetry............................................................................153
  4.5.5 Molecular dynamic simulations..............................................................................154
4.6 Figures.............................................................................................................................156
4.7 References.......................................................................................................................170
CHAPTER 5. Solution Structure of the PhoP DNA-Binding Domain from *Mycobacterium tuberculosis* .................................................................184

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Overview</td>
<td>186</td>
</tr>
<tr>
<td>5.2</td>
<td>Introduction</td>
<td>187</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>188</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>190</td>
</tr>
<tr>
<td>5.5</td>
<td>Experimental Procedures</td>
<td>193</td>
</tr>
<tr>
<td>5.5.1</td>
<td>Protein expression, purification, and NMR sample preparation</td>
<td>193</td>
</tr>
<tr>
<td>5.5.2</td>
<td>NMR spectroscopy, data collection, and chemical shift assignments</td>
<td>194</td>
</tr>
<tr>
<td>5.5.3</td>
<td>NMR structure determination</td>
<td>195</td>
</tr>
<tr>
<td>5.6</td>
<td>Figures</td>
<td>198</td>
</tr>
<tr>
<td>5.7</td>
<td>References</td>
<td>203</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1 The *S. aureus* Isd system .................................................................19
Figure 1.2 IsdB and IsdH Share a Conserved Tri-Domain Unit .........................20
Figure 1.3 Crystal Structure of IsdH\textsuperscript{N2N3-F365Y/A369F/Y642A} in Complex with metHb ........21
Figure 1.4 The Heme Acquisition Pathway in *S. pyogenes* ...............................22
Figure 1.5 The Isd System in *B. anthracis* .........................................................23
Figure 1.6 The Iron Acquisition System in *C. diphtheriae* ....................................24
Figure 2.1 Domain arrangement within Shr and the protein constructs used in this study .64
Figure 2.2 Hb binding studies by NMR .................................................................65
Figure 2.3 Biochemical Studies of Hb binding .......................................................66
Figure 2.4 Hb binding competition experiment .....................................................68
Figure 2.5 Assignments and Dynamics Data on Shr\textsuperscript{HID2} .........................69
Figure 2.6 Crystal structure of Shr\textsuperscript{AHID2} .............................................70
Figure 2.7 Domain topology of Shr\textsuperscript{AHID2} vs. NEAT domains ...................71
Figure 2.8 Hb titration experiments with Shr\textsuperscript{HID1} and Shr\textsuperscript{HID2} mutants by NMR ...72
Figure S2.1 Spectral recapitulation of the NTR .....................................................75
Figure S2.2 Sequence alignment of DUF 1533 domains ....................................76
Figure 3.1 The Crystal Structure of the Shr\textsuperscript{AHID2}:Hb Complex ...................108
Figure 3.2 The Shr\textsuperscript{AHID2}:Hb Interface ................................................109
Figure 3.3 Hb binding studies with Shr\textsuperscript{HID2} .............................................110
Figure 3.4 Shr\textsuperscript{AHID2} slows the rate of hemin release from Hb .................111
Figure 4.1 *S. aureus* uses conserved tridomain containing receptors to capture Hb ....156
Figure 4.2  Stoichiometry of Hb:receptor complexes determined by analytical ultracentrifugation………………………………………………………………………………158

Figure 4.3  Hemin transfer by various Hb and receptors species………………………………………159

Figure 4.4  Hemin transfer from isolated β-globin chains…………………………………………………160

Figure 4.5  Hemin transfer to myoglobin and as function of receptor concentration…………………161

Figure 4.6  Temperature dependence of hemin transfer…………………………………………………..162

Figure 4.7  IsdH used two distinct interfaces to promote hemin transfer…………………………………163

Figure 4.8  Results of molecular dynamics simulations……………………………………………………164

Figure 4.9  Model of the hemin extraction mechanism……………………………………………………165

Figure 5.1  NMR spectra of PhoPC…………………………………………………………………………….198

Figure 5.2  NMR structure of PhoPC and structural comparisons…………………………………………199

Figure 5.3  NMR relaxation, chemical shift and distance restraint data………………………………201
LIST OF TABLES

Table 2.1  Thermodynamic Parameters and Affinity Data for Hb Binding..................73
Table 2.2  Structure Statistics on Shr^{AHID2}..................................................74
Table S2.1  Stoichiometry of Hb:Shr^{HID2} complex determined by analytical
Ultracentrifugation.................................................................77
Table 3.1  Structure Statistics on the Complex....................................................112
Table 3.2  Interactions at the Shr^{AHID2}:Hb Interface........................................113
Table 4.1  Stoichiometry of Hb:receptor complexes determined by analytical
ultracentrifugation.................................................................166
Table 4.2  Apparent rate constants of hemin transfer..........................................168
Table 4.3  Thermodynamic parameters and affinity data for Hb binding...............169
Table 5.1  Statistics for the solution structure of PhoPC ....................................202
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Prof. Robert T. Clubb for accepting me in his laboratory to perform my doctoral research, for his support and guidance over the years, and for instilling in me the confidence and scientific acumen to tackle all future endeavors. I would like to thank Dr. Robert Peterson for his assistance with NMR. My committee members: Prof. David Eisenberg, Prof. Feng Guo, Prof. Joseph Loo for their support, critique, and guidance over the years. Past and present members of the laboratory, Dr. Brendan Amer, Dr. Albert Chan, Dr. Ken Ellis-Guardiola, Dr. Alex Jacobitz, Dr. Michele Kattke, Dr. Brendan Mahoney, Dr. Reza Malmirchegini, Dr. Megan Sjodt, Jason Gosschalk, Scott McConnell, Orlando Martinez, Christopher Sue. My rotation students for their contributions: Colleen Hui, Dr. Kanishk Jain, and Jessica Ochoa. I would like to thank all those who mentored me while I was an undergraduate student: Prof. Maik Hüttemann, Dr. Icksoo Lee, Dr. Petr Pecina, Prof. Ursula Jacob, Dr. Winnie Wholey, Dr. Claudia Cremers, Prof. Erik Zuiderweg, Dr. Atta Ahmad, Dr. Akash Battacharya. I would also like to thank those who mentored me while I was a rotation student: Prof. Juli Feigon, Dr. Catherine Eichhorn, Prof. Mahavir Singh, Prof. Yaqiang Wang, Prof. Wayne Hubbell, Dr. Michael Lerch, Evan Brooks, Prof. David Eisenberg, Dr. Angela Soriaga. I would like to Dr. Elaheh Binabaji and Dr. Alex Jacobitz for their mentorship and guidance during my graduate student co-op at Amgen.

The work presented in this dissertation was supported by the Cellular and Molecular Biology Training Grant.

Chapter two of this dissertation is a version of the publication currently under revision titled: The Streptococcus pyogenes Shr Protein Captures Human Hemoglobin Using Two

Chapter three of this dissertation is a version of the manuscript in preparation titled: The Streptococcus pyogenes ShrHID2 Domain in Complex with Hb Exhibits a Novel Mechanism of Capture within the Field of Gram-positive Bacterial Pathogens. Ramsay Macdonald, Duilio Cascio, Robert T. Clubb., 2018.

Chapter four of this dissertation is a version of the publication: Megan Sjodt, Ramsay Macdonald, Joanna D. Marshall, Joseph Clayton, John S. Olson, Martin Phillips, David A. Gell, Jeff Wereczynski and Robert T. Clubb (2018). Energetics Underlying Hemin Extraction from Human Hemoglobin by Staphylococcus aureus. *Journal of Biological Chemistry*, 293(18):6942-6957. jbc.RA117.000803. [http://doi.org/10.1074/jbc.RA117.000803](http://doi.org/10.1074/jbc.RA117.000803). Reproduced with permission. Copyright 2018, Journal of Biological Chemistry. This work was supported by National Science Foundation Grant MCB-1716948 (to R.T.C. and J.W.) and National Institutes of Health Grants AI52217 (to R.T.C.), R35GM119647 (to. J.W.), F31GM101931 (to M.S.) and P01-HL110900 (to J.S.O.). J.S.O. was also supported by a grant from the Robert A. Welch Foundation (C-0612). R.M. was supported by a Cellular and Molecular Biology Training Grant (Ruth L. Kirschstein National Research Service Award GM007185). NMR equipment used in this research was purchased using funds from shared equipment grant NIH S10OD016336.

Chapter five of this dissertation is a version of the publication: Ramsay Macdonald, Dibyendu Sarkar, Brendan R. Amer, & Robert T. Clubb (2015). Solution Structure of the PhoP DNA-Binding Domain from Mycobacterium tuberculosis. *Journal of Biomolecular NMR*, 63(1):111-7. s10858-015-9965-0. [https://doi.org/10.1007/s10858-015-9965-0](https://doi.org/10.1007/s10858-015-9965-0). Reproduced with permission. Copyright 2015, Journal of Biomolecular NMR. This work was supported by
the National Institutes of Health grant AI52217 to RTC. R.M. was supported by a Cellular and Molecular Biology Training Grant (Ruth L. Kirschstein National Research Service Award GM007185). D.S. was supported by Raman Research Fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India. B.R.A. was supported by a Whitcome Predoctoral Training Grant, University of California – Los Angeles, Molecular Biology Institute. This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421.
VITA

2007-2010  Bachelors of Science, Biochemistry
            University of Michigan, Ann Arbor

2012-2018  Doctoral Research, Clubb Laboratory
            University of California, Los Angeles

2015-2018  Cellular and Molecular Biology Training Grant

2017       Graduate Student Co-op in Process Development
            Amgen

2016-2017  Graduate Student Instructor
            California NanoSystems Institute at UCLA


Chapter 1

Introduction
1.1 Overview

Iron acquisition from the human host by Gram-positive bacteria is a critical step for pathogenesis. Most of the body’s iron supply exists in the oxygen transporter Hb in the form of heme. This introduction surveys the strategies employed by various Gram-positive bacterial pathogens to acquire iron from human Hb. Special attention is paid to *Staphylococcus aureus* and *Streptococcus pyogenes*, as my thesis research focuses on characterizing the Hb receptors in these organisms.
1.2 **Introduction to Iron Acquisition Strategies by Gram-positive Pathogens**

Gram-positive and Gram-negative bacteria use fundamentally different mechanisms to acquire iron from human Hb. The focus of my work is on Gram-positive pathogens as, until recently, the iron acquisition systems employed in these microbes were less well characterized. Gram-positive bacterial pathogens cause a range of infections in humans and can often be lethal. In some cases, conventional antibiotics can still be effectively used to treat infections caused by certain species of Gram-positive pathogens, such as *S. pyogenes* (1). However, *S. aureus* has evolved resistance to almost every known antibiotic through horizontal gene transfer, chromosomal mutation and antibiotic selection (2). In particular, Methicillin-resistant *S. aureus* (MRSA) has emerged as a significant threat to public health (3). It causes thousands of deaths each year and the annual economic burden exceeds one billion dollars in the United States alone (3, 4). This underscores the need for a greater understanding of the mechanisms of bacterial pathogenesis, which could facilitate the development of novel therapeutics. The iron acquisition systems employed by bacteria are a potential drug target, as iron is an essential nutrient that must be actively procured during infections.

Virtually all organisms require iron for growth, as it is a key element that is used in a number of biological processes: oxygen transport, DNA synthesis, and cellular respiration (5–7). Bacteria require iron at concentrations of 0.4 – 4.0 μM to grow (8). However, iron levels are strictly regulated within the human body and the concentration of free iron outside cells is generally only $10^{-18}$ M (9). This is well below the concentration needed by the microbe for growth. Indeed, iron is the only nutrient limiting bacterial growth inside the human host (8). The majority of iron in the human body is sequestered intracellularly and is generally unavailable to the invading pathogen. In fact, this intracellular iron sequestration provides a form of “nutritional
immunity” (10). Most of the body’s iron (~80% of the total) is found in the oxygen transporter hemoglobin (Hb) in the form of heme (protoporphyrin IX + iron) (9). Hb is found within erythrocytes, which contain as many as 280 million molecules of Hb or roughly ~1 billion atoms of iron (10). Many Gram-positive pathogens have evolved protein-based systems to specifically acquire iron from this nutrient-rich source.

Erythrocytes have a typical lifespan of ~120 days (11). During normal iron homeostasis, aged erythrocytes are phagocytosed by macrophages; hemoglobin is then proteolyzed releasing the heme molecule, which is subsequently degraded by enzymes known as heme oxygenases to release iron for recycling (7, 11). Alternatively, erythrocytes can also rupture due to various stress conditions such as vigorous exercise (11). In this scenario, Hb is released into the plasma and dissociates from the tetrameric (α2β2) form into the dimeric form (αβ) due to its dimer/tetramer equilibrium K_{4,2} = 2 μM (11–13). Dimeric Hb is then bound by haptoglobin (Hp), which protects tissue against Hb-stimulated lipid peroxidation (13–15). The Hb-Hp complex is then taken up by macrophages via the CD163 receptor (16). If Hb is not bound by Hp, its heme moieties readily autoxidize to hemin (Fe^{3+}-protoporphyrin IX) and are subsequently released from the (αβ) dimer (17). This form of Hb is known as metHb. At this point free hemin is bound by hemopexin and the hemopexin-hemin complex is then taken up by macrophages via the CD91 receptor (18). Once inside the macrophage, the tetrapyrrole ring is degraded releasing the free iron which is then recycled (11).

During an infection, many bacterial pathogens secrete hemolysins that lyse erythrocytes (19, 20) to enable Hb or Hb-Hp to be utilized as an iron source. In this review, I discuss the specific mechanisms used by Gram-positive pathogens to acquire heme from Hb and the Hb-Hp complex. Specifically, I discuss the heme acquisition systems used by S. aureus, S. pyogenes,
Bacillus anthracis (B. anthracis), Listeria monocytogenes (L. monocytogenes), Corynebacterium diphtheriae (C. diphtheriae), Bacillus cereus (B. cereus), Streptococcus pneumoniae (S. pneumoniae), and Staphylococcus lugdunensis (S. lugdunensis).

1.2.1 Iron Acquisition by Staphylococcus aureus

The heme iron acquisition system used by S. aureus has been characterized in the greatest detail. It is required for pathogenesis and heme has been shown to be S. aureus’ preferred iron source during the initial stages of infection (21). In cell cultures, S. aureus preferentially uses Hb as an iron source (22). Heme is acquired from Hb using the iron-regulated surface determinant (Isd) system, which consists of nine S. aureus proteins (IsdABCDEFGHI) that are expressed under iron limiting conditions. The pathway is shown in Figure 1.1.

Heme acquisition is initiated by the surface exposed IsdB and IsdH proteins, which function as Hb receptors (23, 24). These receptors are then thought to pass heme to IsdA which is partially buried in the cell wall. Holo-IsdA then transfers hemin to apo-IsdC, which is thought to be located proximal to the membrane based on its resistance to proteolysis. Transfer occurs very rapidly, at a rate that is ~70,000 fold faster than the simple dissociation of hemin from IsdA (31). Apo-IsdE then acquires hemin from holo-IsdC (28), enabling the IsdDEF transporter to transfer the hemin molecule across the cell membrane. In the cytoplasm, the tetrapyrrole ring of the hemin molecule is then degraded by the IsdG and IsdI oxygenases to release free iron (32). In vitro biochemical studies of the isolated proteins have demonstrated that the flow of hemin occurs in the following manner: IsdHN3/IsdBN2 -> IsdAN -> IsdCN -> IsdE (28, 33). The specific functions of each of the Isd proteins in the S. aureus iron acquisition system will now be discussed.
Substrate recognition in IsdH, IsdB, IsdA, and IsdC is carried out via NEAr-iron Transporter (NEAT) domains (34). IsdH, IsdB, IsdA, and IsdC contain three, two, one and one NEAT domain(s), respectively (26, 27). Structures of all these NEAT domains in isolation have been reported (26, 27, 35–41) revealing a conserved immunoglobulin-like fold that is comprised of β-strands arranged into two sheets: (β1a-β2-β3-β6-β5) and (β4-β7-β8-β1b). However, there are slight differences in structure depending on the function of the NEAT domain: Hb vs hemin binding. Hemin-binding NEAT domains contain a short 310-helix perpendicular to the β7-β8 hairpin that forms part of the hemin-binding pocket. These domains also usually contain a conserved YXXXY motif within the β8-strand wherein the hydroxyl of the first tyrosine residue in the motif coordinates the iron atom in the hemin molecule (36, 38–41). The hydroxyl within the second tyrosine hydrogen bonds to the first and helps to stabilize the interaction with the iron atom in hemin (39, 40). In most cases, both of these conserved tyrosines have been shown to be essential for hemin binding (39, 42). In contrast, Hb-binding NEAT domains lack this conserved motif and instead contain conserved aromatic residues located in the extended loop between β1b and β2 that has been shown to contact Hb (26, 30, 37, 43).

The IsdH and IsdB receptors use a conserved tri-domain structure to extract hemin from Hb (Figure 1.2). It consists of two NEAT domains separated by an α-helical linker, and is called IsdH^{N2N3} and IsdB^{N1N2} in IsdH and IsdB, respectively (29). Both receptors accelerate hemin release from Hb at least ~580 fold faster than the spontaneous rate of hemin release from Hb (29). In the case of IsdH^{N2N3}, all three of the domains in this tri-domain receptor must be expressed in the same polypeptide for the accelerated kinetics to be observed (29). This contrasts with IsdB^{N1N2}, as a polypeptide that contains only the N1 and linker domains and a polypeptide containing only the N2 domain can work in \textit{in trans} to accelerate hemin release from Hb (37,
This finding is supported by Hb-binding studies which demonstrate that only IsdB_N1N2 binds Hb with high affinity, while IsdH_N2 binds Hb with similar affinity to IsdH_N2N3 (44, 45).

Interestingly, a region at the N-terminus of IsdB (NS), before the N1 domain in the primary sequence, is necessary to achieve the maximum rate of hemin removal from Hb (46). Without this NS region, IsdB_N1N2 exhibits a 10 fold decrease in the rate of hemin extraction from Hb (46). Furthermore, IsdB_{NS-N1-L} and IsdB_N2 in trans are able to remove hemin from Hb at the same rate as IsdB_{NS-N1-L-N2} (46). The corresponding region in IsdH has not been studied. Clearly there are mechanistic differences between IsdH_{N2N3} and IsdB_{N1N2} that are not fully understood.

A crystal structure of IsdH_{N2N3-Y642A} in complex with Hb reveals that all three domains in the receptor engage the very same subunit of Hb (Figure 1.3) (35). Studies of the IsdH_{N2N3-Y642A} receptor in solution reveal that the receptor undergoes a conformational change upon binding Hb (47). The linker and N3 domains form a rigid unit and are connected by a hydrophobic interface, while the N2 domain samples a range of motions relative to the linker and N3 domains (47). The N2 and linker domains are connected by a hydrophilic interface and the N2 domain reorients with respect to the Linker and N3 domains upon binding Hb (47). It has been suggested that this “plasticity” in the N2-linker interface is critical for function (47). In effort to achieve a higher resolution crystal structure of the complex between IsdH_{N2N3} and Hb, the N2 domain was mutated to more closely resemble the N1 domain so as to abolish the weak binding to the ß-subunit of Hb (48). Once this was achieved, a higher resolution crystal structure of IsdH_{N2N3-F365Y/A369F/Y642A} in a pre-transfer complex with the α-subunit of Hb was obtained (48). The structure revealed that the linker and N3 domains of IsdH distort the F-helix of Hb which contains the proximal histidine coordinating the hemin’s iron atom (48). This presumably weakens the affinity of Hb for hemin as the hemin binding pocket of Hb is perturbed (48). A
translation of ~12 Å and a rotation of ~115° in the plane of the porphyrin ring are necessary for the transfer of hemin from Hb to the N3 domain (35, 48). MD simulations of the hemin transfer process between Hb and IsdH_N3 reveal that the hemin molecule may occupy an “intermediate” position which may be rate limiting (49).

A crystal structure of the IsdB_{N1N2} in complex with Hb was obtained (44). Poor electron density between the IsdB_{N1} and IsdB_{Linker} domains indicates flexibility in that region, while well ordered-density observed between the IsdB_{Linker} and IsdB_{N2} domains is likely indicative of a well-ordered region (30). This finding would suggest that the IsdB_{N1} domain is more mobile with respect to the rest of the tri-domain receptor while the IsdB_{Linker} and IsdB_{N2} domains form a more rigid unit as is the case for IsdH_{N2N3} (47). As is the case for IsdH, a conformational change in αHb upon IsdB_{N1N2} binding is observed. The F-helix and part of the E-helix are unwound and the hemin molecule is displaced ~5 Å out of its position in Hb’s hemin binding pocket towards the IsdB_{N2} domain (30).

While the functions of IsdH_{N2N3} and IsdB_{N1N2} seem to be redundant, in vitro and in vivo data suggest that IsdB, and not IsdH, is the primary Hb receptor in S. aureus (22). This finding would seem to indicate that IsdH is superfluous in the S. aureus iron acquisition system. However, recent studies have demonstrated that IsdH_{N2N3} is able to accelerate hemin release from the metHb-Hp complex, whereas IsdB_{N1N2} cannot (27, 30, 50). And unlike IsdB_{N1N2}, IsdH_{N1N2N3} blocks Hb-Hp uptake by the CD163 receptor thereby preventing clearance of this iron source (50). Taken together, these data suggest that IsdB is the primary receptor for Hb, while IsdH is the primary receptor for the Hb-Hp complex.

The remaining NEAT containing proteins in the Isd-system, IsdA and IsdC, are embedded in the cell wall and only bind to hemin. IsdA binds hemin via its single NEAT domain
(IsdA<sup>N</sup>) (39, 51–53). IsdC also binds hemin via its NEAT domain (IsdC<sup>N</sup>) (26, 52–54). Hemin transfer between IsdA<sup>N</sup> and IsdC<sup>N</sup> is accomplished via a weak affinity “hand-clasp” association between the NEAT domains whereby the 3<sub>10</sub>-helices and β7-β8 loops of the two domains are juxtaposed against one another (55). The hemin transfer step from IsdC to IsdE is slower than IsdA to IsdC and seems to be the rate limiting step for importing hemin into the bacterial cytoplasm (56). The purpose of this may be to regulate the amount of hemin being delivered to the bacterial cytoplasm as excessive hemin buildup in the cytoplasm can be toxic (56, 57).

1.2.2 Iron Acquisition by *Streptococcus pyogenes*

In *Streptococcus pyogenes*, heme acquisition is mediated by proteins encoded by the Streptococcus iron acquisition (Sia) locus, which contains 10 genes (*shr-shp-siaA-siaB-siaC-1791-1790-1789-1788-1787*) (58, 59). The pathway is shown in Figure 1.4. Shr and Shp are displayed on the surface of the cell wall in *S. pyogenes* and are anchored via C-terminal transmembrane regions (58, 60). A Δshr mutant of *S. pyogenes* exhibits attenuated virulence in zebrafish and murine models of infection (60, 61). When the Δshr mutation is introduced into *S. pyogenes*, it is no longer able to grow in human blood (61). Furthermore, Shr has been implicated as a vaccine candidate due to the immune response it generates (60, 62).

Shr binds Hb and acquires hemin from Hb (59, 63, 64). Shr is a multi-domain protein which contains two NEAT domains separated by a series of leucine-rich repeats of unknown function (60, 63). Unlike the *S. aureus* Isd system, Hb binding by Shr does not occur at the NEAT domains. Instead Hb binding is mediated by residues in the N-terminal region of the protein that contains two DUF1533 domains (63, 65). The N-terminal region is incapable of binding hemin; hemin binding occurs at the NEAT domains with each domain being capable of
binding one hemin molecule (63).

Shr directly transfers hemin to Shp, presumably facilitated by protein-protein interactions (64, 66). Furthermore, the NEAT1 domain in Shr is able to transfer hemin rapidly to Shp or Shr’s C-terminal NEAT2 domain, yet NEAT2 is only able to rapidly transfer hemin back to NEAT1, but not to Shp (67). Taken together, it has been suggested that NEAT2 functions to store hemin near the cell surface (67). Even though NEAT2 binds hemin, it lacks the conserved tyrosines in the canonical YXXXY motif (63). Therefore, it seems likely that NEAT2 engages hemin in a novel way.

A crystal structure of hemin in complex with Shp reveals a bis-methionine ligation in which the iron atom in the hemin molecule is coordinated on opposite sides by methionine residues (68). While similar to the NEAT domains, the structure of Shp is distinct, leading to its classification as a NEAT-like domain (68). Shp uni-directionally transfers hemin to HtsA (69). In vitro this occurs at a rate that is ~100,000 faster than the spontaneous rate of hemin release from Shp (70). Transfer is mediated by the formation of a pre-transfer holoShp:apoHtsA complex and is believed to occur through a “plug-in” mechanism whereby H229 and M79 hemin axial ligands of HtsA displace the M66 and M153 axial ligands of Shp (70–74). The ABC-transporter HtsABC then presumably transports the hemin molecule across the cell membrane into the bacterial cytoplasm (75). At this point, the hemin molecule is likely degraded by a yet to be identified heme oxygenase that presumably releases iron for subsequent use by the microbe.

1.2.3 Iron Acquisition by Bacillus anthracis

In the heme acquisition system of B. anthracis, the isd locus contains eight protein encoding genes: isdC, isdX1, isdX2, isdE1, isdE2, isdF, srtB, isdG. IsdX1 and IsdX2 are secreted
into the extracellular milieu under iron limiting-conditions and bind hemin and Hb. They have also been shown be capable of acquiring hemin from Hb (76, 77). In vitro, holo-IsdX1 unidirectionally transfers hemin to apoIsdC at a rate that is ~10,000 fold faster than the spontaneous rate of hemin release from IsdX1 (78). Once hemin is bound by IsdC, it is presumably transferred to the ABC-transporter (IsdE1E2F), which then transports it across the cell membrane. Once the hemin molecule has been transported inside the bacterium, IsdG enzymatically degrades it to yield free iron and biliverdin (79). The pathway is shown in Figure 1.5. I will now explain what is known about the various steps in this pathway in more detail.

IsdX1 and IsdX2 contain one and five NEAT domains, respectively. In IsdX2, each of the NEAT domains bind to Hb and each, except for NEAT2, also binds to hemin. The order of hemin affinity is as follows: NEAT5 > NEAT4 > NEAT3 ≈ NEAT1 (81). Interestingly, only NEAT1 and NEAT5 are able to scavenge hemin from Hb (81). In an effort to understand the hemin scavenging properties of NEAT1 and NEAT5, a sequence analysis of all five NEAT domains revealed a glutamine residue conserved in NEAT1 and NEAT5 at position 29 in the alignment (82). However, NEAT2, NEAT3, and NEAT4 had a threonine at position 29 instead (82). A Q29T mutation in NEAT5 did not disrupt hemin binding, but this mutant was no longer able to scavenge hemin from Hb (82). Interestingly, a T29Q mutation in NEAT3 enabled this domain to acquire hemin from Hb (82). This demonstrates the critical role of this glutamine residue in hemin acquisition by the NEAT domains in IsdX2. A crystal structure of the NEAT5 domain revealed that the Q29 is on the solvent exposed face of the 3_10-helix above the hemin binding pocket and is therefore poised to interact with Hb (82). Although the authors demonstrated heme transfer from Hb to the NEAT domains, it is not known if the receptors \textit{B. anthracis} accelerate hemin release from Hb in a similar manner as the \textit{S. aureus} IsdB and IsdH
receptors.

Holo-IsdX1 is also able to transfer hemin to apo-IsdX2 and apo-IsdC at approximately the same rate and it is hypothesized that hemin transfer is facilitated by protein-protein interactions (79). Furthermore, holo-IsdX1 binds apo-IsdC and apo-IsdX2 with $K_D$’s of 5 $\mu$M and 10 nM, respectively (79). It is interesting to note that holo-IsdX1 binds apo-IsdX2 with higher affinity than apo-IsdC. This suggests that IsdX1 preferentially transfers hemin to IsdX2. When expressed in isolation, NEAT domains 1, 3, and 4 of IsdX2 transfer hemin to IsdC, but NEAT5 does not (81). The origin of these differences is not known but may result from the high affinity of NEAT5 for hemin (the highest affinity of all NEAT domains in IsdX2). Based on similarities with the S. aureus system, IsdC presumably acts as the central conduit for relaying hemin acquired by surface exposed hemophores to the ABC-transporter.

Additional hemophores in B. anthracis have also been identified. The B. anthracis S-layer protein K (BslK) is predicted to be expressed under iron limiting conditions and localizes non-covalently to the cell envelope presumably via its three S-layer homology (SLH) domains (83). BslK contains a single NEAT-domain (BslKN), which is capable of binding hemin (83). BslKN directly transfers hemin to IsdC at a rate that is >10,000 fold faster than the spontaneous rate of hemin release from BslKN (83). BslKN is not able to efficiently transfer hemin to IsdX1 (83). It is unclear if the inability to transfer hemin results from BslKN’s weaker affinity for hemin. Furthermore, holo-BslKN interacts with apo-IsdC and it is presumably through these protein-protein interactions that hemin transfer is facilitated (83). The source of the hemin bound by BslK in vivo is currently unknown.

In a different study, the heme-acquisition leucine-rich repeat protein (Hal) was identified. Hal is presumably anchored to the cell wall via a sorting signal at its C-terminus (84). Hal
contains a single NEAT domain at its N-terminus followed by a series of leucine-rich repeats (LLR) of unknown function (84). When the NEAT domain of Hal was expressed in isolation (HalN), it was shown to bind hemin and Hb (84). Furthermore, HalN was able to acquire hemin from Hb (84). A Δhal mutant of B. anthracis exhibited a reduced ability to grow when hemin and Hb were the sole sources of iron (84).

1.2.4 Iron Acquisition by Corynebacterium diphtheriae

In Corynebacterium diphtheriae, there are several receptors displayed on the cell surface which bind hemin and/or Hb. HtaA and HtaB are two proteins displayed on the cell surface whose expression is regulated by iron (85). HtaA is able to bind hemin and Hb, while HtaB is only able to bind hemin (85, 86). HtaA is able to acquire hemin from Hb and transfer it to HtaB (86). Once heme is bound by HtaB, it is presumably transferred to HmuT, the lipoprotein component of the ABC-transporter HmuTUV (87). Once it has been transferred across the bacterial membrane into the cytoplasm, the heme oxygenase HmuO degrades the tetrapyrrole ring to yield free iron (88, 89). The pathway is shown in Figure 1.6. I will now explain what is known about the various steps in this pathway in more detail.

An analysis of the primary sequences revealed that HtaA and HtaB contained regions of roughly 150-200 residues with a high degree of sequence conservation which were subsequently named conserved regions (CR) (85). CR domains have no primary sequence homology to NEAT domains (86). HtaA and HtaB contain two and one CR domain(s), respectively (85). When expressed in isolation, both CR’s in HtaA (HtaA^{CR1} and HtaA^{CR2}) are able to bind hemin and Hb, with HtaA^{CR2} binding both with higher affinity (86). Kinetics studies of the hemin transfer process revealed that HtaA^{CR2} did not seem to accelerate hemin release from metHb and it has
been suggested that this transfer process is passive (90). Interestingly, thermal denaturation studies of HtaA\textsuperscript{CR2} indicate that the hemin-bound form is more stable than the apo-form and suggests that the hemin molecule plays a role in stabilizing the structure of the CR domain (90).

Additional surface exposed receptors involved in iron acquisition have also been identified. ChtA is exposed on the cell surface, its expression is regulated by iron, and it binds hemin and Hb (91). ChtA contains a single CR domain which alone binds hemin, Hb, and the Hb-Hp complex (91, 92). In a recent study, it was shown that Hb and Hb-Hp binds the surface of \textit{C. diphtheriae} (93). In an effort to further understand the interaction of Hb and Hb-Hp with \textit{C. diphtheriae} host proteins, binding studies with surface exposed proteins following fractionation revealed a novel protein (HbpA) (93). HbpA is expressed under low iron conditions, is localized to the cell surface, and does not contain any CR domains (93). HbpA binds Hb and Hb-Hp, but not hemin (93). \textit{ΔhbpA} mutant strains of \textit{C. diphtheriae} exhibited no growth defect when FeCl\textsubscript{3} or heme were the sole sources of iron but did exhibit reduced growth when Hb-Hp was the sole source of iron (93). This highlights the importance of HbpA in utilizing Hb-Hp and suggests that HbpA is the main receptor for this iron source.

1.2.5 Iron Acquisition Systems found in other Gram-positive Pathogens

\textit{Bacillus cereus} uses Hb, hemin, and ferritin as iron sources (94). A protein displayed on the surface was identified, its expression was found to be regulated by iron, and it was named the iron regulated leucine-rich surface protein (IlsA) (Daou 2009). IlsA contains one NEAT domain at its N-terminus followed by a series of leucine-rich repeats (LLR). After the LLRs, there are three S-layer homology domains at the C-terminus (94). While still able to grow in the presence of FeCl\textsubscript{3}, an \textit{ΔilsA} mutant of \textit{B. cereus} was unable to utilize hemin, Hb, or ferritin as iron sources.
(94). IlsA binds hemin, Hb, and ferritin in vitro (94). Apo-IlsA and apo-IlsA\textsuperscript{NEAT} are able to bind hemin and extract hemin form Hb (95). The IlsA\textsuperscript{LLR} domain does not bind hemin nor extract hemin from Hb (95). \textit{B. cereus} contains an \textit{isd} locus with a high degree of homology to the \textit{B. anthracis} \textit{isd} system (95). Within the \textit{B. cereus} \textit{isd} locus, there is a gene which shares 98% sequence homology to \textit{B. anthracis} \textit{isdC} likely indicating that this corresponds to the \textit{B. cereus} IsdC (IsdC\textsubscript{Bc}). Holo-IlsA and holo-IlsA\textsuperscript{NEAT} rapidly transfer hemin to IsdC\textsubscript{Bc} (95).

\textit{Streptococcus pneumoniae} growth can be supported when hemin or Hb are the sole iron sources (96, 97). A novel Hb binding protein of ~37 kDa was identified (Spbhp-37). Expression of Spbhp-37 on the cell surface was increased when Hb was the sole source of iron (98). Interestingly, \textit{S. pneumoniae} secretes a glyceraldehyde-3-phosphate dehydrogenase (GAPDH) which binds hemin and Hb (97). This GAPDH potentially binds hemin through a conserved motif implicated in hemin binding (97). However, the interaction of GAPDH with hemin and Hb has yet to be characterized structurally and the \textit{in vivo} relevance of this interaction in the iron acquisition process by \textit{S. pneumoniae} is not clearly understood (97).

\textit{Staphylococcus lugdunensis} also contains an \textit{isd} locus with the expressed proteins containing a high degree of sequence homology to the one identified in \textit{Staphylococcus aureus} (99). In this Isd system, IsdC and IsdB are displayed on the cell surface and are expressed under iron limiting conditions (99). IsdB, IsdC, IsdJ, and IsdK all bind hemin, but only IsdB binds Hb and Hb-Hp (99). IsdB and IsdJ each contain two NEAT domains while IsdC and IsdK each contain one (99). The conserved aromatic residues in the \(\beta 1\)-\(\beta 2\) loop which had previously been shown to be essential for Hb binding in \textit{S. aureus} IsdH and IsdB were also essential for Hb binding by \textit{S. lugdunensis} IsdB as well (99). As is the case with \textit{S. aureus}, \textit{S. lugdunensis} preferentially grows on human Hb rather than mouse Hb as the sole iron source (99).
Furthermore, a ΔisdB mutant of *S. lugdunensis* exhibited attenuated growth when Hb was the sole source of iron (99).

In the *Listeria monocytogenes* iron acquisition system, the *isd* locus contains the following genes: *hbp1*, *hbp2*, an ABC transporter, *srtB*, and 3 additional genes of unknown function (100). Hbp2 is surface exposed and binds hemin (100–103). Hbp1 and Hbp2 contain one and three NEAT domains (Hbp2N1, Hbp2N2, Hbp2N3), respectively. Each of these NEAT domains is capable of independently binding hemin with the following order of affinities: Hbp1 > Hbp2N3 > Hbp2N2 (103). Hbp1, Hbp2N1, and Hbp2N3 all bind Hb (103). Lacking the conserved YXXXY motif on the β8-strand essential for hemin binding in most NEAT domains, the crystal structure of Hbp2N2 in complex with hemin reveals unexpectedly that a tyrosine residue on the adjacent β7-strand coordinates the hemin molecule (103). A crystal structure of apo-Hbp2N2 reveals that Hbp2N2 undergoes an unusual conformational change during hemin binding not observed by other hemin-binding NEAT domains (103). Once hemin has been bound by Hbp1 and Hbp2, the pathway by which hemin is imported into the bacterial cytoplasm is not well understood. Once inside the cytoplasm, an IsdG-like heme monooxygenase (IsdLmHde) catalyzes the degradation of hemin to biliverdin (104).
1.3 Scope of the Dissertation

Research described in this thesis focuses on the mechanism used by Gram-positive pathogens to acquire heme-iron from human Hb. Hb is an oxygen transporter found in human erythrocytes and contains the majority of the body’s iron. Heme extraction from Hb is needed to acquire iron, an essential nutrient required for pathogenesis. Two microbial Hb receptors are the focus of my research: IsdH found in *Staphylococcus aureus* and Shr found in *Streptococcus pyogenes*.

In CHAPTER 2, we characterize an N-terminal region in the Shr receptor which had previously been shown to bind Hb but was otherwise uncharacterized. Through biophysical methods, we show that this N-terminal region contains three autonomously folded domains, two of which interact with Hb leading us to rename them Hemoglobin Interacting Domains (HIDs). We report on the binding stoichiometry to Hb, the backbone dynamics and crystal structure of HID2. HID2 is a novel Hb binding domain.

In CHAPTER 3, the crystal structure of HID2 in complex with Hb is reported. It reveals a hitherto novel interface between human Hb and a receptor from a bacterial pathogen. In this structure, the HID2 domain binds in front of the heme-binding pocket of Hb and appears to cap it. An alanine-scan is carried out and important residues for the interaction with Hb are revealed. Kinetics data suggest that HID2 slows the rate of heme release from Hb. The implications of this are discussed. This work lays the foundation for understanding how the full-length Shr protein extracts heme from Hb.

In CHAPTER 4, rigorous kinetics and thermodynamics studies of the heme extraction process from human Hb by IsdH reveal that IsdH engages Hb through two energetically distinct interfaces, accelerates heme release from the α-subunit of Hb by up to 13,400-fold and lowers
the activation energy by ~5.9 kcal/mol. Furthermore, MD simulations suggest that IsdH-induced structural distortions of Hb facilitate hydration of Hb’s heme binding pocket and increase the rate of hydrolytic cleavage of the axial HisF8 N$_e$-Fe$^{3+}$ bond. Taken together, a step-by-step reaction mechanism of heme extraction by IsdH is presented.

In CHAPTER 5, the NMR-derived solution structure and backbone dynamics of the C-terminal domain of PhoP (PhoPC) are reported. PhoP is part of the PhoPR two-component system which regulates gene expression and is required for virulence. Our data shows that the polypeptide that connects the DNA-binding domain to the regulatory domain in PhoP is unstructured. This work contributes to an understanding of the mechanism by which PhoP regulates gene expression.
1.4 Figures

![Diagram of the S. aureus Isd system](image)

**Figure 1.1** The *S. aureus* Isd system

On the cell surface, IsdH and IsdB engage with Hb or the Hp-Hb complex. These receptors extract heme from Hb and pass it to IsdA. IsdA then transfers heme to IsdC which is buried in the cell wall. IsdC passes the heme to the ABC-transporter IsdDEF which transports the heme across the cell membrane into the bacterial cytoplasm where the tetapyrrole ring is degraded by IsdG/I to yield free iron. (Figure courtesy of Thomas Spirig).
**Figure 1.2 IsdB and IsdH Share a Conserved Tri-Domain Unit**

There is a conserved tri-domain unit in IsdH and IsdB which shares a high degree of sequence homology and actively extracts heme from Hb, IsdH$^{N2N3}$ and IsdB$^{N1N2}$ respectively. In this conserved receptor, the first NEAT domain binds Hb while the second NEAT binds heme. Both NEAT domains are separated by an α-helical linker. (Figure from Spirig et al. 2013).
Figure 1.3  Crystal Structure of IsdH$^{N2N3}$-F365V/A369F/Y642A in Complex with metHb

A mutated version of the IsdH$^{N2N3}$ receptor bound to the α-subunit of Hb. The N2 domain contacts the A-helix of Hb while the N3 domain binds above the heme binding pocket of Hb. The linker and N3 domains engage and perturb the F-helix of Hb. (Figure from Dickson et al. 2015).
The Shr receptor displayed on the cell surface engages Hb with the N-terminal domain. Heme is then extracted and bound by the NEAT1 domain and transferred to either the NEAT2 domain or Shp. Once bound by Shp, heme is transferred to the ABC-transporter SiaABC (a.k.a. HtsABC). (Figure from Ouattara et al. 2013).
**Figure 1.5 The Isd System in *B. anthracis***

IsdX1 and IsdX2 are secreted on the cell surface. They bind heme and extract heme from Hb. IsdX1 then transfers heme to IsdC which passes it to the ABC-transporter IsdDEF. After heme is imported into the cytoplasm, the tetrapyrrole ring is degraded by IsdG to yield free iron (Figure from Honsa et al. 2011).
Figure 1.6  The Iron Acquisition System in *C. diphtheriae*

HtaA and ChtA bind Hb and acquire heme once it’s released. HtaA then passes heme to HtaB which transfers heme to the ABC-transporter HmuTUV. Once imported into the bacterial cytoplasm, HmuO degrades the tetrapyrrole ring to yield free iron (Figure from Allen et al. 2015).
1.5 References


Binding of Hemoproteins and Acquisition of Iron. *Infect. Immun.* **71**, 1042–1055


Streptococcus pneumoniae requires iron for its viability and expresses two membrane proteins that bind haemoglobin and haem. *Metallomics*. 5, 384–389


Chapter 2

The *Streptococcus pyogenes* Shr Protein Captures Human Hemoglobin Using Two Structurally Unique Binding Domains
The work described in this chapter has been submitted to the Journal of Biological Chemistry and is currently under revision:


Copyright 2018

Journal of Biological Chemistry

I was the first author of this work and conducted the experiments. Duilio Cascio and Martin Phillips assisted with X-ray crystallography and AUC, respectively. Professor Clubb and I wrote the manuscript.
2.1 Overview

In order to proliferate and mount an infection, many bacterial pathogens need to acquire iron from their host. The most abundant iron source in the body is the oxygen transporter hemoglobin (Hb). *Streptococcus pyogenes*, a potentially lethal human pathogen, uses the Shr protein to capture Hb on the cell surface. Shr is an important virulence factor, yet the mechanism by which it captures Hb and acquires its heme is not well understood. Here we show using NMR and biochemical methods that Shr binds Hb using two related modules that were previously defined as domains of unknown function (DUF1533). These Hemoglobin Interacting Domains (HIDs), called HID1 and HID2, are autonomously folded and independently bind Hb. The 1.5 Å crystal structure of HID2 reveals that it is a structurally unique Hb binding domain. Mutagenesis studies reveal a conserved tyrosine in both HIDs that is essential for Hb binding. Our biochemical studies indicate that HID2 binds Hb with higher affinity than HID1, and that the Hb tetramer is engaged by two Shr receptors. NMR studies reveal the presence of a third autonomously folded domain between HID2 and a heme binding NEAT1 domain, suggesting that this linker domain may position NEAT1 near Hb for heme capture.
2.2 Introduction

*Streptococcus pyogenes* (Group A *streptococcus*) each year causes an estimated 700 million infections worldwide (1). While most of these infections lead to non-life threatening acute pharyngitis, many are lethal and cause ~500,000 deaths annually (2). These more severe and invasive infections have a high mortality rate (~30%) and include: necrotizing fasciitis and streptococcal toxic shock syndrome (3, 4). Infections caused by this deadly microbe are presently among the top ten causes of death from an infectious disease world-wide (5). *S. pyogenes* and other pathogenic bacteria require iron to proliferate, as it is a versatile metal that functions as a key biocatalyst and electron carrier in enzymes that mediate metabolism (6). During infections iron is frequently foraged from human hemoglobin (Hb), which contains ~75-80% of the body’s total iron content in the form of heme (iron-protoporphyrin IX). Recent studies have shown that *S. pyogenes* uses an array of surface and membrane associated proteins to acquire Hb’s heme. Understanding how these proteins function at a molecular level is of fundamental importance and could facilitate the discovery of new anti-infective agents that work by limiting microbial access to iron.

Surface and membrane associated proteins are used by *S. pyogenes* to capture heme. The Streptococcal hemoprotein receptor (Shr) is displayed on the bacterium’s surface, where it binds to Hb and acquires the oxidized form of heme (called hemin) (7–11). The Shr Hb receptor is an important virulence factor, as Δ*shr* mutant strains of *S. pyogenes* exhibit reduced growth in human blood and attenuated virulence in murine and zebrafish models of infection (10, 12). *In vitro*, Shr transfers hemin directly to Shp, a cell wall associated hemoprotein that uses a bis-methionine ligation mechanism to interact with hemin (13–15). Shp then passes hemin to HtsA/SiaA, the lipoprotein component of the ABC-transporter HtsABC/SiaABC (16). A binary Shp-HtsA complex mediates this transfer step, in which the methionine axial ligands in Shp are
displaced by methionine and histidine axial ligands within HtsA (17–21). This process occurs very rapidly, >100,000-times faster than the rate at which Shp spontaneously releases hemin into the solvent. Hemin is then presumably transported across the membrane into the bacterial cytoplasm where it is degraded by a yet to be identified oxygenase(s) that releases hemin’s iron for use by the microbe (6). While hemin transfer from Shp to HtsA/SiaA has been extensively studied by Lei and colleagues, the first step in the pathway - the extraction of hemin from human Hb by Shr - is less well understood and the focus of this study.

Many Gram-positive bacterial pathogens display Hb receptors that contain NEAr-iron Transport (NEAT) domains (22, 23). NEAT domains have been shown to bind Hb, hemin, and in many instances are able to rapidly transfer hemin amongst one another (24–35). Hemin acquisition by Staphylococcus aureus has been studied extensively. This microbe uses an array of iron-regulated surface determinant (Isd) proteins that each contain one or more NEAT domains (24, 29, 36–40). Hb is captured on the cell surface by the closely related IsdB and IsdH proteins (41). Our studies have shown that IsdH captures and extracts hemin from Hb using a conserved tri-domain unit that contains two NEAT domains, called N2 and N3, that are separated by a helical linker domain (42). As compared to the rate of hemin released into the solvent by Hb, binding of the tri-domain unit accelerates hemin release 13,400-fold (43). Atomic structures of IsdH-Hb complexes reveal that its N2 and N3 NEAT domains adopt related β-sandwich structures that are decorated by short helices, while the intervening linker domain adopts a three-helix bundle (44, 45). The N2 domain engages Hb’s A-helix with high affinity, enabling the helical and N3 domains to distort Hb’s F-helix, thereby promoting hemin transfer to IsdH’s N3 domain (43, 45, 46). IsdB shares significant primary sequence homology with IsdH and also contains a tri-domain unit that captures hemin from Hb through a similar mechanism (31, 34, 42,
47–50). Other Gram-positive pathogens use NEAT domains to bind hemin and/or Hb, and include among others *Listeria monocytogenes*, *Bacillus anthracis*, *Bacillus cereus*, and *Streptococcus pyogenes* (23). However, the mechanisms used by these domains to capture hemin from Hb may be distinct from that in IsdH and IsdB, since in some instances a single NEAT domain has been proposed to bind to both Hb and heme (51–54). Notable examples are IsdX1, IsdX2, and Hal in which single NEAT domains from these proteins have been reported to bind Hb and extract hemin from Hb (52–54). However, complexes of these NEAT domains bound to Hb have not been visualized at atomic resolution.

Unlike other Hb receptors in Gram-positive bacteria, the *S. pyogenes* Shr protein does not use NEAT domains to engage Hb. Instead, it binds Hb via an N-terminal region (NTR) that is predicted to contain two domains of unknown function, called DUF1533 domains (8, 11) (Fig. 2.1). Released hemin is then bound by residues located within Shr’s C-terminal region (CTR) (residues 365-1275), which contains two hemin binding NEAT domains (NEAT1 and NEAT2) that are separated from one another by a series of Leucine-rich repeats (LRRs) (8). Interestingly, similar to the staphylococcal IsdB and IsdH proteins, transfer experiments have shown that the isolated Shr protein increases the rate of hemin release from Hb *in vitro*, suggesting that the domains within the receptor function synergistically to extract hemin (9). The NEAT domains within Shr can rapidly transfer hemin amongst one another, but only NEAT1 is able to efficiently transfer hemin to Shp (55). These findings led to the suggestion that NEAT2 stores hemin near the cell surface (55).

Here we show that Shr uses two related DUF1533 domains to engage Hb, which we rename Hb Interacting Domains (HIDs). Biochemical studies and the atomic structure of a HID reveal that it is a structurally novel Hb binding module that employs a conserved tyrosine residue
for Hb recognition. Biochemical studies indicate that the HIDs within Shr bind Hb with differing affinities and suggest that each engages a single globin subunit. Interestingly, NMR studies reveal the presence of a third autonomously folded domain that is located between the HID2 and NEAT1 domains. The implications of this arrangement in hemin acquisition from Hb are discussed.

2.3 Results

2.3.1 The Hb interacting N-terminal region within Shr contains three autonomously folded domains

Previous studies have shown the S. pyogenes Shr protein interacts with Hb via residues located at its N-terminus (8). This N-terminal region is formed by residues Lys25-Val364 and is hereafter referred to as the NTR (Fig. 2.1). An inspection of its primary sequence suggests that the NTR contains three folded regions. These include, two domains of unknown function (DUF1533 domains) that are formed by residues Gly61-Phe123 and Ile203-Val269 (8), as well as residues Val299-Ser358 that are predicted to form several \(\alpha\)-helices. Constructs containing residues Gly61-Phe123 and Ile203-Val269 were insoluble. Secondary structural prediction algorithms were used to guide the design of additional constructs until highly soluble ones were obtained and characterized by NMR: Shr\(^{\text{HID1}}\), residues Ser26-Lys148 that correspond to the first DUF1533 domain; Shr\(^{\text{HID2}}\), residues Lys155-Gln285 that correspond to the second DUF1533 domain; and Shr\(^{\text{L}}\), residues Ser294-Val364 that contain the predicted alpha helical region in Shr that links Shr\(^{\text{HID2}}\) to the first NEAT domain (Fig. 2.1). Each polypeptide was uniformly \(^{15}\text{N}\)-labeled, purified, and its \(^{1}\text{H}-^{15}\text{N}\) HSQC spectrum recorded (Fig. 2.2a,c,e). All three polypeptides adopt a folded structure, as the cross peaks in each spectrum are well-dispersed and have appropriate linewidths. Based on the Pfam prediction for their domain boundaries the two
DUF1533 domains within the NTR are separated by ~80 amino acids. To determine if the domains interact with one another a longer polypeptide fragment containing both domains was characterized by NMR (Shr\textsuperscript{HID1-2}, consisting of residues Ser26\textendash Gln285). Superposition of the NMR spectra of \([^{15}\text{N}]-\text{Shr}^{\text{HID1}}\) and \([^{15}\text{N}]-\text{Shr}^{\text{HID2}}\) onto the spectrum of \([^{15}\text{N}]-\text{Shr}^{\text{HID1-2}}\) reveals that the cross-peaks have similar chemical shifts (Fig. S2.1). Thus, in the context of the longer Shr\textsuperscript{HID1-2} polypeptide, the Shr\textsuperscript{HID1} and Shr\textsuperscript{HID2} domains do not significantly interact with one another and they exhibit the same structural fold as they do in isolation. The Shr\textsuperscript{HID2} and Shr\textsuperscript{L} domains also do not interact with one another, as the NMR spectra of a polypeptide containing both of these domains (\([^{15}\text{N}]-\text{Shr}^{\text{HID2-L}},\) residues Gly61\textendash Asn368) can be recapitulated by summing the spectra of the isolated \([^{15}\text{N}]-\text{Shr}^{\text{HID2}}\) and \([^{15}\text{N}]-\text{Shr}^{\text{L}}\) domains (Fig. S2.1). We conclude from the NMR data that the Hb interacting region within Shr contains three independently folded and non-interacting structural domains, two DUF1533 domains and a linker domain that is predicted to contain several \(\alpha\)-helices.

### 2.3.2 DUF1533 domains bind Hb

We used NMR to determine if the domains within the NTR interacted with Hb. The \(^1\text{H}-^{15}\text{N}\) HSQC spectrum of each \(^{15}\text{N}\)-labeled domain was recorded in the presence or absence of Hb (1:1 stoichiometry based on tetramer units) (Fig. 2.2). Upon adding Hb to a solution containing \([^{15}\text{N}]-\text{Shr}^{\text{HID1}}\), a massive loss of signal occurs indicative of complex formation. A similar result is obtained when Hb is added to \([^{15}\text{N}]-\text{Shr}^{\text{HID2}}\) (compare Figs. 2.2a and c with Figs. 2.2b and d) indicating that both DUF1533 domains interact with Hb. In contrast, adding Hb to the \([^{15}\text{N}]-\text{Shr}^{\text{L}}\) domain causes only minimal changes in the NMR spectrum indicating that the proteins do not interact (Fig. 2.2f).

Having established that DUF1533 domains bind to Hb, we next measured their affinities
using ITC (Fig. 2.3a-d and Table 2.1). This data indicates that Shr$^{\text{HID2}}$ binds Hb with a $K_D = 16.0 \pm 3.4 \mu\text{M}$ in a process that is enthalpically favorable (Fig. 2.3b and Table 2.1). Shr$^{\text{HID1}}$ also binds Hb, but the interaction is weaker ($K_D = \sim 140 \mu\text{M}$). The Shr$^{\text{HID1-2}}$ polypeptide containing both DUF1533 domains binds Hb with a $K_D = 5.1 \pm 0.4 \mu\text{M}$ (Fig. 2.3a and Table 2.1). This is only slightly higher in affinity than Shr$^{\text{HID2}}$, suggesting that the domains within this polypeptide independently interact with Hb. From this data, we conclude that both of Shr’s DUF1533 domains bind Hb and the second domain (Shr$^{\text{HID2}}$) is the most important for binding. Based on these results, we propose that the DUF1533 domains within Shr be renamed Hemoglobin Interacting Domains (HIDs).

A NMR experiment was performed to ascertain if the HIDs (DUF1533 domains) bind to the same site on Hb. In this experiment, the ability of unlabeled Shr$^{\text{HID2}}$ to displace Hb-bound $^{15}\text{N}$-Shr$^{\text{HID1}}$ was determined. Initially, unlabeled Hb was titrated into a solution of $^{15}\text{N}$-Shr$^{\text{HID1}}$ such that the signals in its $^1\text{H}$-$^{15}\text{N}$ HSQC spectrum were broadened beyond detection as a result of complex formation (Fig. 2.4b). Unlabeled Shr$^{\text{HID2}}$ was then titrated into the solution and the $^1\text{H}$-$^{15}\text{N}$ HSQC spectrum recorded to determine if it displaces $^{15}\text{N}$-Shr$^{\text{HID1}}$ from Hb. Adding unlabeled Shr$^{\text{HID2}}$ to the solution causes the signals from $^{15}\text{N}$-Shr$^{\text{HID1}}$ to reappear in the HSQC spectrum, meaning that Shr$^{\text{HID2}}$ outcompetes $^{15}\text{N}$-Shr$^{\text{HID1}}$ for its binding site(s) located on Hb. The data show a progressive reappearance of the signals from $^{15}\text{N}$-Shr$^{\text{HID1}}$, starting when a two-fold molar equivalent of Shr$^{\text{HID2}}$ is added (Fig. 2.4c). When four-fold Shr$^{\text{HID2}}$ is added, the spectrum of $^{15}\text{N}$-Shr$^{\text{HID1}}$ is almost completely recovered yet the signal intensities are approximately one third as intense as in the original HSQC (Fig. 2.4d). Importantly, the reappearance of the NMR signals is not caused by simply diluting the NMR sample, as the signals from $^{15}\text{N}$-Shr$^{\text{HID1}}$ remain broadened in a control experiment in which the $^{15}\text{N}$-Shr$^{\text{HID1}}$-
Hb complex was diluted only with buffer. As unlabeled Shr\textsuperscript{HID2} displaces \[^{15}\text{N}]-\text{Shr}\textsuperscript{HID1} from Hb, we conclude that the isolated domains can compete for the same site(s) on Hb. However, the fact that this resulted in only a partial recovery of signal could indicate that Shr\textsuperscript{HID1} binds a second site on Hb and binds to Hb more promiscuously than Shr\textsuperscript{HID2}.

2.3.3 Two Shr\textsuperscript{HID1-2} proteins bind to tetrameric Hb

The binding stoichiometry of the Shr\textsuperscript{HID1-2}:Hb complex was determined using size exclusion chromatography with inline multi-angle light scattering (SEC-MALS). When Shr\textsuperscript{HID1-2} is mixed with the Hb tetramer at a ratio of 2:1 (Shr\textsuperscript{HID1-2} to Hb tetramer units), a complex at 101 kDa is observed (Fig. 2.3e). This is consistent with \(~1.8\) Shr\textsuperscript{HID1-2} proteins binding to one Hb tetramer (calculated as follows: \([101 \text{ kDa}_{\text{complex}} - 56 \text{ kDa}_{\text{Hb alone}}]/25 \text{ kDa}_{\text{Shr}\textsuperscript{HID1-2} \text{ alone}} = 1.8\)). SEC-MALS experiments using higher Shr\textsuperscript{HID1-2}:Hb ratios (4:1 and 6:1) were also performed and yielded similar results, indicating that Hb is saturated and a maximum of two Shr\textsuperscript{HID1-2} proteins bind to one Hb tetramer. Notably, the stoichiometry predicted by SEC-MALS is similar to that obtained from curve-fitting of the ITC data (Table 2.1); two Shr\textsuperscript{HID1-2} proteins bind to one Hb tetramer.

Analytical ultracentrifugation (AUC) was used to determine the stoichiometry of the Shr\textsuperscript{HID2}:Hb complex. This method was used instead of SEC-MALS because the molecular weights of Hb and its complex with Shr\textsuperscript{HID2} are too similar, making it difficult to resolve these species by column chromatography. Similar to Shr\textsuperscript{HID1-2}, Shr\textsuperscript{HID2} is monomeric in solution (Fig. 2.3f); it has a weighted average molecular mass of 13.3 kDa (Table S2.1), which corresponds closely to its theoretical molecular weight of 14.4 kDa (56). In order to simplify the analysis of the AUC data, binding experiments employed Hb0.1, a stabilized tetrameric from of Hb in which
the alpha globin chains are part of a single polypeptide. When Shr$^{\text{HID2}}$ is added to Hb0.1 at a
180-fold molar excess (Hb tetramer units), its weighted average molecular mass is 97.4 kDa
(Fig. 2.3g and Table S2.1), consistent with two Shr$^{\text{HID2}}$ domains binding to one Hb0.1 tetramer.
Hb0.1 is saturated with Shr$^{\text{HID2}}$, since similar molecular weights are measured when Shr$^{\text{HID2}}$ is
present at 60-fold molar excess. Interestingly, the binding stoichiometry obtained from fitting the
ITC data suggests that four Shr$^{\text{HID2}}$ proteins can bind to a tetramer, which differs from the 2:1
stoichiometry determined by AUC (Table 2.1). However, stoichiometry estimates obtained by
ITC may be unreliable because the relatively weak affinity of complex formation causes the
binding isotherm to exhibit non-sigmoidal character.

2.3.4 The HID is a structurally novel Hb-binding domain

To gain insight into the molecular basis of Hb binding we determined the crystal structure
of HID2, the second DUF1533 domain within Shr. To facilitate its crystallization, we first used
NMR to delineate regions within the Shr$^{\text{HID2}}$ polypeptide that are structurally ordered. Triple
resonance methods were applied to [U-13C; U-15N]-Shr$^{\text{HID2}}$, enabling ~94% of its backbone
resonances to be assigned (Fig. 2.5a). A heteronuclear {1H}-15N NOE experiment was then
performed to identify regions within Shr$^{\text{HID2}}$ that are mobile and unstructured (Fig. 2.5b). The
majority of residues within Shr$^{\text{HID2}}$ are structurally-ordered in the absence of Hb, as indicated by
heteronuclear NOE values > 0.6 for backbone amide groups that span residues Ile179-Gln283.
However, ~20 residues located at the N-terminus of Shr$^{\text{HID2}}$ are disordered, as they possess low
magnitude heteronuclear NOEs. We therefore produced a truncated variant of Shr$^{\text{HID2}}$ for x-ray
crystallography studies that removes its unstructured N-terminal residues (Shr$^{\text{AHID2}}$, residues
Asn175-Gln285). Importantly, Shr$^{\text{AHID2}}$ binds to Hb with similar affinity and stoichiometry as
Shr$^{\text{HID2}}$ (Fig. 2.3c and Table 2.1), indicating that the deleted N-terminal residues are not important for function.

The structure of Shr$^{\text{HID2}}$ was determined at 1.5 Å resolution using single-wavelength anomalous diffraction methods applied to native and KI soaked crystals. Two Shr$^{\text{HID2}}$ domains are present in the asymmetric unit (Fig. 2.6a), forming a two-fold symmetry related dimer that buries 517 Å$^2$ of surface area (57). As SEC-MALS and AUC analyses indicate that Shr$^{\text{HID2}}$ is monomeric, the dimer observed in the structure is likely an artifact caused by the high concentration of protein present in the crystal. The two polypeptide chains in the dimer possess nearly identical atomic structures, as their heavy atom coordinates can be superimposed with a r.m.s.d. of 0.80 Å. The polypeptides are well defined by the electron density, with the exception of residue Asp226 in chain A and residues Lys284 and Gln285 in chain B. Complete structural and data statistics are presented in Table 2.2.

Shr$^{\text{HID2}}$ adopts a β-sandwich type fold in which two antiparallel β-sheets are flanked by an α-helix. One face of the sandwich is formed by a sheet containing strands β1, β2, β6, and β5 (β-sheet 1), while the other face is formed by a sheet containing strands β9, β8, β7, β3, and β4 (β-sheet 2) (Fig. 2.6b). Strands β1 and β2 form an antiparallel β-hairpin. The β2 strand is then followed by a long loop that is connected to a two-turn α-helix that is positioned nearly perpendicular with respect to β-sheet 2. Following the helix, strand β3 makes antiparallel contacts with strand β7 in β-sheet 2. Strand β4 then forms part of a β-hairpin with β3 before being followed by an extended loop that winds around, and extends over the top of β-sheet 1. This is followed by strand β5 that forms a β-hairpin with strand β6, and together with strands β1 and β2 comprise β-sheet 1. The structure is completed by a β-meander that is constructed from strands β7, β8, and β9 within β-sheet 2. Shr$^{\text{HID1}}$ presumably adopts a similar structure, as it
shares 27% sequence identity with Shr$^{\text{HID2}}$. The solvent exposed side of β-sheet 1 is predominantly acidic while on β-sheet 2 there is a greater balance of acidic and basic residues (Fig. 2.6c). It is worth mentioning that the HID domain is larger than the predicted DUF1533 domain wherein β1, β2, and the loop connecting it to α1 as well as β9 are all located outside of the predicted DUF1533 domain in the primary sequence.

The HID is a structurally unique bacterial Hb-binding domain. Many species of Gram-positive bacteria that are related to $S$. pyogenes employ NEAT domains to bind Hb. Based on a Dali analysis the Shr$^{\text{AHID2}}$ and NEAT domains adopt distinct structures; their Z-score does not exceed the threshold for a strong match as defined by Holm et al. (58). Moreover, Shr$^{\text{AHID2}}$ and NEAT domains have distinct secondary structure topologies, and Shr$^{\text{AHID2}}$ lacks the conserved aromatic motif that is present in Hb binding NEAT domains (Fig. 2.7a-b). Shr$^{\text{AHID2}}$ is also structurally distinct from Hb receptors found in other human pathogens, such as HpuA (59), ShuA (60) and HpHbR (61). The closest HID structural homolog in the Protein Data Bank is the CBM46 domain present in the Cel5B cellulase from Bacillus halodurans (PDB code 4uz8; the proteins can be superimposed with an r.m.s.d. of 2.1 Å and have a Dali Z-score of 7.9, which is deemed significant (58). However, the proteins have distinct functions, as key residues in CBM46 that mediate carbohydrate binding are poorly conserved in Shr$^{\text{AHID2}}$.

### 2.3.5 The HIDs within Shr$^{\text{HID1}}$ and Shr$^{\text{HID2}}$ engage Hb using a conserved tyrosine residue

We attempted to identify surface residues in the HID that mediate Hb binding by aligning the primary sequences of the DUF1533 domains and their flanking regions from putative Hb receptors. The primary sequences of DUF1533 domains from four Shr orthologues present in $C$. Novyi, $S$. iniae, $S$. equi, and $S$. dysgalactiae were aligned to Shr$^{\text{HID1}}$ and Shr$^{\text{HID2}}$ (Fig. S2.2). Six
residues are universally conserved, of which, only one has a side chain that is completely surface exposed, Y197 (Shr\textsuperscript{HID2} labeling). A Y197A mutation was introduced into Shr\textsuperscript{HID2} (Shr\textsuperscript{HID2-Y197A}) and its ability to bind Hb determined using the aforementioned NMR titration experiment. In marked contrast to wild type Shr\textsuperscript{HID2}, adding Hb to \textsuperscript{15}N-Shr\textsuperscript{HID2-Y197A} does not cause its NMR signals to broaden significantly, even when Hb is in excess (2:1 Hb:Shr\textsuperscript{HID2-Y197A}) (Fig. 2.8a). As the Y197A mutation does not disrupt the structure of Shr\textsuperscript{HID2-Y197A}, the side chain of Y197 presumably forms contacts to Hb that are required for binding. The corresponding tyrosine mutation was also introduced into Shr\textsuperscript{HID1} (Y55A, Shr\textsuperscript{HID1-Y55A}), and was also shown to disrupt Hb binding (Fig. 2.8b). We conclude that a conserved tyrosine residue in Shr\textsuperscript{HID1} and Shr\textsuperscript{HID2} is critical for Hb binding and that the domains likely engage Hb in a similar manner.

### 2.4 Discussion

Human hemoglobin is a rich source of iron that bacterial pathogens capture using surface displayed and secreted Hb receptors (62). In Gram-positive bacteria, these receptors typically interact with Hb via NEAT domains (23). Interestingly, the Hb receptor in \textit{S. pyogenes}, called Shr, appears to bind Hb via a novel mechanism. Its primary sequence has been divided into two regions, an N-terminal region (NTR) that binds Hb and a C-terminal region (CTR) that binds hemin (8, 11). Intriguingly, based on its primary sequence the NTR does not contain a NEAT domain suggesting that it engages Hb in a unique manner. To learn how Shr interacts with Hb we first used NMR spectroscopy to delineate residues within its NTR that form autonomously folded domains. Three regions within the NTR are structured, two modules previously defined as domains of unknown function (DUF1533 domains), and a third module located near the C-terminus of the NTR that is predicted to contain several \(\alpha\)-helices (Fig. 2.1). NMR titration experiments conclusively demonstrate that only the DUF1533 domains within the NTR interact
with Hb (Fig. 2.2). They therefore represent previously uncharacterized Hb Interacting Domains (HIDs). In Shr they are HID1 (residues 26-148) and HID2 (residues 155-285). Interactions originating from HID2 contribute significantly to Shr’s Hb binding affinity, since based on ITC measurements it binds to Hb with a $K_D = 16.0 \pm 3.4 \mu M$, whereas the isolated HID1 domain has significantly weaker affinity ($K_D = \sim 140 \mu M$). Studies of a polypeptide containing both domains (Shr$^{HID1-2}$), suggest that the HIDs interact with Hb independently of one another, as Shr$^{HID1-2}$ binds to Hb with similar affinity as the isolated HID2 domain (Shr$^{HID1-2}$ $K_D = 5.1 \pm 0.4 \mu M$). Moreover, NMR studies reveal that there are no significant domain-domain interactions within apo-Shr$^{HID1-2}$. Tsumoto and colleagues have measured the binding affinity of residues Gln22-Val364 of Shr, which contains both HIDs and the linker domain within the NTR (11). This polypeptide binds to Hb with similar affinity as Shr$^{HID1-2}$ ($K_D = 6.7 \mu M$), further demonstrating that only the HIDs within the NTR mediate Hb binding.

The crystal structure of HID2 reveals that it adopts a compact $\beta$-sandwich type fold that is distinct from previously characterized Hb binding proteins (Fig. 2.6b). Structures of three different NEAT domains in complex with Hb have been determined, including the first and second NEAT domains from the S. aureus IsdH receptor (IsdH$^{N1}$ and IsdH$^{N2}$), and the first NEAT domain from the S. aureus IsdB protein (IsdB$^{N1}$) (44, 50, 63). These NEAT domains share significant primary sequence homology with one another and engage Hb in a similar manner. The globin chain in Hb is contacted by residues in the NEAT domains that are located in an $\alpha$-helix that is positioned between strands $\beta1b$ and $\beta2$, and by residues located within nearby loops (loops connecting strands $\beta3-\beta4$ and $\beta5b-\beta6$). Each NEAT domain contains a conserved aromatic signature sequence ([F/Y]YH[F/Y]) that forms the Hb contacting $\alpha$-helix that when mutated disrupts binding. The HIDs within Shr must bind to Hb in a fundamentally distinct
manner, as their tertiary structures differ substantially from NEAT domains. Moreover, HID primary sequences do not contain the Hb contacting aromatic motif. Atomic structures of other Hb receptors from eukaryotic and Gram-negative microbes have also been determined and do not share significant structural homology with the HID.

To identify the surface on HIDs that contact Hb we compared their primary sequences and performed targeted mutagenesis. Six residues are completely conserved, but only Tyr197 (HID2 numbering) is fully solvent-exposed in the structure of HID2 and conserved. Tyr197 is located in the extended loop that precedes the alpha helix and is surrounded by other less well conserved residues that may mediate Hb interactions. Mutational exchange of the tyrosine to alanine in either HID1 or HID2 abrogates Hb binding, substantiating its functional importance (Fig. 2.8). The HID was originally named DUF1533. At present, more than 150 proteins containing DUF1533 domains from 71 distinct species of bacteria are known. Notably, these conserved tyrosine residues lie outside the predicted DUF1533 domains in the N-terminal regions of the HIDs. It is therefore unclear if all of these sequence-related domains will also bind Hb, as some lack the key tyrosine residue and they are present in proteins that are unlikely to be orthologues of Shr because they contain other domains that are predicted to have enzymatic functions.

In vitro binding studies indicate that a single Hb tetramer is engaged by two Shr receptors. SEC-MALS experiments demonstrate that a Shr\textsuperscript{HID1-2} polypeptide containing both HID1 and HID2 forms a complex with the Hb tetramer that has 2:1 stoichiometry (Fig. 2.3e). It is likely that the intact Shr protein binds with similar stoichiometry, as residues outside the NTR do not mediate interactions with Hb (8). Within the 2:1 Shr\textsuperscript{HID1-2}-Hb complex, we speculate that each HID2 domain binds to a specific site within the tetramer, as AUC studies indicate that two
isolated HID2 modules bind to one Hb tetramer (Fig. 2.3g). The role of HID1 in Hb binding remains unclear, as its weak affinity for Hb precluded accurate stoichiometry measurements. Moreover, NMR studies of the isolated HID1 and HID2 domains indicate that they compete for similar sites on Hb (Fig. 2.4). As ITC experiments indicate Shr$^{\text{HID2}}$ and Shr$^{\text{HID1-2}}$ bind Hb with generally similar affinities, it is possible that within the Hb:Shr$^{\text{HID1-2}}$ complex only HID2 engages Hb, after which the more weakly bound HID1 domain is displaced from Hb. Alternatively, in a model that we favor, both the HID1 and HID2 domains likely engage distinct sites on the Hb tetramer within the Hb:Shr$^{\text{HID1-2}}$ complex. In this scenario, each of the domains would bind a globin subunit, with interactions originating from HID2 being the most important. This model would be consistent with the results of the NMR experiments that showed that the isolated domains competitively bind Hb if it is assumed that the high protein concentrations used in the NMR studies caused the domains to promiscuously bind to the globin subunits and/or the domains bind differently to Hb in isolation as compared to Shr$^{\text{HID1-2}}$. It is interesting to note that HID binding to Hb is entropically unfavorable for both the isolated domains and Shr$^{\text{HID1-2}}$ (Table 2.1). It is possible that residues within the domain(s) undergo a disorder to order transition upon binding Hb, as is the case for the Hb-binding NEAT domains within the IsdH and IsdB S. aureus Hb receptors (50, 63). However, the origin of this effect remains unclear, as the heteronuclear NOE data of apo-HID2 do not reveal the presence of large disordered loops. Structures of the Shr:Hb complex are required to understand the molecular origins of these energetic effects.

Based on the domain architecture that we have elucidated, Shr may use a tri-domain unit to extract hemin from Hb. Only the heme extraction mechanisms used by the S. aureus IsdB and IsdH proteins are well understood. These proteins share similar primary sequences and extract hemin from Hb using a tri-domain unit, in which a helical domain separates N- and C-terminal
NEAT domains that bind to Hb and hemin, respectively (42). Detailed studies of IsdH indicate that its N-terminal NEAT domain (N2 domain) engages the A-helix within a globin chain contributing ~95% of the total binding standard free energy (43). This positions the remaining helical and NEAT domains near the hemin pocket on the same globin subunit, thereby enabling them to distort Hb’s F-helix to trigger hemin release. IsdB also contains a homologous tri-domain unit that binds to Hb in a similar manner (50). Interestingly, Shr contains a hemin binding NEAT domain (NEAT1) immediately following the HID2 and linker domains we have identified (Fig. 2.1). Thus, residues 175-502 in Shr form a D2-linker-NEAT1 tri-domain segment that is similar to those present in IsdB and IsdH. It is possible that Shr uses a similar mechanism as IsdH/B to accelerate hemin release from Hb in which HID2 binds to a globin chain to ‘deliver’ the NEAT1 domain near Hb’s hemin pocket. Interestingly, both the Shr and IsdH receptors supplement their tri-domain segments with an additional N-terminal Hb binding domain (HID1 in Shr and the N1 NEAT domain in IsdH). Recent studies indicate that the supplemental N1 domain in IsdH may function to slow the rate of Hb removal from the blood by obstructing interactions with the macrophage-specific endocytic receptor CD163, thereby advantageously prolonging microbial access to Hb’s heme iron (64). Additional studies will be needed to determine if the HID1 domain in Shr performs a similar function.

Collectively, the data presented in this paper provides new insight into how Shr captures Hb on the cell surface. *S. pyogenes* presumably encounters the dimeric form of Hb, as rupture of the red blood cells is expected to dramatically dilute Hb causing the tetramer to dissociate in αβ dimers. As we have shown that two receptors bind a Hb tetramer, on the cell surface Shr likely engages the Hb αβ dimer with 1:1 stoichiometry, such that its HID1 and HID2 domains each engage a globin subunit. Based on previous hemin transfer studies between Shr and Hb (8, 9), it
seems likely that Hb’s hemin molecule is transferred to Shr’s NEAT1 domain, which is proximally positioned to HID2. The domain architecture of Shr\textsuperscript{HID2-L-N1} is similar to IsdH\textsuperscript{N2N3}, a Hb- and a hemin-binding domain separated by an α-helical linker. We have previously shown that IsdH\textsuperscript{N2N3} accelerates hemin release from Hb by interactions with the linker domain, which distorts the Hb’s F-helix (45, 46). The similar domain architecture between IsdH\textsuperscript{N2N3} and Shr suggests that the linker domain in Shr may also facilitate hemin release after it engages Hb. Further insight into the mechanism of Hb capture will be gained by structural studies on the Shr:Hb complex.

2.5 Experimental Procedures

2.5.1 Cloning, Protein Expression, and Purification.

Standard methods were used to construct expression plasmids that produced the following polypeptides: Shr\textsuperscript{HID1} (residues 26-148), Shr\textsuperscript{HID2} (residues 155-285), Shr\textsuperscript{ΔHID2} (residues 175-285), Shr\textsuperscript{HID1-2} (residues 26-285), and Shr\textsuperscript{L} (residues 294-364). Proteins were expressed from pSUMO-based vectors (LifeSensors) that were transformed into Escherichia coli BL21(DE3) cells (New England Biolabs, Beverly, MA). Cultures were grown at 37°C to an \textit{A}\textsubscript{600} of 0.6 – 0.8 before induction with isopropyl-β-D-thiogalactoside (IPTG) at a final concentration of 1 mM. Induction proceeded at 25°C overnight before harvesting the cells by centrifugation. The cell pellet was then resuspended in lysis buffer containing: 50 mM NaH\textsubscript{2}PO\textsubscript{4}/Na\textsubscript{2}HPO\textsubscript{4} pH 7.0, 300 mM NaCl. A protease inhibitor cocktail (Calbiochem) and phenylmethylsulfonyl fluoride (Sigma) was also added to the lysis buffer. Cells were lysed by sonication and the lysate clarified by centrifugation (27,200 x g). The lysate was then loaded onto a Co\textsuperscript{2+}-chelating column (Thermo Scientific, Waltham, MA) and unbound proteins removed by washing the column with lysis buffer (5 column volume (CV) equivalents). Non-specifically bound proteins
were also removed by applying 5 CVs of wash buffer (lysis buffer + 10 mM imidazole) and 5 CVs of lysis buffer. The N-terminal 6x-His-SUMO tag was then cleaved with the ULP1 protease and eluted by washing the column with lysis buffer. If necessary, proteins were purified using an SEC column as described by the vendor (Sephacryl GE Healthcare Life Sciences). Human Hemoglobin (Hb) was prepared as previously described from the blood of a healthy donor provided by the CFAR Virology Core Lab at the UCLA AIDS Institute (43). Purified Hb0.1 was a generous gift of the Olson lab.

2.5.2 NMR Spectroscopy.

Uniformly $^{13}$C- and $^{15}$N-labeled Shr$^{\text{HID}2}$ was expressed in E. coli BL21(DE3) cells grown in M9 medium supplemented with $^{15}$NH$_4$Cl and $^{13}$C$_6$ glucose. The NMR sample of Shr$^{\text{HID}2}$ contained 1.1 mM $[^{15}\text{N};^{13}\text{C}]]$ Shr(155-285) dissolved in 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.0, 50 mM NaCl, 0.01% NaN$_3$ with 92%/8% H$_2$O/D$_2$O. NMR experiments were performed at 298K on Bruker DRX 500, DRX 600 MHz and AVANCE 800 MHz spectrometers equipped with triple resonance cryogenic probes. NMR spectra were processed using NMRPipe (65) and Topspin3 (66), and analyzed using CARA (67). Backbone $^1$H, $^{13}$C, and $^{15}$N chemical shifts were assigned by analyzing data from the following experiments: 2D $[^{1}\text{H}-^{15}\text{N}]$-HSQC, 2D $[^{1}\text{H}-^{13}\text{C}]$-HSQC, 3D CBCA(CO)NH, HNCACB, HNCO, HNCA, HNHA, HNHB, HBHA(CO)NH, HBNAH, CC(CO)NH (Fig. 2.5a). Assignments have been deposited to the BMRB under accession number 27550. Heteronuclear $[^{1}\text{H}]$-$^{15}$N NOE experiments were collected in an interleaved manner and analyzed using Sparky (68). The average and standard deviation of three replicates are reported (Fig. 2.5b).

NMR Hb titration experiments employed 200 μM uniformly $^{15}$N-labeled Shr constructs
that were dissolved in 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.8, 100 mM NaCl, 92% H$_2$O/8% D$_2$O. The Hb titrant was dissolved in 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.8, 100 mM NaCl and added from a 3.5 – 5.5 mM stock solution (heme units). Hb was added to the sample at a 1:1 ratio of tetramer to $^{15}$N-labeled Shr construct. The NMR competition experiments were carried using conditions that were similar to those used for the Hb titration experiments. A sample of uniformly $^{15}$N-labeled Shr$^{\text{HID1}}$ dissolved in 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.8, 100 mM NaCl, 92% H$_2$O/8% D$_2$O to a final concentration of 200 μM was prepared. Buffer-matched Hb in its carbonmonoxy state was then added such that the final ratio was 1:2, $[^{15}\text{N}]-\text{Shr}^{\text{HID1}}$:Hb. Unlabeled Shr$^{\text{HID2}}$ was then added to achieve a ratio of $[^{15}\text{N}]-\text{Shr}^{\text{HID1}}$:Hb:Shr$^{\text{HID2}} = 1:2:2$ or $1:2:4$. All samples were incubated for 30 min at room temperature prior to data acquisition. For each $^1\text{H}$-$^{15}$N HSQC experiment the number of scans was adjusted to compensate for changes in protein concentration.

2.5.3 Isothermal Titration Calorimetry (ITC).

ITC measurements were performed using a MicroCal iTC200 calorimeter (GE Healthcare) at 25°C. The Hb (in the carbonmonoxy state) and Shr constructs were buffer matched in (20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.5, 150 mM NaCl, 450 mM sucrose). The cell was filled with 25 – 35 μM Hb (concentration reported on a heme/globin chain basis) and the syringe was filled with either 700 μM Shr$^{\text{HID1}}$, 500 μM Shr$^{\text{HID2}}$, 550 – 600 μM Shr$^{\text{AHD2}}$, or 350 – 500 μM Shr$^{\text{HID1-2}}$. Twenty injections were performed using 2.0 μL injection volumes at 180 s intervals. For each of the Shr constructs, an experiment was carried out to control for the heats of dilution in which the respective Shr construct was titrated into buffer and the control data subtracted from the experimental data. ORIGIN software was used to fit the data to a single-site binding model.
2.5.4 Size Exclusion Chromatography with In-line Multi-Angle Light Scattering (SEC-MALS).

The analytical size-exclusion column WTC-030S5 (Wyatt Technology) was equilibrated in (20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.5, 150 mM NaCl) using an AKTA pure (GE Healthcare). For each protein or complex, a 100 μL (buffer-matched) sample was loaded: 1.2 mM HbCO (heme basis), 600 μM Shr$^{HID1-2}$ and the complex of HbCO at 1.2 mM and Shr$^{HID1-2}$ at 600 μM, 1.2 mM, and 1.8 mM. During elution, light scattering was measured with a miniDAWN TREOS (Wyatt Technology), and the refractive index (n) was measured with an Optilab T-rEX system (Wyatt Technology). Data were analyzed using ASTRA software (version 6.1) to obtain average molecular weights. The $dn/dc$ value (where c is the concentration) for the calculation was set to 0.185 mL/g for Hb (64). Based on the primary sequence, the $dn/dc$ values for Shr$^{HID1-2}$ and Shr$^{ΔHID2}$ were calculated to be 0.187 mL/g and 0.188 mL/g, respectively. For the complex of Shr$^{HID1-2}$ with Hb, the $dn/dc$ value was calculated to be 0.186 mL/g based on the equation for the weight average sum. Theoretical molecular weights were obtained using the ProtParam server based on the primary amino acid sequence (56).

2.5.5 Analytical Ultracentrifugation (AUC).

Sedimentation equilibrium experiments were performed at 20°C using a Beckman Optima XL-A analytical ultracentrifuge equipped with an An60-Ti rotor. Native HbA has a dimer/tetramer equilibrium which complicates the determination of the Shr$^{HID2}$:Hb binding stoichiometry by AUC. To simplify the analysis, a mutant variety of Hb which stabilizes the tetrameric form was employed (Hb0.1)(69, 70). Absorption optics at 280 nm were carried out for
Shr$^{\text{HID2}}$ alone, and at 412 nm for Hb0.1 alone or in complex with Shr$^{\text{HID2}}$ so that only the heme molecule is detected. All experiments were carried out in (20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.5, 150 mM NaCl, 450 mM sucrose). To prevent heme release, Hb0.1 in the carbonmonoxide state was used. Three mm pathlength double sector cells were used for all samples and were purged with CO before sealing the cells to prevent oxidation. Sedimentation equilibrium profiles were measured at 8,000 (~5,200 x g), 11,000 (~9,700 x g) and 26,000 rpm (~54,500 x g) for Shr$^{\text{HID2}}$ alone, with the last being used for the analysis. For Hb0.1 alone and in complex with Shr$^{\text{HID2}}$, sedimentation equilibrium profiles were measured at 11,000 (~9,700 x g) and 13,000 rpm (~13,600 x g) with the latter being used for the analysis. A third measurement was made at 30,000 rpm (~72,500 x g) and was used to determine the baseline. Samples of Hb0.1 and Shr$^{\text{HID2}}$ alone contained 5.5 μM and 248 μM of protein, respectively. For data collected at a 15:1 ratio of Shr$^{\text{D2}}$:Hb0.1, the sample contained 6.2 μM and 89 μM of Hb0.1 and Shr$^{\text{HID2}}$, respectively. For the 45:1 ratio of Shr$^{\text{HID2}}$:Hb0.1, the sample contained 5.5 μM and 248 μM of Hb0.1 and Shr$^{\text{HID2}}$, respectively. Weight-average molecular masses were determined by fitting with a non-linear least-squares exponential fit for a single ideal species using Beckman Origin-based software (version 3.01). Partial specific volumes were calculated from the amino acid compositions and corrected to 20°C. They were 0.740 for Shr$^{\text{HID2}}$, 0.749 for Hb0.1, and 0.746 for the complex.

2.5.6 X-ray Crystallography.

Crystals of Shr$^{\text{AHD2}}$ were produced from a stock of 52 mg/mL (4.2 mM) dissolved in 20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.5. Crystals were grown using the hanging drop, vapor diffusion method against a mother liquor of 0.1 M sodium acetate pH 4.5, 0.1 – 0.2 M lithium sulfate, 30 – 50% (w/v) PEG 400. Crystals developed after ~24 hours at room temperature. Data sets were
collected on two crystals: Shr$^{\text{AHiD2}}$ (native) and Shr$^{\text{AHiD2}}$ soaked for 30 s in mother liquor containing 35.9% (w/v) PEG 400 and 460 mM KI (Shr$^{\text{AHiD2}}$ + KI) (71). The KI dataset was collected on a Rigaku FRE+ generator with a copper anode equipped with a Rigaku HTC image plate detector ($\lambda = 1.5418$ Å). The Matthews coefficient indicated that there were two Shr$^{\text{AHiD2}}$ molecules in the asymmetric unit using the online server MATTPROB (72). XDS/XSCALE was used to index, integrate, and scale the data (73). Conservative resolution limits were applied to balance the calculated I/σ, R$_{\text{sym}}$, and CC1/2 in the highest resolution shell. The structure of Shr$^{\text{AHiD2}}$ was solved by single-wavelength anomalous dispersion (SAD) on Shr$^{\text{AHiD2}}$ + KI using HKL2MAP (74), the graphical user interface for the SHELXC/D/E programs. Four iodide sites were found by SHELXD (75). SHELXE (76) was used to assign the hand, produce the first set of phases, and perform solvent flattening. The final figure of merit was 0.767. Further density modification and automatic model building was performed by SHARP (77). SHARP built residues 2-111 of chain A and 8-109 of chain B in the initial model out of 112 residues in the construct. The Shr$^{\text{AHiD2}}$ + KI model was refined to 1.8 Å. This model was further refined against the high resolution 1.5 Å native data set obtained at the Advanced Photon Source on beamline 24-ID-C on a DECTRIS-PILATUS 6M detector ($\lambda = 0.9793$ Å). Initial rounds of structure refinement were carried out using Coot (78) and PHENIX (79). Later rounds of refinement were carried out using Coot (78) and BUSTER (80) with TLS refinement (81). Structure statistics are presented in Table 2. The coordinates of the final model and structure factors have been deposited in the Protein Data Bank under PDB code 6DKQ.
2.6 Figures

Figure 2.1  Domain arrangement within Shr and the protein constructs used in this study.

The top panel shows the domains that have been identified in Shr. Previous studies have shown that Shr contains an Hb-binding N-terminal region (NTR) and a hemin binding C-terminal region (CTR). Polypeptides used in this study contain regions within the NTR and include: Shr$^{\text{HID1}}$ and Shr$^{\text{HID2}}$ that contain the first and second DUF 1533 domains (renamed Hb Interacting Domains (HIDs)), respectively; Shr$^{\text{L}}$, which is predicted to contain several alpha helices; and Shr$^{\text{HID1-2}}$, which contains both of the DUF 1533 domains. The DUF 1533 domains are Hb Interacting Domains (HIDs) based on findings reported in this paper.
Figure 2.2  Hb binding studies by NMR.

Panels showing the $^1$H-$^{15}$N HSQC spectra of Shr proteins in the presence and absence of Hb. Panels (A) and (B) show the $^1$H-$^{15}$N HSQC spectra $[^{15}$N$]$-Shr$^{HID1}$ in the presence and absence of Hb (1:1 molar equivalent in tetramer units), respectively. Panels (C) and (D) show the $^1$H-$^{15}$N HSQC spectra $[^{15}$N$]$-Shr$^{HID2}$ in the presence and absence of Hb (1:1 molar equivalent in tetramer units), respectively. Panels (E) and (F) show the $^1$H-$^{15}$N HSQC spectra $[^{15}$N$]$-Shr$^L$ in the presence and absence of Hb (1:1 molar equivalent in tetramer units), respectively. Only polypeptides containing the DUF 1533 domains interact with Hb.
Figure 2.3  Biochemical Studies of Hb binding.

Panels (A) to (D) show representative ITC Hb binding data for (A) Shr$^{\text{HID1-2}}$, (B) Shr$^{\text{HID2}}$, (C) Shr$^{\text{AHID2}}$, and (D) Shr$^{\text{HID1}}$. In each titration experiment, Shr polypeptides were injected into a cell containing Hb. In each panel, at the top is shown the time course of the titration (black) and baseline (red). While the bottom part of the panel shows the integrated isotherms (squares) and the curve fit (line). As is standard, the first data point in each experiment was eliminated prior to analysis. The fifth data point in the Shr$^{\text{HID1}}$ titration was also eliminated due to a spurious double peak. Origin software was used to analyze the data. (E) SEC-MALS data defining receptor-Hb
interactions. Elution profiles of Hb alone, Shr$^{\text{HID1-2}}$ alone, and Hb in combination with Shr$^{\text{HID1-2}}$. Refractometer voltage is indicated on the left y-axis and the corresponding trace is shown in black. The elution volume is shown on the x-axis. ASTRA software was used to analyze the data. The molecular weight is indicated on the right y-axis. Interestingly, the measured Hb molecular weight is slightly smaller than its predicted molecular weight of 64.5 kDa based on its primary sequence, a discrepancy that has also been observed by others (63, 64). (F-G) Sedimentation equilibrium data defining receptor-Hb interactions. The panel shows profiles of Shr$^{\text{HID2}}$ alone (grey) at 26,000 rpm, as well as Hb alone (red) and in complex with Shr$^{\text{HID2}}$ (blue) at a 1:45 ratio of Hb0.1 (heme basis):Shr$^{\text{HID2}}$ at 13,000 rpm. The top panels show the residuals of the fit of the experimental data, while the bottom panels show the absorbance readings on the y-axis versus the radial position on the x-axis. Circles and curved black lines correspond to experimental data and calculated fits to binding models described in the text, respectively.
NMR was used to demonstrate that the isolated HIDs domain can compete for the same site on Hb. Data is similar to that shown in Figure 2. The panels show the $^1$H-$^{15}$N HSQC spectra of $[^{15}$N]-Shr$^{\text{HID1}}$. Panel (A) shows the spectrum of the $[^{15}$N]-Shr$^{\text{HID1}}$ protein in isolation, while panel (B) shows its spectrum in the presence of two-fold excess of Hb (heme basis). Panels (C) and (D) are similar to panel (B). They show the spectrum of a 2:1 (C) and a 4:1 (D) equivalent of unlabeled Shr$^{\text{HID2}}$. $[^{15}$N]-Shr$^{\text{HID1}}$ to the sample.
Figure 2.5  Assignments and Dynamics Data on Shr\textsuperscript{HID2}.

Backbone assignments of Shr\textsuperscript{HID2} overlaid onto the $^{1}\text{H}-^{15}\text{N}$ HSQC spectrum (A). To the left, the entire HSQC spectrum is shown. To the right, an inset shows an enlarged view of the center which has the greatest amount of spectral overlap. Of the Shr\textsuperscript{HID2} construct, ~94% of the residues were assigned. Unassigned residues correspond to two prolines, two residues at the N-terminus, and five residues scattered throughout the domain. $^{1}\text{H}-^{15}\text{N}$ heteronuclear NOE values (on the y-axis) shown on a per residue basis (on the x-axis) (B). Residues with missing heteronuclear NOE values were eliminated from the analysis due to either missing backbone assignments or severe spectral overlap. Shown are the average and standard deviation of three replicates.
Figure 2.6  Crystal structure of Shr$^{\Delta\text{HID2}}$.

Cartoon representation of the two Shr$^{\Delta\text{HID2}}$ molecules in the asymmetric unit (A). Chain B of the crystal structure with secondary structure elements labeled (B). Electrostatic potential calculated using PyMOL (C), shown is the same face of the domain as in (B). Areas shaded in red and blue correspond to negative and positive potentials, respectively.
Figure 2.7  Domain topology of $\text{Shr}^{\Delta\text{HID2}}$ vs. NEAT domains.

Secondary structural elements of $\text{Shr}^{\Delta\text{HID2}}$ showing the domain topology (A) compared with that of the canonical Hb-binding NEAT domains from *S. aureus* (B). Both domains are colored purple to red from the N- to C-terminus, respectively.
Figure 2.8  Hb titration experiments with Shr\textsuperscript{HID1} and Shr\textsuperscript{HID2} mutants by NMR.

Normalized height of peaks in the $^1$H-$^{15}$N HSQC spectra (y-axis) of $[^{15}$N]-Shr\textsuperscript{HID2} and $[^{15}$N]-Shr\textsuperscript{HID1} wild type (black) as well as $[^{15}$N]-Shr\textsuperscript{HID2-Y197A} and $[^{15}$N]-Shr\textsuperscript{HID1-Y55A} (green) upon titrating the sample with Hb (x-axis). Individual data points represent the average height of peaks used in the analysis and error bars represent the standard deviation of those peaks used to calculate the average. Data points were normalized to the corresponding $^1$H-$^{15}$N HSQC spectrum prior to adding Hb. Data were analyzed using Sparky and Microsoft Excel 2010. Figure was made using GraphPad Prism 5 and Adobe Illustrator CS5.
Table 2.1  Thermodynamic Parameters and Affinity Data for Hb Binding

<table>
<thead>
<tr>
<th></th>
<th>$K_D$</th>
<th>$\Delta H^\circ$</th>
<th>$\Delta S^\circ$</th>
<th>$\Delta G^\circ$</th>
<th>n $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu M$</td>
<td>kcal/mol</td>
<td>cal/mol·K</td>
<td>kcal/mol</td>
<td></td>
</tr>
<tr>
<td>Hb + Shr$^{HID1}$</td>
<td>5.1 ± 0.4</td>
<td>-13.2 ± 1.1</td>
<td>-19.9 ± 3.9</td>
<td>-7.2 ± 0.1</td>
<td>0.58 ± 0.03</td>
</tr>
<tr>
<td>Hb + Shr$^{HID2}$</td>
<td>16.0 ± 3.4</td>
<td>-7.8 ± 1.0</td>
<td>-4.3 ± 3.6</td>
<td>-6.6 ± 0.1</td>
<td>1.06 ± 0.08</td>
</tr>
<tr>
<td>Hb + Shr$^{\Delta HID2}$</td>
<td>20.8 ± 0.7</td>
<td>-9.2 ± 1.1</td>
<td>-9.5 ± 3.7</td>
<td>-6.4 ± 0.1</td>
<td>1.14 ± 0.14</td>
</tr>
<tr>
<td>Hb + Shr$^{HID1}$</td>
<td>~143</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

$^a$ Samples were in buffer containing 20 mM sodium phosphate, 150 mM NaCl, 450 mM sucrose, pH 7.5, 298K

$^b$ n refers to the molar ratio protein:Hb.

$^c$ ND refers to “not determined”.

Errors represent the standard deviation of three replicates.
# Table 2.2 | Structure Statistics on Shr^AHID2

<table>
<thead>
<tr>
<th></th>
<th>Native</th>
<th>KI soak (SAD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>APS</td>
<td>UCLA-HTC</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9793</td>
<td>1.5418</td>
</tr>
<tr>
<td>Space Group</td>
<td>P2_12_1_2</td>
<td>P2_12_1_2</td>
</tr>
<tr>
<td><strong>Cell Dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>32.47, 59.14, 102.71</td>
<td>31.51, 57.90, 102.00</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>19.4 – 1.5 (1.554 – 1.5)</td>
<td>19.4 – 1.818 (1.883 – 1.818)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>0.04259 (0.5419)</td>
<td>0.0295 (0.08212)</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt;</td>
<td>0.04649 (0.5921)</td>
<td>0.03443 (0.1068)</td>
</tr>
<tr>
<td>R&lt;sub&gt;pim&lt;/sub&gt;</td>
<td>0.01836 (0.2356)</td>
<td>0.01761 (0.06735)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>23.21 (3.34)</td>
<td>28.78 (8.12)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.75 (99.63)</td>
<td>99.66 (99.61)</td>
</tr>
<tr>
<td>No. of Total Reflections</td>
<td>206028 (19825)</td>
<td>114307 (3963)</td>
</tr>
<tr>
<td>No. of Unique Reflections</td>
<td>32457 (3194)</td>
<td>32554 (1808)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.3 (6.2)</td>
<td>3.5 (2.0)</td>
</tr>
<tr>
<td>Wilson B-factor</td>
<td>19.84</td>
<td>18.98</td>
</tr>
<tr>
<td>CC&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>0.999 (0.914)</td>
<td>0.999 (0.985)</td>
</tr>
<tr>
<td>CC*</td>
<td>1 (0.977)</td>
<td>1 (0.996)</td>
</tr>
</tbody>
</table>

**Refinement**

|                          |                         |                         |
| Resolution (Å)           | 19.7 – 1.5 (1.55 – 1.5) |                         |
| No. of Reflections       | 32457                   |                         |
| R<sub>work</sub>/R<sub>free</sub> | 0.197/0.222 (0.2388/0.2685) |                 |
| No. of atoms             | 1824                    |                         |
| Protein                  | 1696                    |                         |
| Sulfate                  | 15                      |                         |
| Water                    | 113                     |                         |
| Average B-factors        |                         |                         |
| Protein                  | 25.22                   |                         |
| Sulfate                  | 54.15                   |                         |
| Water                    | 29.91                   |                         |
| r.m.s.d.                 |                         |                         |
| Bond lengths (Å)         | 0.010                   |                         |
| Bond angles (°)          | 1.17                    |                         |
| Ramachandran Plot        |                         |                         |
| Favored (%)              | 95.35                   |                         |
| Allowed (%)              | 4.65                    |                         |
| Outliers (%)             | 0.00                    |                         |
| **PDB code**             | 6DKQ                    |                         |

†Statistics for the highest resolution shell are shown in parentheses.
Figure S2.1  Spectral recapitulation of the NTR

$^1$H-$^{15}$N HSQC spectrum of $[^{15}$N]-Shr$^{\text{HID1-2}}$ alone (top left) and with $^1$H-$^{15}$N HSQC spectra of $[^{15}$N]-Shr$^{\text{HID1}}$ and $[^{15}$N]-Shr$^{\text{HID2}}$ overlaid on top of it (top right). $^1$H-$^{15}$N HSQC spectrum of $[^{15}$N]-Shr$^{\text{HID2-L}}$ alone (bottom left) and with $^1$H-$^{15}$N HSQC spectra of $[^{15}$N]-Shr$^{\text{HID2}}$ and $[^{15}$N]-Shr$^L$ overlaid on top of it (bottom right).
Sequence alignment of the first two DUF 1533 domains from five Shr orthologues: (Sp) Streptococcus pyogenes (Accession: ABW80932); (Cn) Clostridium novyi (Accession: WP_039225525); (Si) Streptococcus iniae (WP_003102035); (Se) Streptococcus equi (Accession: WP_012678482); (Sd) Streptococcus dysgalactiae (Accession: WP_041789141). Universally conserved residues are highlighted in black while conserved mutations are highlighted in grey.

Sequence alignment carried out using Clustal Omega and figure made using BoxShade.
Table S2.1  Stoichiometry of Hb:Shr$^{\text{HID2}}$ complex determined by analytical ultracentrifugation

<table>
<thead>
<tr>
<th></th>
<th>Speed (rpm)$^{a}$</th>
<th>Weighted average molecular mass (Da)$^{b,c}$</th>
<th>Receptor/globin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shr$^{\text{HID2}}$</td>
<td>26,000</td>
<td>13,300</td>
<td></td>
</tr>
<tr>
<td>Hb0.1$^{d}$</td>
<td>13,000</td>
<td>63,900</td>
<td></td>
</tr>
<tr>
<td>Hb0.1 + Shr$^{\text{HID2}}$</td>
<td>13,000</td>
<td>92,400</td>
<td>2.0</td>
</tr>
<tr>
<td>Hb0.1$^{d}$</td>
<td>13,000</td>
<td>67,200</td>
<td></td>
</tr>
<tr>
<td>Hb0.1 + Shr$^{\text{HID2}}$</td>
<td>13,000</td>
<td>97,400</td>
<td>2.1</td>
</tr>
</tbody>
</table>

$^{a}$An (An60-Ti) rotor was used.

$^{b}$Samples were centrifuged at 20°C in buffer containing 20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, pH 7.5, 150 mM NaCl, 450 mM sucrose.

$^{c}$Molecular Masses were obtained by non-linear least squares exponential fitting of data using Beckman Origin-based software (Version 3.01).

$^{d}$Globin chains were held at a concentration of 5.5 μM

$^{e}$Receptor concentrations were in a 15-fold molar excess (83 μM)

$^{f}$Receptor concentrations were in a 45-fold molar excess (248 μM)
2.8 References


heme-IsdC complex, the central conduit of the Isd iron/heme uptake system in 

488

Staphylococcus aureus heme iron transport components IsdA and IsdB. *Infect. Immun.* **77**, 
2624–2634

42. Spirig, T., Malmirchegini, G. R., Zhang, J., Robson, S. A., Sjodt, M., Liu, M., Krishna 
Kumar, K., Dickson, C. F., Gell, D. A., Lei, B., Loo, J. a, and Clubb, R. T. (2013) 
Staphylococcus aureus uses a novel multidomain receptor to break apart human 

43. Sjodt, M., Macdonald, R., Marshall, J. D., Clayton, J., Olson, J. S., Phillips, M. L., Gell, 
Extraction from Human Hemoglobin by *Staphylococcus aureus*. *J. Biol. Chem.* **293**, 
jbc.RA117.000803

44. Dickson, C. F., Krishna Kumar, K., Jacques, D. A., Malmirchegini, G. R., Spirig, T., 
Hemoglobin-IsdH Complex Reveals the Molecular Basis of Iron Capture by 
Staphylococcus aureus. *J. Biol. Chem.* 10.1074/jbc.M113.545566


Staphylococcus aureus Protein IsdH inhibits host hemoglobin scavenging to promote heme acquisition by the pathogen. *J. Biol. Chem.* **291**, 23989–23998


80. Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A.,
Vonrhein, C., and Bricogne, G. (2012) Exploiting structure similarity in refinement:
Biol. Crystallogr.* **D68**, 368–380

Chapter 3

The *Streptococcus pyogenes* Shr$^{HID2}$ Domain in Complex with Hemoglobin Exhibits a Novel Mechanism of Capture within the Field of Gram-positive Bacterial Pathogens
The work described in this chapter will be published soon:

Ramsay Macdonald, Duilio Cascio, Robert T. Clubb. The *Streptococcus pyogenes* Shr$^{\text{HID2}}$ Domain in Complex with Hemoglobin Exhibits a Novel Mechanism of Capture within the Field of Gram-positive Bacterial Pathogens.

I was the first author of this work and conducted all the biochemical experiments and analyzed that data. Duilio Cascio and I solved the crystal structure of the complex.
3.1 Overview

In this chapter, I describe the structure of the $\text{Shr}^{\text{HID2}}:\text{Hb}$ complex. The atomic structure was determined at 2.3 Å resolution. In the structure, HID2 appears to cap the heme binding pocket of Hb. The mechanism of binding is unique and has not been reported previously. Biochemical experiments confirm this mode of binding. Furthermore, we present evidence that these interactions slow the rate of hemin release from Hb.
3.2 Introduction

*Streptococcus pyogenes* (*S. pyogenes*) is a Gram-positive opportunistic pathogen and is among the top ten causes of mortality from an infectious disease (1). Pharyngitis (aka strep throat) is the most common disease caused by this bacterial pathogen and is usually treated using penicillin (1, 2). However, this pathogen can also cause invasive infections such as Streptococcal Toxic Shock Syndrome and Necrotizing fasciitis that are much more severe resulting in mortality rates of 30% (3, 4). To combat infections caused by this deadly microbe, new therapeutics must be developed. One attractive target is the iron acquisition system in *S. pyogenes*, as the pathogen must acquire iron from its host in order to mount an infection. Human hemoglobin (Hb) is the most abundant source of iron in the body and contains ~80% of the body’s total iron supply in the form of heme (iron-protoporphyrin IX) (5). *S. pyogenes* has evolved a sophisticated protein pathway to acquire heme from Hb which is initiated by Hb binding the Shr receptor. Understanding the mechanism through which heme is extracted is of fundamental importance and could facilitate the creation of small molecule anti-infectives that work by disrupting heme acquisition.

Amongst Gram-positive pathogens, the iron acquisition system in *Staphylococcus aureus* (*S. aureus*) is the best characterized. This microbe uses the Iron-regulated Surface Determinant (Isd) system to acquire iron in the form of heme from Hb. Once *S. aureus* enters the bloodstream, it secretes hemolytic agents which lyse red blood cells (6). At this point, Hb becomes dimeric due to the dimer/tetramer equilibrium and is oxidized to metHb (7–9). Oxidized heme is hereafter referred to as hemin. Displayed on the surface of the cell wall are the two Hb-receptors in *S. aureus*: IsdH and IsdB. Both bind and accelerate hemin release from metHb (10–16). In the Isd system, hemin and Hb binding are mediated by NEAr-iron Transporter (NEAT) domains.
IsdH, IsdB, IsdA, and IsdC contain three, two, one, and one NEAT domain, respectively. Structures of all of these NEAT domains have been reported and all exhibit a generally conserved immunoglobulin like β-sandwich type fold (12, 13, 18–24). IsdH and IsdB each contain a conserved tri-domain unit consisting of two NEAT domains separated by an alpha helical linker (IsdH^{N2N3} and IsdB^{N1N2}) (12, 15). These tri-domain units alone are able to accelerate hemin release from Hb (15). In this tri-domain receptor, the first NEAT domain binds the A-helix of Hb with high affinity (16, 18, 20, 25, 26). The linker and second NEAT domains form a rigid unit which engages Hb and perturbs the hemin binding pocket (16, 27, 28). The second NEAT domain is positioned over the hemin binding pocket of Hb poised to receive the hemin molecule once it has been released into the solvent. IsdH/IsdB then pass hemin to IsdA which directly transfers it to IsdC via a “hand-clasp” mechanism (14, 29–31). IsdC transfers hemin directly to the ABC-transporter IsdDEF which transports the hemin molecule into the bacterial cytoplasm where the tetrapyrrole ring is subsequently degraded by heme oxygenases (IsdG and IsdI) to yield free iron (14, 29, 32). Similar systems exist in S. pyogenes, Bacillus anthracis, Listeria monocytogenes, and Corynebacterium diphtheriae.

The mechanism of heme acquisition by S. pyogenes is only beginning to be understood. Initially, Hb is bound by the Shr receptor on the cell surface (33). Shr has been implicated as a virulence factor as a Δshr knockout strain of S. pyogenes exhibits impaired growth when Hb is the sole source of iron and attenuated virulence in zebrafish and murine models of infection (33, 34). The purified Shr protein binds Hb, hemin, and extracts hemin from Hb (35–37). It contains two NEAT domains separated by a series of Leucine-rich repeats (36). Each NEAT domain binds hemin and have been shown to be important for S. pyogenes’s ability to grow in cultures in which Hb is the sole source of iron (36). Interestingly, Hb binding by Shr does not occur via its
NEAT domains, but instead is mediated by residues within Shr’s N-terminal region (NTR) (36, 38). The NTR contains two Domains of Unknown Function (DUF1533). It has been reported that a polypeptide containing the NTR and the first NEAT domain (NEAT1) is capable of extracting hemin from Hb (36). After hemin is removed from Hb by Shr, it is transferred to the Shp protein, which is thought to be embedded in the cell wall. Interestingly, the second NEAT domain in Shr (NEAT2) that is located at the C-terminus of the protein can rapidly transfer hemin to NEAT1, but not Shp (39). It has therefore been suggested that NEAT2 in Shr acts to store hemin near the cell surface, while Shp receives hemin from Shr’s NEAT1 domain. Heme loaded Shp then transfers hemin to HtsA at a rate that is ~100,000 fold faster than the rate of spontaneous hemin release from Shp (40, 41). This hemin transfer step involves the formation of a holoShp:apoHtsA pre-transfer complex after which the histidine and methionine residues in HtsA displace the bis-methionine ligation of the hemin-iron in Shp (42–45). Once bound by HtsA, the HtsABC ABC transporter presumably transports the hemin across the cell membrane into the bacterial cytoplasm where the tetrapyrrole ring is degraded by as of yet unidentified heme oxygenases (46).

Previously, we demonstrated that the NTR within Shr engages Hb via two Hemoglobin Interacting Domains (HIDs), called HID1 and HID2. Using isothermal titration calorimetry, we demonstrated that HID2 binds Hb with higher affinity than HID1. Moreover, we determined the crystal structure of HID2 revealing that it is a novel Hb binding domain. However, the mechanism of Hb binding remains unknown. Herein, we report the crystal structure of HID2 in complex with Hb at 2.3 Å resolution. The crystal structure reveals a unique binding mechanism in which Shr directly contacts Hb’s hemin, and its surrounding E- and F-helices. Binding studies using NMR validate the binding mechanism observed in the crystal structure. Interestingly, heme
transfer kinetics studies demonstrate that Shr interactions with Hb’s hemin effectively slows the rate at which hemin is released into the solvent. We propose that Shr engages Hb in this manner to ensure that only hemin-loaded Hb is captured by Shr for subsequent hemin removal and use by the microbe.

3.3 Results

3.3.1 The Hb:Shr$^{ΔHID2}$ Complex Exhibits an Unusual Interface.

In order understand how Shr acquires hemin from Hb, we sought to obtain a crystal structure of the Shr$^{ΔHID2}$:metHb complex. The ratio of Shr$^{ΔHID2}$:metHb set down in the drops was 2:1 with Hb in tetramer units. This ratio was selected based in the stoichiometry determined previously by AUC. Initially, crystal screens were carried out keeping the total protein concentration similar to that employed by David Gell and colleagues to obtain crystal structures of IsdH in complex with Hb (~10 mg/mL) (28, 47, 48). After these initial screens proved unsuccessful, we increased the total concentration of protein to ~45 mg/mL while maintaining the same ratio of Shr$^{ΔHID2}$ to metHb. Once this concentration of protein was used, crystals were observed (Fig. 3.1a). Given the red/brownish color of the crystals, it presumably contained metHb, but we were uncertain if the crystal contained Shr$^{ΔHID2}$ as well. In order to resolve this, one of these crystals was selected and soaked in a fresh drop of mother liquor to wash it and remove Shr$^{ΔHID2}$ that may coat the outside of the crystal. The crystal was then transferred to a fresh drop of buffer without mother liquor in which it instantly dissolved. An SDS-PAGE analysis was then carried out on this drop and revealed that the crystal in fact contained both metHb and Shr$^{ΔHID2}$ (Fig. 3.1b).
The crystal structure of the complex at 2.3 Å was obtained (Fig. 3.1c). In the asymmetric unit, there were two Hb tetramers and three ShrΔHID2 domains. Two of the ShrΔHID2 domains were bound to βHb subunits and one to an αHb. A fourth ShrΔHID2 domain exhibited partial occupancy at a second αHb subunit, but the density was insufficient to model it in. ShrΔHID2 engages the α- and β-subunits in the same manner. The structure of the complex reveals an unexpected interface. ShrΔHID2 appears to bind directly over the hemin-binding pocket of Hb, almost as if the domain is capping it. ShrΔHID2 engages Hb via the β2-α1, β4-β5, and β5-β6 loops (Fig. 3.2a-d). Previously, we identified a conserved tyrosine (Y197 in the β2-α1 loop) necessary for Hb binding. In the structure of the complex, the hydroxyl group of the tyrosine makes electrostatic contacts directly with the propionate group of the hemin moiety. The aromatic ring of Y197 appears to make hydrophobic contacts with Leu91 in the F8-helix of βHb and Leu86 in the F8-helix of αHb. Other contacts to Hb from the β2-α1 loop were observed (Fig. 3.2e-f). R196 makes electrostatic contacts with the other propionate group of the hemin molecule. Also, D199 makes electrostatic contacts with Lys95 in the F-helix of βHb. However, this contact is not observed with the αHb subunit.

Contacts to Hb were also observed from the β4-β5 loop. I224 seems to make hydrophobic contacts with A62 in βHb, yet this residue is G57 in αHb. Furthermore, D226 makes electrostatic contacts with K65 of βHb and K60 of αHb. The interaction of the β5-β6 loop with Hb occurs through M238. M238 appears to be positioned in between three hydrophobic residues: T87, L88, and L91 in βHb; and A82, L83, and L86 in αHb. In addition to these direct contacts, there are indirect contacts between ShrΔHID2 and Hb which are mediated by hydrogen bonds through water molecules. In the interaction of ShrΔHID2 with the α-subunit, the backbone carbonyl oxygen of S236 in ShrΔHID2 and the R-NH3+ group of the K61 side chain in αHb both hydrogen bond to the
same water molecule. This same water molecule also forms a hydrogen bond with the O1A atom of the propionate group in the heme moiety and furthermore forms a fourth hydrogen bond with another water molecule which itself hydrogen bonds with the carbonyl oxygen of I224 in Shr^{AHID2}. A similar hydrogen bond network is observed in the interaction of Shr^{AHID2} with βHb.

The apo form of Shr^{AHID2} overlays with the three Shr^{AHID2} chains in the structure of the complex with a backbone r.m.s.d. of 0.588 – 0.610 Å. The main difference lies in the β4-β5 loop which is essentially flipped. In the case of I224, the Cα atom was displaced ~3 Å closer with reference to the heme molecule whereas D226 is displaced ~4 Å away with reference to the heme molecule. Presumably structural distortion of the β4-β5 loop occurs to facilitate contacts between residues I224 and D226 with the above-mentioned residues in Hb. According to heteronuclear NOE data presented previously, the β4-β5 loop did exhibit slightly greater mobility. This is further supported by the fact that D226 was missing density in one of the two chains in the apo structure, likely because it was not well defined structurally. Ignoring the β4-β5 loop, the apo and holo structures of Shr^{AHID2} overlay almost identically with a backbone r.m.s.d. of 0.270 – 0.345 Å. Taken together, only the β4-β5 loop undergoes a conformational change upon binding Hb.

Neither the α- nor β-subunits of Hb are structurally perturbed upon binding of Shr^{AHID2}. Furthermore, the hemin molecule is not displaced from its binding pocket in Hb. This would seem to indicate that Shr^{AHID2} serves merely to capture the Hb and is not responsible for accelerating hemin release in any way. The same amount of surface area is buried by the interface of Shr^{AHID2} with αHb (574 Å²) as with βHb (563 and 599 Å²). However, contacts from Shr^{AHID2} with hemin account for part of that surface area. The surface area of the hemin molecule buried by Shr^{AHID2} accounts for 100 Å² in the case of the αHb and 134 – 158 Å² in the case of the
βHb. It should be noted that the surface area of the hemin molecule contacted by Shr<sup>AHID2</sup> is significantly less than that contacted by the Hb globin chain 550 – 600 Å<sup>2</sup>. It is interesting to note that the main Hb-binding domain in Shr simultaneously makes contacts with hemin and Hb. This is quite distinct from the IsdH and IsdB receptors in <i>S. aureus</i>, wherein the high-affinity Hb-binding NEAT domain engages the A-helix of Hb and the linker domain positions the hemin-binding NEAT domain above the hemin binding pocket poised to receive the hemin molecule once it has been released from Hb. To the best of our knowledge, this is a novel mode of Hb recognition for bacterial pathogens.

### 3.3.2 Identification of Important Residues in the Hb:Shr<sup>AHID2</sup> Interaction.

To ascertain which residues in Shr<sup>HID2</sup> are important for the interaction with Hb, an alanine-scan was carried out on the residues listed in Table 3.2. The effect on binding affinity was tested by titrating the respective <sup>15</sup>N-Shr<sup>HID2</sup> mutant with Hb and monitoring the broadening of chemical shifts as described previously. The D226A mutation appeared to have no effect on the binding affinity (Fig. 3.3b). This is not entirely surprising as this contact with Lys of Hb was only observed in two of the three interfaces. Surprisingly, the D199A mutation also had no effect on the ability of Shr<sup>HID2</sup> to engage Hb even though it was observed in all three interfaces (Fig. 3.3b). M238A, and to a lesser extent I224A, reduced the binding affinity of Shr<sup>HID2</sup> for Hb (Fig. 3.3a). Still, higher ratios of Hb to Shr<sup>HID2</sup> (2:1 – with Hb expressed in globin units) result in an almost complete broadening of signal (Fig. 3.3a). The R196A mutation completely abolished the interaction with Hb (Fig. 3.3a). It is interesting to note that both residues which contact the hemin molecule in Hb (R196 and Y197) are essential for the interaction. From this data, the most important contacts made to Hb are through residues: R196, Y197, I224, and M238. Of these,
only the tyrosine residue is conserved in Shr\textsuperscript{HID1}. It would seem likely that other important residues in Shr\textsuperscript{HID1} are making contacts with Hb.

### 3.3.3 \textit{Shr}\textsuperscript{ΔHID2} slows hemin release from Hb.

Due to the unusual nature of the interaction where Shr\textsuperscript{ΔHID2} occludes the hemin binding pocket of Hb and almost seems to hinder the transfer of hemin from metHb to the NEAT1 domain, the presumed binding domain of the hemin molecule once it has been released from metHb, we sought to determine if Shr\textsuperscript{ΔHID2} binding to metHb could slow the rate of hemin release from metHb. To test this, an experiment was designed to measure hemin transfer from metHb to apo-myoglobin (apoMb\textsuperscript{H64Y/V68F}) in the presence and absence of Shr\textsuperscript{ΔHID2}. Importantly, a control experiment was carried out to ensure this fragment of Shr did not interact Mb\textsuperscript{H64Y/V68F} as previously it has been reported that Shr interacts with Mb (35). In the presence of Shr\textsuperscript{ΔHID2}, the rate of hemin transfer from metHb to apoMb\textsuperscript{H64Y/V68F} is slower than in the absence of Shr\textsuperscript{ΔHID2} over the course of 30 minutes (Fig. 3.4a). As the transfer reaction is monitored over the course of ~90 hours, the transfer of hemin from metHb to apoMb\textsuperscript{H64Y/V68F} reaches completion with an expected absorbance change of 0.35 units (Fig. 3.4b). In the presence of Shr\textsuperscript{ΔHID2} however, the reaction does not reach completion with an absorbance change of only 0.15 units. Therefore, only part of Hb’s hemin was transferred to apoMb\textsuperscript{H64Y/V68F} indicating that Shr\textsuperscript{ΔHID2} slows hemin release likely by capping the hemin binding pocket and sterically occluding the release of the hemin molecule from Hb.

### 3.4 Discussion
The Shr$^{\text{HID2}}$ domain binds above the heme-binding pocket of Hb and makes contacts directly with Hb and the hemin molecule. This is a novel mode of Hb binding by Gram-positive pathogens. However, a similar mode of recognition has been reported for the Hp-Hb receptor (HpHbR) from *Trypanosoma brucei*, a single-celled eukaryotic human pathogen transmitted via the tsetse fly (49). In the case of the HpHbR, electrostatic contacts were made to the hemin propionate groups primarily through a K56 residue (49). A HpHbR$^{\text{K56A}}$ mutation bound the HpHb complex with significantly weaker affinity, but the interaction was not completely abrogated (49). It has been suggested that HpHbR makes contacts directly with the hemin moiety to allow the receptor to preferentially engage heme-bound Hp-Hb (49). In the case of Shr, contacts with the hemin molecule are essential for binding. It is possible that Shr$^{\text{HID2}}$ makes contacts directly with the hemin molecule to ensure that only hemin-loaded Hb is captured by the receptor for subsequent iron utilization by the pathogen.

In addition to the direct protein-protein and protein-hemin contacts between Shr$^{\text{AHID2}}$:Hb, there are indirect interactions mediated through hydrogen bonding with water molecules. The contribution of this hydrogen bond network to the strength of the Shr$^{\text{AHID2}}$:Hb interaction has yet to be properly tested as contacts in Shr$^{\text{AHID2}}$ are through backbone carbonyl oxygens, which cannot be removed via site-directed mutagenesis. An I224A mutation in Shr$^{\text{AHID2}}$ was shown to weaken the interaction, but it is unclear if the hydrogen bond network formed by the backbone carbonyl is maintained in this mutant and therefore it is unknown if the main contribution of I224 to the interaction is through side-chain hydrophobic contacts or through backbone hydrogen bonding.

Our kinetics data has shown that Shr$^{\text{AHID2}}$ slows the rate of hemin release from Hb. It is interesting to note that in the presence of Shr$^{\text{AHID2}}$, the Soret of metHb shifts from 405-406 nm to
407.5 nm along with an ~16% decrease in absorbance. As the binding of Shr$^{\text{AHID2}}$ to Hb does not perturb the hemin binding pocket, the exact cause of this spectral change is unclear. It is interesting to note that there is a large density between the hemin molecule and the distal histidine in all eight globin chains of the asymmetric unit. Typically for metHb, the small molecule ligand bound to hemin is water. Even with the water molecules modeled in and the distances restrained, errors of ~5σ on average appear in the difference map (Fo-Fc) strongly suggesting the presence of an additional heavy atom. This appears in every subunit of Hb regardless of the presence of Shr$^{\text{AHID2}}$. We are confident based on the UV-vis spectrum that Hb was in the met- form when the crystal trays were set up. Yet seeing as how the Hb was original purified in the carbon monoxide (CO) form, we tried to model in CO to determine if the density would accommodate this diatomic molecule with the correct geometries and distances upon refinement. While the error in the difference map was reduced, the geometries and distances were not correct for CO. It is unclear if the shift in Soret upon addition of Shr$^{\text{AHID2}}$ is related to the unexpectedly large density above the hemin molecule but it seems unlikely as this extra density is observed in every globin chain, not just the ones bound by Shr$^{\text{AHID2}}$. Taken together, we are confident that hemoglobin is in the met- form but the additional density between the hemin and distal histidine residue cannot be explained, yet it is not due to any structural distortion of the hemin binding pocket.

While the vast majority (>97%) of φ and ψ torsion angles exhibit favorable geometry in the Ramachandran plot, two residues exhibit backbone angles outside the allowed region of the Ramachandran plot. The torsion angles of I224 in Shr$^{\text{AHID2}}$ were outliers in all three molecules in the asymmetric unit. As mentioned above, the β4-β5 loop containing I224 undergoes a structural rearrangement upon binding Hb presumably to facilitate the interaction. Therefore, it seems
likely that the backbone torsion angles of I224 become kinked in such a way to accommodate the Hb ligand. Also, K87 in one of the Shr^{AHID2} molecules exhibits φ and ψ torsion angles outside of the allowed regions. K87 does not make contacts with Hb and is solvent exposed.

Taken together, we have shed light on the mechanism by which Shr captures Hb using the Shr^{AHID2} domain. Contacts with Hb’s hemin molecule are critical in this interaction. Additional contacts observed in the crystal structure confirmed the in vivo relevance of the interface observed therein. Binding of this domain to Hb slows the rate of hemin release indicating that Shr^{AHID2} occludes the hemin binding pocket preventing release of the hemin molecule.

3.5 Future Directions

Further characterization of the kinetics of hemin release will be carried out to determine if the reduced rate of hemin release from Hb by Shr^{AHID2} is dose dependent and saturateable. At this point, the maximal rate inhibition of hemin release can be determined. The nature of the Shr^{AHID2}:Hb interaction raises questions as to how the full-length Shr protein is able to extract hemin from Hb, how the NEAT1 domain presumably captures hemin, and what role the linker plays in this process. In the case of IsdH^{N2N3}, the N2 domain is flexible with respect to the rest of the protein while the linker and N3 domains form a rigid unit (18, 27, 28). It is believed that the greater flexibility of the N2 domain and the rigid junction between the linker and N3 domains is critical for function. Based on the importance of domain-domain flexibility in IsdH^{N2N3}, we are interested in probing the flexibility between the various domains in the Shr(HID2-Linker-NEAT1) tri-domain unit using Small-angle X-ray scattering (SAXS). In one case, we are interested in studying the tri-domain unit in isolation. We would like to determine if the three
domains behave like “beads-on-a-string” in that the respective domains tumble completely independently of one another; or if two of the domains form a rigid unit as in the case of IsdH$^{N2N3}$. We are also interested in studying this tri-domain unit in complex with Hb. Our goal is to determine if this tri-domain unit forms a rigid complex with Hb or if the linker and NEAT1 domains are loosely associated once the HID2 domain has bound Hb.

3.6 Experimental Procedures

3.6.1 Cloning, Protein Expression, and Purification.

Site-directed mutagenesis was used to introduce point mutations into pSUMO-Shr$^{HID2(155-285)}$ in order to make the following mutants: pSUMO-Shr$^{HID2(155-285)R196A}$, pSUMO-Shr$^{HID2(155-285)Y197A}$, pSUMO-Shr$^{HID2(155-285)D199A}$, pSUMO-Shr$^{HID2(155-285)I224A}$, pSUMO-Shr$^{HID2(155-285)D226A}$, and pSUMO-Shr$^{HID2(155-285)M238A}$. Constructs were transformed into *Escherichia coli* BL21(DE3) cells (New England Biolabs, Beverly, MA). Cultures were grown at 37°C to an $A_{600}$ of 0.6 – 0.8 before induction with isopropyl-$\beta$-D-thiogalactoside (IPTG) at a final concentration of 1 mM. Induction proceeded at 25°C overnight before harvesting the cells by centrifugation. The cell pellet was then resuspended in lysis buffer containing: 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.0, 300 mM NaCl. A protease inhibitor cocktail (Calbiochem) and phenylmethylsulfonyl fluoride (Sigma) was also added to the lysis buffer. Cells were lysed by sonication and the lysate clarified by centrifugation (27,200 x g). The lysate was then loaded onto a Co$^{2+}$-chelating column (Thermo Scientific, Waltham, MA) and unbound proteins removed by washing the column with lysis buffer (5 column volume (CV) equivalents). Non-specifically bound proteins were also removed by applying 5 CVs of wash buffer (lysis buffer + 10 mM imidazole) and 5 CVs of lysis buffer. The N-terminal 6x-His-SUMO tag was
then cleaved with the ULP1 protease and eluted by washing the column with lysis buffer. Human Hemoglobin (Hb) was prepared as previously described from the blood of a healthy donor provided by the CFAR Virology Core Lab at the UCLA AIDS Institute (25). Hb was purified in the HbCO form and converted to the metHb form as previously described (25).

### 3.6.2 X-ray Crystallography

Purified Shr\textsuperscript{ΔHID2} and metHb were dialyzed against 20 mM sodium phosphate pH 6.5 before being mixed at a 2:1 ratio of Shr\textsuperscript{ΔHID2}:metHb. Crystals were produced from a total protein concentration of ~45 mg/mL of the complex. Crystals were grown using the hanging drop, vapor diffusion method against a mother liquor of 11 – 15% PEG 6000, 240 mM lithium chloride, 100 mM sodium acetate pH 5.0. Crystals of sufficient size developed after a few days. Prior to data acquisition, crystals were cryoprotected by a 20 second soak in 20% PEG 6000, 200 mM lithium chloride, 100 mM sodium acetate pH 5.0, and 20% glycerol. Data was collected at the Advanced Photon Source on beamline 24-ID-C on a DECTRIS-PILATUS 6M detector (λ = 0.9793 Å). XDS/XSCALE was used to index, integrate, and scale the data (50). Conservative resolution limits were applied to balance the calculated I/σ, R$_{sym}$, and CC1/2 in the highest resolution shell. Phases were obtained by molecular replacement using PHASER (51) with human Hb (PDB code: 1IRD) and Shr\textsuperscript{ΔHID2} (PDB code: 6DKQ) as search models. The Matthews coefficient indicated that there were two Hb tetramers and four Shr\textsuperscript{ΔHID2} molecules present in the asymmetric unit using the online server MATTPROB (52). The density for one of the Shr\textsuperscript{ΔHID2} molecules was blurred because it fell at the two-fold axis of symmetry and was subsequently removed from the analysis. The Shr\textsuperscript{ΔHID2}:metHb complex was refined to 2.3 Å. Structure
refinement was carried out using Coot (53) and BUSTER (54) with TLS refinement (55).

Complete statistics on the structure are presented in Table 3.1.

3.6.3 NMR spectroscopy.

Uniformly $^{15}$N-labeled Shr$^{HID2}$ mutants were expressed in E. coli BL21(DE3) cells grown in M9 medium supplemented with $^{15}$NH$_4$Cl. After purification, the mutants were dialyzed against 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 6.8, 100 mM NaCl with 92%/8% H$_2$O/D$_2$O and concentrated to 200 $\mu$M. Buffer-matched HbCO (in the carbonmonoxy form) was titrated into the samples to make Hb(heme basis):Shr$^{HID2}$ at the following ratios: 0:1, 0.2:1, 0.4:1, 0.6:1, 1:1, and 2:1. An $^1$H-$^{15}$N HSQC spectrum was recorded at each titration point. NMR experiments were performed at 298K on Bruker DRX 500 and 600 MHz spectrometers equipped with triple resonance cryogenic probes. NMR spectra were processed using NMRPipe (56) and Topspin3 (57) and analyzed using Sparky (58).

3.6.4 Heme Transfer Experiments.

Purified metHb, apoMb$^{H64Y/V68F}$, and Shr$^{AHID2}$ were buffer matched in 20 mM sodium phosphate pH 7.5, 150 mM sodium chloride, and 450 mM sucrose. The apoMb double mutant (apoMb$^{H64Y/V68F}$) was used for these experiments as it exhibits a distinct absorption spectrum from native Mb and Hb allowing transfer to be more easily monitored (59). 5 $\mu$M metHb was mixed with 50 $\mu$M apoMb$^{H64Y/V68F}$ in the presence and absence of 50 $\mu$M Shr$^{AHID2}$. The change in absorbance at 405 nm (for metHb) was monitored over the course of ~90 hrs. Absorbance readings were recorded using a Shimadzu UV-1800 spectrophotometer (Shimadzu).
3.7 Figures

Figure 3.1 The Crystal Structure of the Shr<sup>ΔHID2</sup>:Hb Complex

Microscopic image of the Shr<sup>ΔHID2</sup>:Hb crystal (A). Crystals are brownish in color due to the presence of hemin. An SDS-PAGE of one of the crystals after it was washed in mother liquor and dissolved in buffer indicating that the crystal in fact contained both Shr<sup>ΔHID2</sup> and Hb (B). The asymmetric unit in the crystal structure containing two Hb tetramers and three Shr<sup>ΔHID2</sup> domains. The α-subunits are colored in green, the β-subunits are colored in beige, the Shr<sup>ΔHID2</sup> domains are colored in blue.
Figure 3.2  The Shr$^{\Delta}$HID2:Hb Interface

The interface of Shr$^{\Delta}$HID2 with the α- (A) and β-subunits (B) of Hb. The α-subunits are colored in green, the β-subunits are colored in beige, the Shr$^{\Delta}$HID2 domains are colored in blue, and the hemin molecule is colored in red. The same interfaces as in A and B but rotated by ~90° (C) and (D), respectively. Residues in Shr$^{\Delta}$HID2 important for the interaction with Hb are highlighted in grey sticks (E-F).
Figure 3.3  Hb binding studies with Shr\textsuperscript{HID2}

Based on the crystal structure of the complex, select residues were mutated to alanine in Shr\textsuperscript{HID2} to determine the importance of these residues in the interaction with Hb. In this case, each mutant was \textsuperscript{15}N-labeled and Hb was titrated into the sample at the following ratios of Shr\textsuperscript{HID2}:Hb(heme basis): 1:0, 1:0.2, 1:0.4, 1:0.6, 1:1, 1:2. The normalized average peak height on the y-axis at the given titration point on the X-axis. The errors represent the standard deviation of the peaks selected for the analysis. Mutations R196A, I224A, and M238A had the most significant effect on the interaction with Hb (A). Mutations D199A and D226A had little to no effect on the interaction (B).
Figure 3.4  Shr\textsuperscript{AHID2} slows the rate of hemin release from Hb.

On the y-axis, absorbance at 405 nm (Soret of Hb) normalized to the total absorbance change for the transfer of hemin from metHb to apoMc\textsuperscript{H64Y/V68F}. Time points are on the X-axis. The first thirty minutes are shown in (A) and the time points for the full 90 hrs. are shown in (B). Dots and squares represent the presence and absence of Shr\textsuperscript{AHID2}, respectively.
### Table 3.1 Structure Statistics on the Complex

<table>
<thead>
<tr>
<th></th>
<th>Sh\textsuperscript{AHD\textsuperscript{2}}:metHb</th>
</tr>
</thead>
</table>

**Data Collection**

<table>
<thead>
<tr>
<th></th>
<th>APS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.9793</td>
</tr>
<tr>
<td>Space Group</td>
<td>C121</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td>215.99, 116.45, 70.95</td>
</tr>
<tr>
<td></td>
<td>90, 97.21, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>60.9 – 2.3 (2.4 – 2.3)†</td>
</tr>
<tr>
<td>R\textsubscript{merge}</td>
<td>0.051 (0.676)</td>
</tr>
<tr>
<td>R\textsubscript{meas}</td>
<td>0.061 (0.845)</td>
</tr>
<tr>
<td>R\textsubscript{pim}</td>
<td>0.033 (0.428)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>14.02 (1.69)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.5 (95.7)</td>
</tr>
<tr>
<td>No. of Total Reflections</td>
<td>260587 (18671)</td>
</tr>
<tr>
<td>No. of Unique Reflections</td>
<td>76359 (5533)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.4 (3.4)</td>
</tr>
<tr>
<td>Wilson B-factor</td>
<td>53.80</td>
</tr>
<tr>
<td>CC\textsubscript{1/2}</td>
<td>0.999 (0.597)</td>
</tr>
<tr>
<td>CC*</td>
<td>1 (0.878)</td>
</tr>
</tbody>
</table>

**Refinement**

|                   | 62.5 – 2.3 (2.32 – 2.30) |
| No. of Reflections | 76358 |
| R\textsubscript{work}/R\textsubscript{free} | 0.220/0.239 (0.3043/0.3538) |
| No. of atoms       | 11625 |
| Protein            | 11236 |
| Ligands            | 344 |
| Water              | 45 |
| Average B-factors  | |
| Protein            | 65.48 |
| Ligands            | 56.56 |
| Water              | 48.62 |
| R.M.S.D.           | |
| Bond lengths (Å)   | 0.010 |
| Bond angles (°)    | 1.10 |
| Ramachandran Plot  | |
| Favored (%)        | 97.31 |
| Allowed (%)        | 2.41 |
| Outliers (%)       | 0.28 |

†Statistics for the highest resolution shell are shown in parentheses.
<table>
<thead>
<tr>
<th>Shr(^{\Delta H ID2})</th>
<th>methHb</th>
<th>Distance (Å)*</th>
<th>Contact Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>β2-α1 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg196(NH(_2))</td>
<td>heme(O1D)</td>
<td>3.0</td>
<td>electrostatic</td>
</tr>
<tr>
<td>Tyr197(O(_\beta))</td>
<td>heme(O1A)</td>
<td>2.7</td>
<td>electrostatic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu86(C(_\beta))</td>
<td>4.0</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu86(C(_\beta))</td>
<td>3.4</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu86(C(_\beta))</td>
<td>3.6</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu86(C(_\beta))</td>
<td>3.8</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>4.1</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>4.1</td>
<td>hydrophobic</td>
</tr>
<tr>
<td><strong>β4-β5 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile224(C(_\beta))</td>
<td>Gly57(C(_\beta))</td>
<td>3.9</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Asp226(O(_\delta))</td>
<td>Lys60(N(_\gamma))</td>
<td>2.8</td>
<td>electrostatic</td>
</tr>
<tr>
<td><strong>β5-β6 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met238(C(_\gamma))</td>
<td>Leu83(C(_\beta))</td>
<td>4.0</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Met238(C(_\gamma))</td>
<td>Ala82(C(_\beta))</td>
<td>3.8</td>
<td>hydrophobic</td>
</tr>
<tr>
<td><strong>β2-α1 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg196(NH(_2))</td>
<td>heme(O2D)</td>
<td>2.8 - 3.1</td>
<td>electrostatic</td>
</tr>
<tr>
<td>Tyr197(O(_\beta))</td>
<td>heme(O2A)</td>
<td>2.7</td>
<td>electrostatic</td>
</tr>
<tr>
<td>Tyr197(O)</td>
<td>Lys95(N(_\gamma))</td>
<td>3.1 - 3.2</td>
<td>electrostatic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>4.0 - 4.1</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu96(C(_\beta))</td>
<td>3.8 - 4.1</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>3.5</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu96(C(_\beta))</td>
<td>3.8 - 4.0</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>3.7 - 3.9</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Tyr197(C(_\beta))</td>
<td>Leu91(C(_\beta))</td>
<td>3.9</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Asp199(O(_\delta))</td>
<td>Lys95(N(_\gamma))</td>
<td>2.8 - 3.3</td>
<td>electrostatic</td>
</tr>
<tr>
<td><strong>β4-β5 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile224(C(_\beta))</td>
<td>Ala62(C(_\beta))</td>
<td>4.0 - 4.1</td>
<td>hydrophobic</td>
</tr>
<tr>
<td>Asp226(O(_\delta))</td>
<td>Lys65(N(_\gamma))</td>
<td>3.8</td>
<td>electrostatic</td>
</tr>
<tr>
<td><strong>β5-β6 loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met238(C(_\gamma))</td>
<td>Leu91(C(_\beta))</td>
<td>3.7 - 3.9</td>
<td>hydrophobic</td>
</tr>
</tbody>
</table>

*Only distances less than ~ 4 Å were incorporated in this table.
3.8 References


the envelope of Staphylococcus aureus. *Science.*** 299, 906–9


duplicated in genes near the components of a putative Fe3+ siderophore transporter from Gram-positive pathogenic bacteria. *Genome Biol.* 3, research0047.1-research0047.5


116
complexation through a specific protein-heme interaction: The case of the third NEAT domain of IsdH from Staphylococcus aureus. *J. Biol. Chem.* **283**, 28649–28659


118


Chapter 4

Energetics Underlying Hemin Extraction from Human Hemoglobin

by Staphylococcus aureus
This research was originally published in the Journal of Biological Chemistry:

[http://doi.org/10.1074/jbc.RA117.000803](http://doi.org/10.1074/jbc.RA117.000803)

© the American Society for Biochemistry and Molecular Biology  
Copyright 2018  
Journal of Biological Chemistry

I contributed to the published work by collecting and analyzing the ITC and NMR data. I also wrote part of the results and discussion sections and assisted with the revisions to the manuscript.
4.1 Overview

*Staphylococcus aureus* is a leading cause of life-threatening infections in the United States. It actively acquires the essential nutrient iron from human hemoglobin (Hb) using the iron-regulated surface-determinant (Isd) system. This process is initiated when the closely related bacterial IsdB and IsdH receptors bind to Hb and extract its hemin through a conserved tri-domain unit that contains two NEAr iron Transporter (NEAT) domains that are connected by a helical linker domain. Previously, we demonstrated that the tri-domain unit within IsdH (IsdH\textsuperscript{N2N3}) triggers hemin release by distorting Hb’s F-helix. Here, we report that IsdH\textsuperscript{N2N3} promotes hemin release from both the α and β subunits. Using a receptor mutant that only binds to the α-subunit of Hb and a stopped-flow transfer assay, we determined the energetics and micro-rate constants of hemin extraction from tetrameric Hb. We found that at 37°C, the receptor accelerates hemin release from Hb up to 13,400-fold, with an activation enthalpy of 19.5 ± 1.1 kcal/mol. We propose that hemin removal requires the rate-limiting hydrolytic cleavage of the axial HisF8 Nε-Fe\textsuperscript{3+} bond, which, based on molecular dynamics simulations, may be facilitated by receptor-induced bond hydration. Isothermal titration calorimetry experiments revealed that two distinct IsdH\textsuperscript{N2N3}–Hb protein–protein interfaces promote hemin release. A high-affinity receptor:Hb(A-helix) interface contributed ~95% of the total binding standard free energy, enabling much weaker receptor interactions with Hb’s F-helix that distort its hemin pocket and cause unfavorable changes in the binding enthalpy. We present a model indicating that receptor-introduced structural distortions and increased solvation underlie the IsdH-mediated hemin extraction mechanism.
4.2 Introduction

*Staphylococcus aureus* is an opportunistic Gram-positive pathogen that causes a wide range of illnesses such as pneumonia, meningitis, endocarditis, toxic shock syndrome, bacteremia, and septicemia (1). Methicillin resistant *S. aureus* (MRSA) and other drug resistant strains are a leading cause of life-threatening hospital- and community-acquired infections in the United States (2, 3). *S. aureus* requires iron to proliferate, which it actively procures from its human host during infections (4-7). Heme (protoporphyrin IX + iron) present within hemoglobin (Hb) accounts for ~75-80% of the total iron found in the human body and is preferentially used by *S. aureus* as an iron source (6, 7). Consequently, *S. aureus* and other microbial pathogens have evolved elaborate heme-acquisition systems to exploit this rich iron source. An understanding of the molecular mechanism of heme scavenging will provide insight into how this pathogen survives and persists within its human host and it could lead to new anti-infective agents that work by limiting microbial access to iron.

*S. aureus* uses the iron-regulated surface determinant (Isd) system to extract the oxidized form of heme from Hb (hereafter called hemin). The Isd system is comprised of nine proteins that form a hemin relay system that captures Hb and rapidly extracts its hemin, transfers hemin across the cell wall, and then pumps the hemin into the cytoplasm where it is degraded to release free iron (8–10). Four Isd proteins (IsdA, IsdB, IsdC, and IsdH/HarA) are covalently linked to the bacterial cell wall via sortase transpeptidases (11, 12). The two surface receptor proteins, IsdH and IsdB, first bind Hb and extract its hemin (13–17). Then the hemin is passed to IsdA, which is partially buried in the cell wall (18). Holo-IsdA subsequently transfers hemin to the fully buried IsdC protein via an ultra-weak handclasp complex (19–22). Lastly, holo-IsdC delivers the hemin to the bacterial ABC transporter complex, IsdDEF, which pumps hemin into the cytoplasm where it is degraded to release free iron.
degraded to release free iron by the hemin oxygenases, IsdG or IsdI (23, 24). The Isd system is important for *S. aureus* virulence, as strains lacking genes that encode for components of the system display attenuated infectivity in mouse models and are impaired in their ability to utilize hemin or Hb as the sole source of iron (7, 11, 25–27). Related hemin acquisition systems are used by other Gram-positive pathogens to obtain iron.

Hemin is extracted from the oxidized form of Hb (called methemoglobin) using the *S. aureus* IsdB and IsdH receptors (15, 16, 28–31). The receptors share similar primary sequences and contain NEAr iron Transporter (NEAT) domains, binding modules originally named because they were found in bacterial genes that are proximal to putative Fe$^{3+}$ siderophore transporter genes (32). Our previously published studies of IsdH have shown that it extracts hemin using a tridomain unit formed by its second (N2) and third (N3) NEAT domains, which are connected by a helical linker domain (IsdH$^{N2N3}$, residues A326-D660) (Fig. 4.1a). Studies using native MetHb have shown that IsdH$^{N2N3}$ rapidly extracts hemin from Hb, ~500-fold faster than the rate at which hemin spontaneously dissociates from Hb. Recent NMR and crystallography studies of the Hb-free and bound forms of IsdH have provided insight into how it promotes hemin release from Hb (29–31). Crystal structures of the Hb:IsdH$^{N2N3}$ complex reveal that the receptor adopts an extended dumbbell shape structure in which the N2 domain engages the A- and E-helices of Hb, while the N3 domain is positioned within 12 Å of the hemin molecule that is located on the same globin chain (Fig. 4.1b). Hemin release may in part be triggered by the helical linker domain, which distorts the F-helix in Hb that houses the iron-coordinating proximal histidine residue (HF8) (Fig. 4.1c). NMR studies of the apo-receptor uncovered the presence of inter-domain motions between the N2 and linker domains, revealing that the receptor adaptively recognizes Hb. IsdB shares significant primary sequence homology with IsdH and actively acquires hemin from Hb,
suggesting that it also employs a semi-flexible tridomain unit to extract Hb’s hemin (16, 22, 28, 33).

While the structural basis of IsdH binding to Hb is now understood, we lack a quantitative understanding of the mechanism of hemin transfer. This is because the heterogeneous nature of the transfer reaction, which has thus far limited its detailed biophysical characterization. For example, at the concentrations used to monitor hemin release, Hb adopts distinct tetrameric and dimeric quaternary states that have different propensities to release hemin (34, 35). Moreover, IsdH binds to both the α- and β-globin chains within Hb, which also release hemin at different rates (36, 37). Finally, hemin extraction causes Hb to dissociate into its component globin chains, which irreversibly aggregate and release hemin faster than intact Hb (38, 39). Thus, previously measured rates of hemin transfer using wild-type Hb and IsdH report on an amalgamation of distinct extraction processes, hindering a quantitative understanding of the energetics that underlie hemin extraction. To overcome these limitations we used a stabilized tetrameric form of Hb and a bacterial receptor mutant that only binds to Hb’s α-subunit. We show using stopped-flow experiments that the tridomain unit accelerates hemin release from the α-subunit by more than 10,000 fold, a rate that is significantly faster than previously reported values that were measured using heterogeneous components. Transfer requires a Gibbs activation energy of 18.1 ± 1.1 kcal/mol to be surmounted, whose high enthalpic component suggests that in the transition state the proximal histidine-hemin bond (Nε-Fe) bond is hydrolytically cleaved. Molecular dynamics (MD) simulations suggest that the receptor’s ability to lower the activation energy by ~6 kcal/mol is due to its ability to induce strain in the bond and to increase the concentration of nearby water molecules that compete with Hb’s HisF8 Nε atom for the iron atom within hemin. Quantitative measurements of Hb-receptor binding interactions using Isothermal titration calorimetry (ITC)
indicate that two energetically distinct receptor-Hb interfaces form the pretransfer complex. An interface between the receptor’s N2 domain and the A-helix of Hb is the primary contributor to overall binding affinity, tethering IsdH to Hb to enable lower affinity interactions that distort Hb’s hemin pocket. A model of the hemin extraction process is presented that incorporates the existing structural, kinetics and thermodynamic binding data.

4.3 Results

4.3.1 Development of a quantitative assay to study hemin extraction

We developed an assay to quantitatively measure the kinetics of IsdH-mediated hemin removal from the α-globin chain within tetrameric Hb. In the assay a mutant of IsdH$^{N2N3}$ that only binds to the α-subunit of Hb is used (α-IsdH, residues A326-D660, with $^{365}$FYHYA$^{369} \rightarrow ^{365}$YYHYF$^{369}$ mutations). α-IsdH contains F365Y and A369F mutations within the alpha-helix of the Hb binding N2 domain, which selectively weakens its affinity for the β-subunit in Hb (31). The assay also employs Hb0.1, a tetramer stabilized recombinant form of Hb in which the two α-subunits are present within a single polypeptide (the C-terminus of one α-subunit is joined via a glycine residue to the N-terminus of the second α-subunit (40)). Hb0.1 also contains a V1M mutation that facilitates expression in E. coli, but is otherwise structurally and functionally identical to wild-type Hb. Because Hb0.1 has a significantly reduced propensity to dissociate into αβ Hb dimers, hemin transfer experiments employing it and α-IsdH enables hemin extraction from the α-globin chain within tetrameric Hb to be selectively monitored (35, 40).

As summarized in Table 1, analytical ultracentrifugation (AUC) sedimentation equilibrium experiments confirm that α-IsdH binds tetrameric Hb0.1 with the appropriate 2:1 stoichiometry
via interactions with its α-subunits. In the AUC experiments, to prevent complications caused by hemin transfer, Hb0.1 was maintained in its carbonmonoxy ligated state and hemin-binding deficient IsdH proteins were used that contain a Y642A mutation that removes the iron coordinating tyrosine residue (41). In all AUC experiments the concentration of Hb0.1 was monitored by measuring its absorbance at 412 nm, the Soret band of the bound heme molecule. As expected, isolated Hb0.1 forms a stable tetramer at concentrations that are used in hemin transfer experiments; at 5 µM (heme units), Hb0.1 has a weighted average molecular mass of 64,212 ± 829 Da at 13,000 × g. We next used AUC to determine the Hb0.1:α-IsdH\textsuperscript{Y642A} binding stoichiometry when α-IsdH\textsuperscript{Y642A} is present at a 15-fold molar excess. This data indicates that ~1.8 α-IsdH\textsuperscript{Y642A} proteins bind to each Hb0.1 tetramer, compatible with it only binding to the α-subunits. Increasing the ratio of α-IsdH\textsuperscript{Y642A} to Hb0.1 in the AUC experiments to 30:1 does not increase this stoichiometry, indicating that Hb0.1 is saturated (Fig. 4.2a, red curve). In contrast, IsdH\textsuperscript{Y642A}, which is identical to α-IsdH\textsuperscript{Y642A} except that it does not contain mutations in the Hb contacting alpha-helix, binds to Hb0.1 with a stoichiometry of ~3.2 (Fig. 4.2a, blue). This is consistent with previous studies that have shown that IsdH\textsuperscript{Y642A} binds to both the α- and β-subunits of Hb (31, 42). To confirm that α-IsdH\textsuperscript{Y642A} interacts with only the α-subunits within Hb0.1, AUC experiments were performed using the β-CO tetramer, which only contains the isolated β-globin protein (Fig. 4.2b). Consistent with previously reported studies, AUC analysis of the β-CO form indicates that it adopts a tetrameric structure; molecular mass of 56,700 ± 724 Da at 6000 × g (56,300 ± 508 Da at 13,000 × g) (35, 43). When the AUC studies are repeated in the presence of excess IsdH\textsuperscript{Y642A}, the receptor binds to the β-CO tetramer with ~2:1 stoichiometry indicating that it can bind to the β-globin chain, but steric effects caused by tetramerization may limit receptor
occupancy (molecular mass of 134,000 ± 1321 Da at 6000 × g) (Fig. 4.2b, blue). Interestingly, when similar experiments are performed using α-IsdH<sup>Y642A</sup>, no detectable binding is observed (Fig. 4.2b, red). Overall, the results of these studies indicate that the α-IsdH<sup>Y642A</sup> receptor only binds to the α-subunits within Hb0.1 and that both of these subunits are nearly saturated with the receptor at a ratio of 15:1, α-IsdH<sup>Y642A</sup>:Hb0.1 (Hb0.1 concentration 5 μM heme units). These conditions and reagents were therefore used in subsequent stopped-flow hemin transfer experiments to preferentially measure the rate at which the receptor extracts hemin from the α-subunit of tetrameric Hb0.1.

### 4.3.2 Hb tetramerization slows hemin removal from the α-subunit

Stopped-flow UV-Vis spectroscopy was used to measure the rate at which α-IsdH extracts hemin from either the stabilized Hb0.1 tetramer or native Hb (results summarized in Table 4.2). In these experiments, the proteins were rapidly mixed, and the amount of hemin transfer was determined by monitoring UV absorbance changes at 405 nm (29, 35). A 30-fold molar excess of receptor was employed (5 and 150 μM of Hb and receptor, respectively) to ensure that α-IsdH saturates the α-globin subunit (Table 1). Upon mixing, a rapid shift in the UV-spectrum occurs indicating hemin transfer from Hb0.1 to α-IsdH (Fig. 4.3a). The time-dependent spectral changes exhibit biphasic kinetics, with approximately half of the total magnitude of the spectral change characterized by a rapid event defined by the rate constant \( k_{\text{fast}} \) and the remaining, much slower changes, described by \( k_{\text{slow}} \) (Fig. 4.3b). Importantly, hemin transfer is completed during the time course of the experiment, as minimal additional spectral changes are observed when the proteins are incubated overnight. As the receptor is in excess and has higher affinity for hemin than hemoglobin (described below), we conclude that all of the hemin in Hb0.1 is captured by α-IsdH.
The observed pseudo-first-order rate constants, $k_{\text{fast}}$ and $k_{\text{slow}}$, are $0.34 \pm 0.01 \text{ s}^{-1}$ and $0.026 \pm 0.001 \text{ s}^{-1}$ at $25^\circ\text{C}$, respectively. As $\alpha$-IsdH only binds to the $\alpha$-subunits of the stabilized Hb0.1 tetramer, we conclude that the rapid changes characterized by $k_{\text{fast}}$ describe hemin removal from the $\alpha$-subunit within tetrameric Hb0.1, while the slower changes result from indirect hemin capture from the $\beta$-subunit after it has first been released into the solvent. Indirect hemin capture from the $\beta$-subunit is consistent with the value of $k_{\text{slow}}$, as it is similar in magnitude to the previously reported rate at which the $\beta$-subunit spontaneously releases hemin into solvent from the semihemoglobin form of tetrameric Hb ($0.007-0.013 \text{ s}^{-1}$) (35, 39). Moreover, the value of $k_{\text{slow}}$ is independent of receptor concentration (Fig. 4.5b). The tetrameric $\alpha$-subunit spontaneously releases hemin into the solvent at a rate of $\sim 0.000083 \text{ s}^{-1}$ at $37^\circ\text{C}$. Thus, $\alpha$-IsdH accelerates the rate of hemin release from the $\alpha$-subunit $\sim 13,400$-fold, and scavenges hemin from the $\beta$-subunit indirectly via the solvent.

To investigate the impact of the dimer-tetramer Hb equilibrium on hemin capture we measured transfer rates to $\alpha$-IsdH from Hb, which exists as a mixture of dimeric ($\alpha\beta$) and tetrameric ($\alpha_2\beta_2$) species that have differing propensities to release hemin (35). Stopped-flow experiments indicate that hemin is rapidly removed from Hb, with $k_{\text{fast}}$ and $k_{\text{slow}}$ rate constants of $3.27 \pm 0.07 \text{ s}^{-1}$ and $0.041 \pm 0.001 \text{ s}^{-1}$ at $25^\circ\text{C}$, respectively. As for Hb0.1, $k_{\text{fast}}$ accounts for approximately half of the overall spectral change that occurs upon receptor mixing with Hb, consistent with it reporting on the rate of hemin removal from the $\alpha$-subunit. Interestingly, as compared to Hb0.1, the $k_{\text{fast}}$ for hemin removal from Hb is increased $\sim 9$-fold (Fig. 4.3c). This is consistent with the presence of dimeric species in the reaction containing Hb that are known to release hemin more rapidly than the tetrameric Hb, and the fact that upon hemin removal, Hb can fully dissociate into monomeric $\alpha$-globin chains that have even weaker affinity for hemin (15, 35). More modest 2-fold increases in $k_{\text{slow}}$ are also observed when Hb is used instead of Hb0.1 (Fig. 4.5c).
4.3d. They are consistent with the production of isolated β-subunits produced upon Hb dissociation that are known to more rapidly release hemin than tetrameric Hb0.1 (35). However, as this process occurs slowly, α-IsdH presumably extracts hemin from the β-subunits through an indirect process in which the hemin is first released into the solvent.

Hemoglobin tetramer dissociation also accelerates the rate of hemin capture by the native IsdHN2N3 receptor that binds to both the α- and β-globin chains. This was demonstrated by measuring the rate of hemin transfer to IsdHN2N3 from either Hb or the stabilized Hb0.1 tetramer (Fig. 4.3b). With IsdHN2N3, as with α-IsdH, hemin is removed from Hb0.1 more slowly than it is from Hb, consistent with the reduced propensity of Hb0.1 to dissociate into more labile dimeric and monomeric species upon hemin removal (Fig. 4.3c).

Therefore, the IsdHN2N3 receptor’s faster kinetics of hemin removal from Hb0.1 may be caused by its ability to also remove hemin from hemoglobin’s β-subunit, which is known to release hemin more readily than the α-subunit (35). Combined, these data indicate that the propensity of Hb to dissociate during the extraction process increases the overall observed rate of transfer to the receptor by producing partially hemin ligated states (semi-forms) or dissociated forms of hemoglobin that have a greater propensity to release their hemin molecules.

4.3.3 IsdH extracts hemin from both the α- and β-subunits in an effectively irreversible process.

The stopped-flow and AUC data indicate that IsdH actively extracts hemin from the α-subunit (Figs. 4.2-3), but it is not known if the receptor also triggers hemin release from the β-subunit. We therefore probed the ability of IsdHN2N3 to remove hemin from tetrameric Metβ, which only contains β-globin chains bound to hemin (Fig. 4.4). Active removal is observed when
IsdH\textsuperscript{N2N3} is added to Met\(\beta\), with rapid changes in the absorbance spectrum observed within 5 seconds of mixing. A detailed interpretation of the kinetics data is not possible because the \(\beta\)-globin chains aggregate after hemin is removed. However, extraction from Met\(\beta\) occurs at a rate of \(\sim 3 \text{ s}^{-1}\), which is similar to the rate at which IsdH\textsuperscript{N2N3} removes hemin from Hb and about 4-times faster than Hb0.1. This is compatible with previous studies that have shown that hemin loss from Met\(\beta\) tetramers is comparable to that of \(\beta\)-subunits within tetrameric Hb (35), and it suggests that IsdH\textsuperscript{N2N3} may remove hemin more rapidly from the \(\beta\)-subunit within Hb than from its \(\alpha\)-subunit.

In contrast to IsdH\textsuperscript{N2N3}, only modest spectral changes are observed when \(\alpha\)-IsdH is added to Met\(\beta\) (Fig. 4.4). This is compatible with \(\alpha\)-IsdH’s inability to bind the \(\beta\)-CO tetramer (Fig. 4.2b), such that it must indirectly scavenge hemin after it has first been released into the solvent. This supported by control transfer studies using recombinant H64Y/V68F apo-myoglobin (apo-Mb containing H64Y and V68F mutations), which does not physically interact with Met\(\beta\) and like \(\alpha\)-IsdH slowly scavenges its hemin (Fig. 4.4) (39). Combined, these results indicate that native IsdH\textsuperscript{N2N3} actively removes hemin from both the \(\alpha\)- and \(\beta\)-globin chains within human hemoglobin.

To the best of our knowledge, the relative hemin binding affinities of IsdH\textsuperscript{N2N3} and Hb have not been reported. We therefore performed hemin binding competition assays in which each protein was separately mixed with the apo-Mb reagent, which has a known affinity for hemin and whose hemin binding can readily be measured at 600 nm (39). Mixing either 5 \(\mu\)M of Hb or Hb0.1 with a 10-fold molar excess of apo-Mb results in a measurable increase in absorbance at 600 nm indicating that hemin is transferred to apo-Mb (Fig. 4.5a). As expected, both transfer processes occur slowly, compatible with hemin first being released into the solvent by hemoglobin before it is subsequently bound by the apo-Mb reagent. These data indicate that apo-Mb has significantly higher affinity for hemin than Hb and Hb0.1. This is compatible with previously reported hemin
dissociation constants ($K_H$) of Hb and apo-Mb; the $K_H$ of apo-Mb is ~0.03 pM, whereas $K_H$ for dimeric Hb is 1.7 and 42 pM for its $\alpha$ and $\beta$ chains, respectively, and the $K_H$ for tetrameric Hb is 0.8 and 4.2 pM for its $\alpha$ and $\beta$ chains, respectively (35). Notably, the transfer data reveal that Hb0.1 releases hemin more slowly than Hb (Fig. 4.5a), consistent with previous studies that have shown that the tetrameric form of Hb releases hemin more slowly than the dimeric form (35). In contrast, when hemin containing IsdH$^{N2N3}$ is mixed with the apo-Mb reagent only small absorbance changes are observed at 600 nm, indicating that unlike Hb and Hb0.1, the receptor has higher affinity for hemin than apo-Mb. Similar to IsdH$^{N2N3}$, mixing $\alpha$-IsdH with apo-Mb results in minimal absorbance changes (data not shown). This is expected as mutations in $\alpha$-IsdH that confer $\alpha$-globin binding selectivity are located in the N2 domain and are distal to the hemin binding pocket located in the N3 domain. Taken together, we conclude that IsdH$^{N2N3}$ and $\alpha$-IsdH bind hemin with $K_H$ values < ~0.03 pM, such that they have substantially higher affinity for hemin than Hb or Hb0.1. Therefore, during the time-frame of the stopped-flow experiments transfer of hemin from hemoglobin to the receptor is effectively unidirectional and irreversible.

4.3.4 Determination of the micro-rate constants describing hemin removal from the $\alpha$-subunit.

To estimate the micro-rate constants that describe hemin removal, the kinetics data were interpreted using scheme I. In this reaction, hemin-bound Hb0.1 first forms a complex with $\alpha$-IsdH described by the association and dissociation rate constants, $k_1$ and $k_{-1}$, respectively. Within the Hb[hemin]:$\alpha$-IsdH pretransfer complex, hemin is transferred from the $\alpha$-subunit of tetrameric Hb0.1 to $\alpha$-IsdH in a process described by $k_{\text{trans}}$. This yields a semi-form of Hb0.1 in which one hemin molecule has been removed and the protein remains bound to $\alpha$-IsdH.
At the conditions used in our stopped-flow experiments, measured values of $k_{\text{fast}}$ describe the kinetics of this entire process and account for the simultaneous removal of hemin from both $\alpha$-subunits. Contributions to $k_{\text{fast}}$ caused by hemin removal from the $\beta$-subunit are also possible, but are expected to be minimal as $\alpha$-IsdH does not bind to this subunit or actively extract its hemin (Fig. 4 and Table 1). To estimate the micro-rate constants, stopped-flow hemin transfer experiments were performed using varying amounts of $\alpha$-IsdH in its hemin-free form. At all receptor:Hb0.1 stoichiometries $k_{\text{fast}}$ accounted for ~50% of the total spectral change consistent with it monitoring hemin removal from only the $\alpha$-subunit (Fig. 4.3b), and the pseudo first order values for $k_{\text{fast}}$ increased hyperbolically with increasing receptor concentrations (Fig. 4.5b, dark grey circles). This is compatible with the rapid formation of a Hb:$\alpha$-IsdH complex, followed by rate-limiting transfer of hemin to the receptor. In the transfer experiments, [heme-free $\alpha$-IsdH] is $\geq 10[Hb0.1]$, such that the rate of hemin transfer is pseudo-first-order and the fraction of pre-transfer complex is small and in steady state during the reaction. The dependence of $k_{\text{fast}}$ on the micro-rate constants and receptor concentration can therefore be described by Eqn. 1 (44),

$$k_{\text{fast}} = \frac{k_1 k_{\text{trans}}[\alpha-\text{IsdH}]}{k_1[\alpha-\text{IsdH}] + k_{-1} + k_{\text{trans}}} \quad \text{(Eqn. 1)}$$

where $k_1$, $k_{-1}$, and $k_{\text{trans}}$ are the micro-rate constants of the individual reaction steps. As shown in Fig. 4.5b, fitting of the kinetics data using Eqn. 1, yields values of $2.8 \times 10^7$ M$^{-1}$s$^{-1}$, $2.4 \times 10^3$ s$^{-1}$, and $0.57 \pm 0.02$ s$^{-1}$ for $k_1$, $k_{-1}$, and $k_{\text{trans}}$ respectively. The fitted value of $k_{-1}$ is much larger than $k_{\text{trans}}$, and as result, only the ratio $k_{-1}/k_1$ and not the absolute values of these constants are well defined.
by the observed data. The equilibrium dissociation constant (K_D) for the Hb0.1:α-IsdH pretransfer complex calculated from the micro-rate constants is 85.7 ± 3.8 μM (k_1/k_1 = K_D), a value that is consistent with dissociation constants measured using ITC (described below). The k_slow values do not depend on the concentration of the receptor, consistent with it describing absorbance changes caused by hemin release into the solvent from the β-globin chain followed by uptake by α-IsdH (Fig. 4.5b, light grey circles).

4.3.5 An enthalpic barrier must be surmounted to capture hemin.

Stopped-flow experiments were performed at temperatures ranging from 15 to 37°C using a molar ratio of 1:40 Hb0.1:α-IsdH (Fig. 4.6a). At this Hb:receptor ratio the contribution of k_1/k_1 to k_fast becomes negligible, such that k_fast ≈ k_trans (44). A plot of ln(k_trans/T) versus 1/T enables the activation parameters to be determined using the linearized form of the Eyring equation,

\[
\ln \frac{k}{T} = \ln \left(\frac{k_B}{h}\right) + \frac{\Delta S^\ddagger}{R} - \frac{\Delta H^\ddagger}{RT} \quad \text{(Eqn. 2)}
\]

where k_B, h, and R are the Boltzmann, Planck, and ideal gas constants, respectively. The activation enthalpy (ΔH^\ddagger) and entropy (ΔS^\ddagger) are calculated from slope and intercept, respectively (Fig. 4.6b). This analysis yielded ΔH^\ddagger and ΔS^\ddagger values of 19.45 ± 1.1 kcal/mol and 4.45 ± 0.31 cal/mol·K, respectively, and a ΔG^\ddagger value of 18.07 ± 1.1 kcal/mol at 37°C. Previous studies have shown that hemin dissociates from the α-globin chain of tetrameric Hb at a rate of 0.000083 s^{-1} (35), which based on the Arrhenius equation indicates that the ΔG^\ddagger for spontaneous hemin release is 23.96 kcal/mol at 37°C. Thus, the receptor lowers the activation energy needed to release hemin from the α-subunit by ~5.9 kcal/mol, accelerating the rate ~10,000-fold (k_trans = 0.88 s^{-1} vs. 0.000083 s^{-1} at 37°C).
4.3.6 Two energetically distinct receptor-Hb interfaces are used to bind Hb and distort its hemin pocket.

The crystal structure of the Hb:IsdH$^{N2N3}$ complex reveals that the receptor alters the structure of the hemin binding pocket within Hb, presumably weakening its affinity for hemin (31). Two receptor-Hb interfaces are present in the structure, the N2 domain binds to Hb’s A-helix, while the linker and N3 domains contact the F-helix within the same globin (Fig. 4.1b). These contact surfaces are extensive, burying 618 and 688 Å$^2$ of surface area, respectively. Previously, we have shown that polypeptides containing only the N2 (IsdH$^{N2}$) and linker-N3 (IsdH$^{LN3}$) domains of the receptor are autonomously folded, and that IsdH$^{LN3}$ forms a rigid structure (15, 29). Using these protein constructs ITC experiments were performed to gain insight into how each interface contributes to the energetics of Hb binding (Table 4.3). Initially, binding of $\alpha$-IsdH$^{Y642A}$ to the carbonmonoxy form of Hb (HbCO) was investigated. As expected, no heme transfer occurs between these proteins when monitored by UV-vis spectroscopy (not shown) enabling the protein binding energetics to be quantified. In the ITC experiments, a syringe filled with 200 µM $\alpha$-IsdH$^{Y642A}$ was injected incrementally into a cell containing 30 µM of HbCO, and the ensuing heat changes were used to generate a binding isotherm (Fig. 4.7a). Consistent with the AUC data, $\alpha$-IsdH$^{Y642A}$ binds Hb at a ratio of ~2:1 (Tables 4.1 and 4.3). The measured $K_D$ value for complex is 4.0 $\pm$ 0.6 µM and is of similar magnitude as the value obtained from an analysis of the kinetics data. At 25°C, the standard enthalpy ($\Delta H^o$) and entropy ($\Delta S^o$) of receptor binding are $-6.0 \pm 0.5$ kcal/mol and $4.6 \pm 1.9$ cal/mol-deg, respectively ($\Delta G^o = -7.4 \pm 0.1$ kcal/mol). To selectively probe binding by the N2 domain, a polypeptide containing only the N2 domain of $\alpha$-IsdH with the appropriate amino acid mutations that confer binding selectivity for the $\alpha$-globin chain was studied.
($\alpha$-IsdH$^\text{N2}$, residues A326-P466 of $\alpha$-IsdH). ITC measurements indicate that $\alpha$-IsdH$^\text{N2}$ binds HbCO with a $K_\text{D}$ of $7.5 \pm 1.1$ $\mu$M and similar ~2:1 stoichiometry (Fig. 4.7b). Interestingly, the entropy change associated with HbCO binding is unfavorable for $\alpha$-IsdH$^\text{N2}$ ($\Delta S^\circ = -11.1 \pm 2.0$ kcal/mol-deg), but favorable for intact $\alpha$-IsdH$^\text{Y642A}$. This difference likely arises from unfavorable ordering of the alpha-helix within the N2 domain that accompanies Hb binding (30, 45), which, in the context of the longer $\alpha$-IsdH$^\text{Y642A}$ protein, are masked by other entropic changes caused by interactions originating from the linker and N3 domains (29, 45).

ITC binding studies were also performed using a polypeptide that contains only the linker and N3 domains of IsdH (IsdH$^\text{LN3}$, residues P466-D660, containing a Y642A mutation in the N3 domain to prevent hemin binding). IsdH$^\text{LN3}$ binds to HbCO, but these interactions are too weak to be accurately quantified. Nevertheless, it is evident from the ITC data that HbCO binding by IsdH$^\text{LN3}$ is endothermic (Fig. 4.7c), consistent with the crystal structure of the complex that revealed that this portion of the receptor partially unwinds Hb’s F-helix. NMR spectroscopy experiments confirm that IsdH$^\text{LN3}$ and HbCO interact weakly, as cross-peaks in the $^1\text{H}-^1\text{N}$ HSQC spectrum of $^{15}\text{N}$-enriched IsdH$^\text{LN3}$ are attenuated and broadened only when large amounts of unlabeled HbCO are added (Fig. 4.7d, right). The binding affinity is very weak, with an estimated $K_\text{D}$ in the mM range. Weak HbCO-IsdH$^\text{LN3}$ interactions require both the linker and N3 domains in IsdH$^\text{LN3}$, as no interactions are observed when HbCO is added to a $^{15}\text{N}$-enriched IsdH$^\text{N3}$ polypeptide that only contains the N3 domain (Fig. 4.7d, left). Thus, IsdH uses two energetically distinct interfaces to bind Hb. A high affinity interface formed between the IsdH N2 domain and Hb’s A-helix tethers the receptor to Hb, enabling a second, much weaker interface to form in which the F-helix is distorted.
4.3.7 Molecular dynamics simulations reveal the receptor promotes hemin solvation.

Hemin release from Hb presumably involves the competitive displacement of the axial histidine Ne atom with a water molecule, enabling formation of a new water-Fe$^{3+}$ bond. The observed rate enhancements may therefore in part be due to distortion and increased solvation of the histidine-Fe$^{3+}$ linkage upon receptor binding. To investigate solvation effects of IsdH binding, two explicit-solvent 200ns molecular dynamics (MD) simulations were performed on hemin containing methemoglobin systems: one with, and one without IsdH$^{N2N3}$ bound to the $\alpha$ subunit (Fig. 4.8). Over the course of the simulations, the overall complex structures did not exhibit any large-scale structural rearrangements, and the F-helix in the IsdH$^{N2N3}$ bound structure remained distorted (data not shown). Two analyses were performed to assess the effects of receptor binding on hemin solvation. First, the radial distribution functions (RDFs) of the water hydrogen and oxygen atoms were computed relative to the midpoints of the histidine-iron bonds (Fig. 4.8b). The RDF plots were normalized such that a value of 1.0 is equivalent to bulk solvent. It was observed that in both systems water molecules were allowed to approach this bond, however, in the IsdH$^{N2N3}$ bound system, the waters were more frequently observed at locations close to the bond. For example, in the IsdH$^{N2N3}$ bound system water hydrogens were ~17 times more likely to be located within a distance of 2.1 Å, with the integral of the RDF from 0 to 2.1 Å being 0.016 and 0.001 for the bound and free system respectively, and hydrogens were ~2 times more likely to be located within 2.5 Å of the bond center.

To further investigate the location and stability of solvent molecules surrounding the iron-nitrogen bond, we performed a grid inhomogeneous solvation theory (GiST) analysis of our simulations in the region surrounding the receptor bound hemin. This analysis computes the density, energies, and entropies of water molecules in a three-dimensional grid for the specified
region over the equilibrated portion of a simulation. We defined a “hydration site” as a position in space in which the occupancies of water oxygen molecules were at least three times greater than water in bulk, and which the estimated free energies of water molecules were less than zero relative to bulk. This definition effectively chooses locations where it is favorable for a water molecule to occupy relative to being in the bulk, which is accounted for by both a thermodynamic analysis, as well as an observed high occupancy. Results of this analysis show that in both the IsdH\textsuperscript{N2N3} free and bound forms of Hb, hydration sites exist within the heme binding pocket. However, when IsdH\textsuperscript{N2N3} binds to Hb more of these hydration sites occur near the axial His87 residue in Hb (compare Figs. 4.8c and e). Combined with the RDF analysis, these data indicate that the hemin pocket in Hb is never completely devoid of water molecules, and show that when IsdH binds to Hb, closer and more stable hydration occurs near atoms that are proximal to the iron-nitrogen covalent bond. We conclude that receptor introduced distortions in the axial bond that weaken its affinity combined with increased solvation, favors water in the competitive displacement equilibrium.

4.4 Discussion

*Staphylococcus aureus* acquires iron from human hemoglobin using two related surface receptors, IsdH and IsdB (46). The receptors extract Hb’s hemin using a conserved tridomain unit that is formed by two NEAT domains that are separated by a helix linker domain (Fig. 4.1a) (15). The mechanism of extraction is best understood for IsdH. Previously, we’ve shown that the NEAT domains within its tridomain unit (IsdH\textsuperscript{N2N3}) synergistically extract hemin from Hb at a rate that is significantly faster than the rate at which Hb spontaneously releases hemin into the solvent. A 4.2 Å resolution crystal structure of the Hb:IsdH\textsuperscript{N2N3} complex revealed that each receptor engages a single globin
chain, with the N2 domain contacting the globin’s A-helix and the N3 domain positioned near its hemin molecule (30). The hemin needs to slide only ~12 Å from the globin into the receptor’s N3 domain where its iron atom is coordinated by a tyrosine residue. A subsequent higher resolution 2.55 Å structure of the complex contained the receptor bound to the α-globin and revealed it may trigger hemin release by distorting the F-helix that contains the axial histidine residue (HisF8) (31). While atomic structures have provided great insight into the process of hemin extraction, the energetics underlying this process are poorly understood. This is because the native IsdH-Hb system is heterogeneous, as Hb exists in dimeric and tetrameric forms that have different propensities to release hemin, and once hemin is removed from Hb it dissociates into its component globins that aggregate and more rapidly release hemin. Furthermore, our stopped-flow results indicate IsdH<sup>N2N3</sup> extracts hemin from both the α- and β-subunits within Hb (Figs. 4.3 and 4.4) and that the native IsdH binds to tetrameric Hb with sub-saturating stoichiometry (Table 4.1). Thus, the heterogeneous properties of the native system limit its detailed biophysical characterization.

In order to quantitatively analyze hemin extraction from Hb we developed an assay that preferentially monitors hemin transfer from only its α-subunit. This was accomplished by using a stabilized tetramer of Hb (Hb0.1) and a variant of IsdH<sup>N2N3</sup> (α-IsdH) that we show only binds to the α-subunit (Table 4.1). Stopped-flow UV-Vis experiments reveal that α-IsdH captures hemin in a biphasic manner; a rapid phase characterized by <i>k</i><sub>fast</sub> describes the active extraction of hemin from the α-globin subunit, while slower spectral changes defined by <i>k</i><sub>slow</sub> report on the indirect capture of hemin from the β-subunits after it has first been released into the solvent. The latter conclusion is supported by our finding that the <i>k</i><sub>slow</sub> values are independent of α-IsdH concentration (Fig. 4.5b).
It is possible that some rapid, activated hemin transfer occurs from β-subunits and contributes to the fast phase in the IsdH<sup>N2N3</sup> experiments because our spectral and kinetic measurements do not distinguish between these subunits. The notion of accelerated heme dissociation from Hb β-subunits is suggested by the results of AUC and hemin transfer experiments both of which indicate that IsdH<sup>N2N3</sup> binds to isolated metβ subunits and rapidly acquires its hemin. However, the amplitude of the fast phase is 50% in the transfer experiments with both α-IsdH, where only α chains interact with the receptor, and IsdH<sup>N2N3</sup>, where some binding to β subunits does occur. This result implies that the fast phase represents activated heme transfer from α subunits in both cases. However, it is conceivable that hemin release from the β-subunit also contributes passively to $k_{\text{fast}}$, as our spectral and kinetic measurements can’t distinguish easily between these subunits. Additional experiments using valence hybrid mutants of Hb are required to fully resolve this issue (39).

Our results suggest that the receptor removes hemin at a faster rate from dimeric MetHb (αβ) than it does from tetrameric MetHb (α<sub>2</sub>β<sub>2</sub>), as α-IsdH removes hemin 10-fold faster from native Hb than it does from Hb0.1 ($k_{\text{fast}}$ 0.34 and 3.27 s<sup>-1</sup>, respectively) (Table 4.2 and Fig. 4.3c). This conclusion is supported by our AUC data that indicate that Hb0.1 adopts a tetrameric state at the concentrations used in our stopped-flow studies (Table 4.1), whereas Hb is a mixture of dimeric and tetrameric forms; at the 5 μM hemoglobin concentrations used in this study, ~65% of Hb is tetrameric and ~35% dimeric ($K_{4,2}$ of 1.5 μM) (35). The larger $k_{\text{fast}}$ for native Hb is presumably caused by hemin removal from its dimeric form, which is known to release hemin more rapidly than the tetramer (35). ESI-MS studies have also shown that hemin removal by the receptor causes native Hb to dissociate into its monomeric globin chains (15). This latter process further explains why α-IsdH removes hemin more rapidly from Hb than from Hb0.1, as the isolated
globins caused by native Hb dissolution release hemin at faster rates than Hb dimers and tetramers (35). Similar effects are observed for IsdH^{N2N3}, which like α-IsdH, removes hemin more rapidly from native Hb than it does from the stabilized Hb0.1 tetramer. Interestingly, when the $k_{\text{fast}}$ values for hemin removal from Hb0.1 are compared, IsdH^{N2N3} removes hemin ~2-fold faster than α-IsdH (Fig. 4.3c). We speculate that this difference is caused by the IsdH^{N2N3} receptor’s unique ability to also extract hemin from the β-subunit of Hb0.1, which has weaker affinity for hemin than the α-subunit. This notion is substantiated by the results of AUC and hemin transfer experiments that indicate that IsdH^{N2N3} binds to Metβ and rapidly acquires its hemin. The measured $k_{\text{fast}}$ for hemin transfer from Hb0.1 to IsdH^{N2N3} is therefore likely an amalgam of the rates of hemin removal from both types of globin chains within hemoglobin, whereas the $k_{\text{fast}}$ of transfer to α-IsdH only reflects hemin removal from the α-globin that binds hemin more tightly. No differences in $k_{\text{fast}}$ are observed between the receptors when Hb is employed as the hemin donor, as tetramer dissociation accelerates hemin release thereby masking subtle globin-specific contributions to $k_{\text{fast}}$ (Fig. 4.3b). These complications highlight the benefits of using Hb0.1 and α-IsdH to selectively monitor hemin removal, as the native receptor captures hemin from both globins in a process that is facilitated by Hb dissociation into its component monomeric globins and dimeric Hb (Table 4.2).

The micro-rate constants and energetics of hemin extraction were estimated using data from the α-specific transfer assay. Extraction can be envisioned as occurring through scheme 1, with the ratio of $k_{-1}/k_1$ describing the equilibrium dissociation constant for formation of the Hb:α-IsdH complex, and $k_{\text{trans}}$ representing the process that moves hemin from Hb into the receptor. Transfer is essentially unidirectional, as competition studies with apoMb indicate that IsdH^{N2N3} binds hemin with significantly higher affinity than Hb (Fig. 4.5a). Fitting of this data yielded values of $85.7 \pm 3.8 \mu M$ and $0.57 \pm 0.02 \text{ s}^{-1}$ for $k_{-1}/k_1$, and $k_{\text{trans}}$, respectively (Fig. 4.5). The value
of \( k_{\text{trans}} \) is ~60% larger than the measured value of \( k_{\text{fast}} \), presumably because Hb is only partially saturated with the receptor at the protein concentrations used in the transfer assay. An Eyring plot of the temperature dependence of \( k_{\text{trans}} \) reveals that hemin transfer is limited by an enthalpic barrier at 37°C, with \( \Delta G^\ddagger = 18.1 \pm 1.1 \text{ kcal/mol} \), and \( \Delta H^\ddagger \) and \( \Delta S^\ddagger \) values of 19.5 \( \pm 1.1 \text{ kcal/mol} \) and \( \Delta S^\ddagger = 4.5 \pm 0.3 \text{ cal/mol-K} \), respectively. The activation energy required for spontaneous hemin release from the \( \alpha \)-subunit of tetrameric Hb is ~24 kcal/mol (47). Thus, \( \alpha \)-IsdH effectively lowers \( \Delta G^\ddagger \) by ~6 kcal/mol, consistent with the ~10,000-fold rate enhancement observed in our hemin transfer experiments as compared to the rate of spontaneous hemin release from Hb.

We hypothesize that breakage of the HisF8-iron is rate limiting in the hemin transfer reaction, and that it occurs through a displacement mechanism in which a water molecule outcompetes HisF8 as a ligand for the Fe(III) atom within hemin (48). The Hb:IsdH\( ^{N2N3} \) crystal structure may resemble the pretransfer complex (scheme 1), since the hemin in this structure remains bound to Hb via a HisF8-iron bond. After bond breakage, the mostly non-polar hemin would then travel ~12 Å through a predominantly hydrophobic pathway to the receptor’s N3 domain to form the post-transfer complex (scheme 1) (31). In the Hb:IsdH\( ^{N2N3} \) crystal structure the receptor unwinds the F-helix, which may lower the energy required to break the axial bond breakage by inducing strain. Receptor induced pocket distortions could also lower the activation energy by increasing the concentration of water molecules near the bond, increasing the probability that they can outcompete HisF8’s Ne atom for hemin’s iron atom. Indeed, our MD simulations reveal that receptor binding to Hb causes more stable and closer hydration to occur near the axial bond, perhaps favoring water in the competitive displacement equilibria (Fig. 4.8). Similar hydration of the space near the Fe-HisF8 bond was clearly observed in the crystal structure of the
Leu89(F8)Gly mutant of Mb, which showed a similar 3,000-fold increase in the rate of hemin dissociation at pH 7 (48).

The MD simulations also suggest that the receptor alters the positioning of the HisF8-iron bond, but its quantitative impact on axial bond strength will require the application of more sophisticated computational methods. Interestingly, altering the conformation of the F-helix may be a general strategy used to modulate Hb’s hemin affinity, as helical fluctuations in globin chains have been shown to govern overall affinity (38, 49–53). Moreover, this strategy is employed by the α-Hb stabilizing protein (AHSP), which promotes autooxidation of α-globin chains by binding to a distal site that causes localized distortions in the F-helix (54–56).

Receptor-hemoglobin binding measurements employing ITC reveal that IsdH\textsuperscript{N2N3} uses two energetically distinct binding interfaces with Hb to form the pretransfer complex (Table 4.3). The first interface is comprised of the N2 domain that binds to the A- and E-helices of Hb (N2-Hb interface), while the second interface contains the linker and N3 domains that interact with F-helix, EF- and FG-corners of Hb (LN3-Hb interface) (Fig. 4.1b). Using receptor truncation mutants we demonstrated that the N2-Hb interface drives formation of the pretransfer complex. This is evident from our finding that the IsdH\textsuperscript{N2} polypeptide binds to Hb with similar affinity as the intact tridomain receptor IsdH\textsuperscript{N2N3}, $K_D = 7.5 \pm 1.1$ and $4.0 \pm 0.6$ µM, respectively (Table 4.3). Conversely, the LN3-Hb interface that contains the distorted F-helix forms with very weak affinity as judged by ITC and NMR ($K_D > 1$ mM). ITC data indicate that unwinding Hb’s F-helix is enthalpically unfavorable, as IsdH\textsuperscript{LN3} binding to Hb causes heat to be absorbed (Fig. 4.7), a comparison of the binding data for α-IsdH and α-IsdH\textsuperscript{N2} reveals less favorable ΔH\textsuperscript{o} changes occur for the intact α-IsdH receptor as compared to α-IsdH\textsuperscript{N2} ($ΔH^o = -6.0 \pm 0.5$ and $-10.3 \pm 0.5$ kcal/mol for α-IsdH and α-IsdH\textsuperscript{N2}, respectively) (Table 4.3). In contrast, formation of the higher affinity
N2-Hb interface causes favorable enthalpic changes and unfavorable changes in binding entropy ($\Delta S^\circ = -11.0 \pm 2$ cal/mol-K). The latter effect is likely caused by ordering of loop 2, which undergoes a disordered-to-ordered transition upon binding Hb’s A helix (30, 42, 45). Thus, despite burying similar solvent exposed surface areas, the N2-Hb interface is the primary determinant for overall binding affinity, whereas the LN3-Hb interface forms with very weak affinity because of the large enthalpic penalty that must be paid to distort Hb’s F-helix. Recently, binding of IsdH to the Hb-Hp complex was studied by surface plasmon resonance (SPR) and it was reported that native IsdH$^{N2N3}$ binds Hb-Hp with a $K_D \sim 120$ nM (57), which is an order of magnitude higher in affinity than our ITC measured affinity of the Hb-IsdH$^{N2N3}$ complex. This difference may result from the distinct methods that were employed. However, it is interesting to note that both the IsdH$^{N1}$ domain and IsdB have been reported to bind the Hb-Hp complex with higher affinity than Hb alone (58). As these receptors do not contact Hp directly, it is possible that subtle Hp-induced changes in the structure of Hb may increase its binding affinity for IsdH and IsdB. However, remarkably Hp binding prevents hemin transfer from Hb to IsdB in the ternary complex (57). Thus, the Hp induced conformational changes prevent hemin dissociation with and without bound receptor (59).

Our measured binding energetics for $\alpha$-IsdH are generally similar to those previously reported for IsdB$^{N1N2}$, which is the analogous tridomain unit within IsdB (33). Nevertheless, there are clear differences. Most notably, IsdB$^{N1N2}$ was reported to bind HbCO with 10-fold higher affinity than IsdH$^{N2N3}$, and unlike IsdH$^{N2}$ that binds Hb with high affinity, the analogous isolated domain from IsdB (IsdB$^{N1}$) binds HbCO with weak affinity. Thus, despite a high degree of sequence homology shared by the two receptors, there may be significant biochemical differences whose molecular basis needs to be deciphered.
Our studies enable a free energy reaction coordinate diagram to be constructed that describes the process of receptor-mediated hemin extraction from the α-subunit of Hb (Fig. 4.9). Prior to engaging Hb, the IsdH^{N2N3} receptor undergoes interdomain motions in which the position of the N2 domain moves with respect to the linker and N3 domains that form a rigid unit (29). In step #1, the higher affinity N2-Hb interface forms ($\Delta G^o = -7.0 \pm 0.1$ kcal/mol), contributing ~95% of the total binding standard free energy for the receptor-Hb complex. In step #2, the much weaker Hb-LN3 interface forms, unravelling the F-helix to form the pretransfer complex in a process that is enthalpically unfavorable. It is unclear if Hb binding quenches N2-LN3 inter-domain motions within the receptor. In principle, the two distinct receptor-Hb binding interfaces could form in a concerted manner, with binding at the higher affinity N2-Hb interface nucleating new receptor inter-domain and Hb-receptor interactions that immobilize the receptor on Hb and drive formation of the lower affinity Hb-LN3 interface. However, in the crystal structure of the Hb-IsdH^{N2N3} complex no significant N2-Linker interdomain interactions are observed and electron density for residues Asn465-Glu472 connecting these domains is poorly resolved. Thus, it is possible that inter-domain motions at the N2-Hb interface persist after Hb-IsdH^{N2N3} complex formation, with the N2-Hb interface holding the mobile LN3 unit near Hb’s hemin to increase its effective molarity and the likelihood of forming weak contacts to Hb in which the F-helix is unwound (60). Persistent receptor inter-domain motions within the Hb-IsdH^{N2N3} complex may also facilitate downstream hemin transfer from IsdH to IsdA, since in principle, they would not be reliant upon the complete dissociation of the complex. In step #3, Hb’s hemin molecule is transferred to IsdH following 12 Å path that is exposed to the solvent. Transfer requires a Gibbs activation energy of $18.1 \pm 1.1$ kcal/mol to be surmounted. We hypothesize that this energy is needed to hydrolytically cleave the proximal histidine-hemin bond (Nε-Fe) bond, as $\Delta H^\ddagger$ is $19.5 \pm 1.1$ kcal/mol. The structure of the
complex and our MD simulations suggest that the receptor lowers the activation energy by ~6 kcal/mol by inducing strain in the bond and by increasing the concentration of nearby water molecules that compete with HisF8’s Nε atom for the iron atom within hemin. Based on studies of spontaneous hemin release from Mb and Hb, bond breakage may be facilitated by formation of a hemichrome (61). Indeed, a bis-His hemin complex has recently been observed in the structure of the IsdB-Hb complex (62). The receptor likely extracts hemin from the β-subunit using a similar mechanism, but the energetics of this process remain ill-defined. Once hemin is transferred, the hemin-bound receptor dissociates from Hb, presumably through the two-step process in which the weaker Hb-LN3 interface disengages first (step #4), followed by dissociation of the stronger Hb-N2 interface (step #5). The structure of the post-hemin transfer complex is not known, so it is possible that hemin transfer to the N3 domain alters its structure further weakening the Hb-LN3 interface. Transfer is effectively unidirectional as IsdH binds hemin with significantly higher affinity than Hb based on competition studies with apo-Mb.

Hemin removal by the native receptor from wild-type Hb is expected to be significantly more complex and could involve simultaneous removal from both the α- and β-subunits, as well as receptor-induced dissociation of Hb into dimeric and monomeric species. Based on unfolding studies with purified Hbs, receptor mediated hemin removal from wild-type Hb could cause the formation of a dimeric molten globular intermediate in which the α1β1 interface remains intact, but the heme pocket has partially melted, facilitating the hemin removal from both subunits (63). However, in our structure of the complex only the α F-helix has unfolded. Although we now have an idea about the mechanism and energies that underlie hemin transfer from the α-subunit, it remains unknown how hemin moves between the proteins, how interdomain motions affect this
process, and whether there are true mechanistic differences between IsdH and IsdB. The new assay we describe in this paper could be a valuable tool to help answer these questions.

4.5 Experimental Procedures

4.5.1 Cloning and protein preparation.

IsdH proteins were expressed from pET-28b-based plasmids and initially contained a removable N-terminal hexahistidine-small-ubiquitin-like modifier (SUMO) tag facilitate purification (pHis-SUMO) (64, 65). Protein expressing plasmids, include: pRM208 encoding A326-D660 in IsdH (IsdH\textsuperscript{N2N3}), pRM216 encoding A326-D660 that contains a Y642A substitution (IsdH\textsuperscript{N2N3}-Y642A), pMMS322 encoding A326-D660 with substitution of FYHYA (365-369) \rightarrow YYHYF (365-369) that confers selectivity for the α-globin chain of Hb (α-IsdH\textsuperscript{N2N3}) and pMMS323 containing the same substitution of FYHYA (365-369) \rightarrow YYHYF (365-369) as well as a Y642A substitution (α-IsdH\textsuperscript{N2N3}-Y642A) (15, 31, 66). In addition, the pET-28b-based plasmid pSUMO-Mb was constructed that expresses recombinant sperm whale myoglobin H64Y/V68F with a removable SUMO tag (Mb\textsuperscript{H64Y/V68F}). The expression plasmids were constructed using standard approaches and made use of the QuickChange site-directed mutagenesis kit, (Stratagene, La Jolla, CA). The proteins were expressed by transforming the aforementioned plasmids into E. coli BL21-DE3 cells (New England Biolabs, Beverly, MA). Protein expression and purification procedures have been described previously (15). Briefly, expression proceeded overnight at 25°C by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG) to cell cultures. The bacterial cells were then harvested by centrifugation, lysed by sonication, and the protein purified using a Co\textsuperscript{2+}-chelating column (Thermo Scientific, Waltham, MA). The amino terminal 6x-His-SUMO tag was then cleaved using the ULP1 protease and mixture reapplied to the Co\textsuperscript{2+}-chelating column to
remove the protease and cleaved SUMO tag. The hemin-free forms for the IsdH and Mb proteins were generated by hemin extraction with methyl ethyl ketone followed by buffer exchange into 20 mM NaH$_2$PO$_4$, 150 mM NaCl, pH 7.5 (67). Hemin saturated proteins were produced by adding a 1.5 molar excess of hemin to purified protein solutions followed by removal of excess hemin using a Sephadex G-25 column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 8.0. Protein-hemin stoichiometries were determined using UV-visible spectroscopy and calculating the $A_{\text{Soret}}/A_{280}$.

Human hemoglobin (Hb) was prepared as previously described from the blood of a healthy donor provided by the CFAR Virology Core Lab at the UCLA AIDS Institute (68). Briefly, red blood cells were washed in 0.9% NaCl solution and then lysed under hypotonic conditions. Hb was purified from the hemolysate by two ion-exchange chromatography steps, cation-exchange (SP Sepharose, GE Healthcare Life Sciences) followed by anion-exchange chromatography (Q Sepharose, GE Healthcare Life Sciences). During the purification, the globin chains were maintained in the carbon monoxide-liganded state to inhibit autooxidation and subsequent denaturation. Recombinant genetically stabilized human hemoglobin was expressed from plasmid pSGE1.1-E4 in E. coli BL21-DE3 cells, and contains V1M mutations in both $\alpha$- and $\beta$-globin chains and a glycine linker between the C-terminus of one $\alpha$-subunit and the N-terminus of another $\alpha$-subunit (called Hb0.1) (40, 69). Cells were grown in M9 minimal media containing 3 g/L glucose and 1 g/L NH$_4$Cl as the sole source of carbon and nitrogen, respectively. Prior to induction, exogenous hemin was added to the cell cultures to a final concentration of 8 $\mu$M. Expression was performed at 16°C overnight by adding IPTG to cell cultures to achieve a final concentration of 1 mM. Cells were then harvested by centrifugation, and washed with lysis buffer (50 mM Tris-HCl, 17 mM NaCl, pH 8.5) twice to remove excess hemin. The cell suspension was then purged with a
steady stream of CO for 15 min in an ice-water bath. All buffers used during the purification were kept at pH 8.5 and were purged with CO to prevent heme oxidation. The cells containing recombinant Hb0.1 were lysed by sonication and purified in a single step using a Co²⁺-chelating column (Thermo Scientific, Waltham, MA) via interactions with the naturally occurring His residues on the surface of Hb0.1. The column was washed using three buffers prior to elution with imidazole: 5 column volumes (CV) of lysis buffer (vida supra), 5 CVs wash buffer #1 (20 mM Tris-HCl, 500 mM NaCl, pH 8.5), and 5 CVs wash buffer #2 (20 mM Tris-HCl, pH 8.5). Column-bound Hb was then eluted with 5 CVs of elution buffer (20 mM Tris-HCl, 100 mM imidazole, pH 8.5).

4.5.2 Hemin transfer experiments.

Ferric Hb and Hb0.1 (methemoglobin) were used in the hemin transfer experiments. They were produced by converting the carbonmonoxy-form of each pure protein into the ferric form using established methods and described for Hb. Briefly, HbCO was converted into HbO₂ by delivering a pure stream of oxygen gas over HbCO solution kept cold in an ice-water bath and simultaneously illuminated with a high-intensity (100 W) light source. Removal of CO from HbCO and conversion to HbO₂ was considered complete when the ratio of $A_{577}/A_{560}$ was ~1.8 (70). Oxidation of HbO₂ to generate ferric Hb was achieved by incubating HbO₂ with a five-fold molar excess of potassium ferricyanide followed by application of the sample to a Sephadex G-25 column (GE Healthcare) to remove the excess ferricyanide. Ferric Hb (or Hb0.1) was then buffer exchanged into 20 mM NaH₂PO₄, 150 mM NaCl, pH 7.5. The concentration of hemin within each form of the protein was determined using the extinction coefficient of 179 mM⁻¹ cm⁻¹ at a wavelength of 405 nm (71).
The kinetics of hemin transfer from ferric Hb (donor) to apo-IsdH proteins (acceptor) were measured using an Applied Photophysics SX18.MV stopped-flow spectrophotometer (Applied Photophysics, Surrey, UK). Acceptor and donor proteins were dissolved in 20 mM NaH$_2$PO$_4$, 150 mM NaCl, pH 7.5 supplemented with 0.45 M sucrose to prevent absorbance changes caused by apoprotein aggregation (39). Ferric Hb was mixed with apo-acceptor present at ≥ 10 molar excess (in hemin units) in order to maintain pseudo-first-order conditions. After rapid mixing (dead-time 3 ms) absorbance changes at 405 nm were monitored for 1000 s. Experiments were performed in triplicate and the data analyzed by fitting the observed time courses to a double-phase exponential expression using the GraphPad Prism program (GraphPad Software, version 5.01). As described in the text, micro-rate constants describing the hemin transfer were obtained by fitting the dependence of the observed fast phase kinetic rate constant ($k_{\text{fast}}$) on the concentration of apo-receptor (Eq. 1) using Mathematica software (Wolfram Mathematica 9.0, Wolfram Research). For slower transfer reactions involving apo-Mb$^{H64Y/V68F}$ (39) a 10-fold molar excess of apo-Mb$^{H64Y/V68F}$ was used and absorbance changes were measured using a conventional UV–vis spectrophotometer. For these reactions the entire UV-Vis spectrum was recorded over 17 h period.

4.5.3 Analytical Ultracentrifugation (AUC)

Sedimentation equilibrium runs were performed at 25 °C in a Beckman Optima XL-A analytical ultracentrifuge using absorption optics at 412 nm such that only the heme molecule was detected. The protein concentrations in samples analyzed by AUC were the same as those used for the stopped-flow hemin transfer experiments (20 mM sodium phosphate, 150 mM NaCl, 0.45 M sucrose, pH 7.5). In order to limit hemin transfer during the AUC experiment, IsdH proteins contained Y642A mutation that limits heme binding and the carbon monoxide-liganded form of
Hb0.1 (or β-Hb) was used to limit heme release. Three mm pathlength double sector cells were used for all samples and were purged with CO before sealing the cells to prevent oxidation. Sedimentation equilibrium profiles were measured at speeds of 9,000 (~6,000 × g) and 13,000 rpm (~13,000 × g). Samples contained Hb0.1 or β-Hb (isolated β-globin chains) at a concentration of 5 μM and 0, 75 μM, and 150 μM α-IsdH\textsuperscript{Y642A}. Experiments using native IsdH\textsuperscript{Y642A} made use of a 150 μM sample. Weight-average molecular masses were determined by fitting with a non-linear least-squares exponential fit for a single ideal species using the Beckman Origin-based software (version 3.01). Partial specific volumes were calculated from the amino acid compositions and corrected to 25°C (72, 73). They were 0.749 for the αβ Hb heterodimer and β-globin chains, 0.735 for apo-receptor, and 0.739 for the complex. At higher concentrations of receptor, the fitted molecular weights were slightly lower than those obtained when lower receptor concentrations were used, which is compatible with the increased non-ideality of the solution.

4.5.4 Isothermal Titration Calorimetry (ITC).

ITC measurements were performed using a MicroCal iTC200 calorimeter (GE Healthcare) at 25°C. To prevent heat changes due to hemin transfer the Hb was maintained in the carbonmonoxy ligated state and hemin binding deficient receptor proteins containing an alanine substitution for Y642. The HbCO and the apo-receptor proteins were buffer matched in ITC buffer (20 mM NaH\textsubscript{2}PO\textsubscript{4}, 150 mM NaCl, 0.45M sucrose, pH 7.5). The cell was filled with 30 μM HbCO and the syringe was filled with 200, 300, and 850 μM α-IsdH\textsuperscript{Y642A}, α-IsdH\textsuperscript{N2}, IsdH\textsuperscript{LN3(Y642A)}, respectively. Twenty injections were performed using 2.0 μL injection volumes at 180 s intervals. ORIGIN software was used to fit the data to a single-site binding model. The heat (Q) is related to the number of sites (n), the binding constant (K), and the enthalpy (ΔH) via the following equation,
\[ Q = \frac{nM_t \Delta H V_o}{2} \left[ 1 + \frac{X_t}{nM_t} + \frac{1}{nK M_t} - \sqrt{(1 + \frac{X_t}{nM_t} + \frac{1}{nK M_t})^2 - 4\frac{X_t}{nM_t}} \right] \]

where \( X_t, M_t, \) and \( V_o \) represent the bulk concentration of ligand, the bulk concentration of macromolecule in \( V_o \), and the active cell volume respectively. A correction is made to account for the displaced volume after the \( i^{th} \) injection,

\[ \Delta Q(i) = Q(i) + \frac{dV_i}{V_o} \left[ \frac{Q(i) + Q(i - 1)}{2} \right] - Q(i - 1) \]

Calculated \( \Delta Q(i) \) values for each injection are compared to the corresponding experimental values. The fit is improved over successive iterations by altering values for \( n, K, \) and \( \Delta H \) (74).

### 4.5.5 Molecular dynamic simulations.

The protein structure for the IsdH-bound system was taken directly from a methemoglobin dimer-IsdH complex (PDB 4XS0, resolution 2.55 Å (31)), while the coordinates of native methemoglobin was taken from a tetramer deoxy-hemoglobin structure (PDB 1A3N, resolution 1.8 Å (75)). Missing residues in each structure were added using ModLoop (76). Both systems were neutralized with \( \text{Na}^+ \) ions and solvated with TIP3P (77) waters and a 150 mM NaCl environment in a periodic box, which had a minimum of 10 Å from protein heavy atoms to the box edge. Following minimization for 4,500 steps, both systems were heated to 300K with restraints of 10 kcal/mol/Å\(^2\) on the solute heavy atoms applied for 20 ps, and the restraints were then gradually removed over 350 ps. Production simulations of 200 ns were then performed in the NPT ensemble with temperature maintained via a Langevin thermostat and pressure regulated with a Berendsen barostat (78). Nonbonded interactions were truncated at 10 Å, and long ranged electrostatics were computed with the particle-mesh Ewald method using a maximum spacing of 1.0 Å. All simulations were performed with the GPU accelerated version of PMEMD in the
Amber16 package (79). The AMBERff14SB force field was used for the proteins, while parameters for hemin and the $N_{e}$-$Fe^{3+}$ bond were calculated following the protocol of Shahrokh et al (80). In short, geometry optimization, frequency calculations, and electrostatic potential calculations were done in Gaussian09 (81), and a RESP fit was performed in antechamber Gaussian and tleap scripts were generated using MCBP.py from AmberTools16. Bonded and van der Waals parameters were derived from the General Amber Force Field (GAFF)(82). Analysis of each system was performed on the final 150 ns of simulations using VMD (83), MDAnalysis (84), and CPTRAJ (85). The GIST analysis was done with a grid spanning 33x41x61 Å and a grid spacing of 0.5 Å.
4.6 Figures

Figure 4.1  *S. aureus* uses conserved tridomain containing receptors to capture Hb.

(a) Schematic showing the domain organization within the *S. aureus* Hb receptors, IsdH and IsdB. NEAT domains (N) that bind Hb and hemin (oxidized form of heme) are shown in gray and black, respectively. The helical domain that connects them is labeled “linker”. Residue numbers that define the boundaries of the functionally homologous NEAT domains are indicated. (b) Crystal structure of the Hb:α-IsdH$^{N2N3}$ complex (pdb 4XS0) with only contacts to
the α-subunit shown. Proteins are shown in ribbon format, while the hemin group is represented by a space filling model. Density for residues in the segment connecting the N2 and linker domains are absent in the electron density (dashed line).  (c) Close up view of the LN3-Hb interface that is distorted. α-IsdH is shown in blue and ribbon diagram of αHb is shown in yellow (α-IsdH, residues A326-D660 from IsdH containing 365FYHYA369 \rightarrow 365YYHYF369 mutations). The F-helix in the receptor-Hb complex that is distorted (colored red) is overlaid with the native F-helix observed in the isolated Hb protein (colored green) (PDB code: 2DN2). The hemin group is represented by a space filling model (grey) with the iron atom colored red.
Figure 4.2   Stoichiometry of Hb:receptor complexes determined by analytical ultracentrifugation.

(a) Representative absorbance scans at 412 nm at equilibrium are plotted versus the distance from the axis of rotation for a mixture of Hb0.1 + IsdH$^{Y642A}$ (blue circles) and Hb0.1 + α-IsdH$^{Y642A}$ (red triangles). Hb0.1 was maintained in the carbonmonoxy ligated state, and both receptors contained Y642A mutation to prevent heme transfer. Protein samples were mixed at a ratio of 5 µM Hb0.1 and 150 µM receptor and were centrifuged at 25 °C for at least 48 h at 13,000 rpm. The solid lines represent the global nonlinear least squares best-fit of all the data to a single molecular species with a baseline fit. Residuals of the fit are shown above each panel.

(b) Representative scans identical to the conditions described in (a) with the exception that 5 µM β-CO (isolated β-chains) was used and samples were centrifuged at a speed of 9,000 rpm.
Figure 4.3  Hemin transfer by various Hb and receptors species.

(a) Spectral changes in the UV–Vis spectrum of the reaction containing 5 μM Hb0.1 with buffer (solid black line), 150 μM IsdH^{N2N3} (dotted blue line), and 150 μM α-IsdH (solid red line). Spectral traces were recorded at 1000 s post mixing. (b) Representative stopped-flow time courses showing absorbance changes at 405 nm (ΔA_{405}) after mixing 5 μM Hb or Hb0.1 with 150 μM apo-receptor. The data were fit to a double-exponential equation to obtain k_{fast} and k_{slow} hemin transfer rates. Spectral traces were normalized for comparison using the following equation: y = (y_{i} - y_{0}) / (y_{i} + y_{0}). Panels (c) and (d) show the measured k_{fast} and k_{slow} rate constants, respectively. Values were derived from the time courses in (b). A two-way ANOVA was used to access the significance in the difference of rates. NS and *** corresponds to p-values > 0.05 and < 0.001, respectively.
Figure 4.4  Hemin transfer from isolated β-globin chains.

Representative stopped-flow time courses showing absorbance changes at 408 nm (ΔA_{408}) after mixing 5 μM isolated Metβ-globin chains with 150 μM apo-receptor or 50 μM apo-Mb. Relative spectral traces are shown for comparison in which all starting absorbance were normalized to a value of 1. Only the first 10 s are shown as absorbance artifacts were observed after this time as a result of the slow process of apo-globin denaturation.
Figure 4.5  Hemin transfer to myoglobin and as function of receptor concentration.

(a) Spontaneous hemin release from the receptor and Hb was measured using H64Y/V68F apomyoglobin. Time courses of the absorbance change at 600 nm ($\Delta A_{600}$) of a mixture of H64Y/V68F apomyoglobin (Mb) with Hb (black line), Hb0.1 (dark grey line), and IsdH$^{N2N3}$ (light grey line) at a final concentration of 50 μM apo-Mb to 5 μM holo-protein. The receptor has higher affinity for hemin as compared to Hb or Hb0.1. (b) Plots of the observed rate constants, $k_{fast}$ (dark gray circles) and $k_{slow}$ (light gray circles), versus the ratio of [$\alpha$-IsdH] to [Hb0.1]. The concentration of Hb0.1 was held constant at 5 μM. The observed values were obtained in experiments similar to those described in the legend in Fig. 3. For $k_{fast}$, fits to the data using Eqn. 1 are shown.
Figure 4.6  Temperature dependence of hemin transfer.

(a) Representative normalized stopped-flow time courses as function of ΔA_{405}. They were obtained by mixing 5 μM Hb0.1 with 200 μM α-IsdH. Spectral traces were normalized using the following equation: \( y = \frac{(y_t - y_0)}{(y_t + y_0)} \). (b) A plot of ln(\(k_{\text{fast}}/T\)) versus 1/T is shown. The observed \(k_{\text{fast}}\) values were obtained in experiments similar to those described in the legend in Fig. 3. The correlation factor, \(R^2\), is indicated on the plot.
Figure 4.7  IsdH used two distinct interfaces to promote hemin transfer.

(a) Representative ITC data for the titration of 200 μM α-IsdH<sup>Y642A</sup> into 30 μM HbCO (heme/globin-chain basis). Injections were made at 180 s intervals at 25°C. (b) Representative ITC data for the titration of 300 μM α-IsdH<sup>N2</sup> into 30 μM HbCO. (c) Representative ITC data for the titration of 850 μM α-IsdH<sup>LN3-Y642A</sup> into 30 μM HbCO. The top panels show the time course of the titration (black) and baseline (red). The bottom panels show the integrated isotherms. ORIGIN software was used to fit the data to a single-site binding model to derive thermodynamic parameters. Receptors containing the N3 domain have a Tyr642Ala mutation that prevents hemin binding. (d) NMR titration data showing the effects of Hb binding on 15N-labeled polypeptides that contain only the N3 domain (IsdH<sup>LN3-Y642A</sup>, left) or the linker and N3 domains (IsdH<sup>LN3-Y642A</sup>, right). The receptors contain a Tyr642Ala mutation that prevents hemin binding. 1H-15N HSQC spectra of the proteins in the absence (top) or presence (bottom) of a 10-fold excess of HbCO tetramer are shown. Significant peak broadening is only observed when Hb is added to IsdH<sup>LN3-Y642A</sup>, indicating that linker and N3 domains bind Hb, whereas the N3 domain in isolation does not bind to Hb.
Figure 4.8  Results of molecular dynamics simulations.

(a) The hemoglobin dimer (gold) bound to hemin molecules (grey) and an IsdH<sup>N2N3</sup> molecule (green and blue) was simulated, along with a system lacking IsdH<sup>N2N3</sup> (not shown). (b) Radial distribution functions of solvent hydrogen (solid lines) and oxygen (dotted lines) atoms show that in the receptor bound system, solvent molecules are more frequently found at closer locations to the Nε-Fe bond than in the receptor free system. Analysis of simulations by a Grid Inhomogeneous Solvation Theory (GIST) analysis show that in IsdH<sup>N2N3</sup> free systems (c&amp;d) the number of hydration sites (cyan) surrounding the hemin group is limited, however, in the IsdH<sup>N2N3</sup> bound systems (e&amp;f) significantly more hydration sites are identified around the Nε-Fe bond.
Figure 4.9 Model of the hemin extraction mechanism.

Proposed hemin transfer reaction coordinate diagram. The IsdH^{N2N3} receptor recognizes Hb via two energetically distinct interfaces (N2-Hb and LN3-Hb interfaces). Top, schematic showing the steps in hemin transfer. Bottom, relative free energies of the intermediates. Free energies calculated from this study are labeled along reaction coordinate diagram. Breakage of the axial Fe-Ne bond is presumably rate limiting (step #3). The following color scheme was used: N2, green; Linker-N3, blue; Hb α-subunit, gold.
### Table 4.1  Stoichiometry of Hb:receptor complexes determined by analytical ultracentrifugation

<table>
<thead>
<tr>
<th></th>
<th>Weighted average molecular masses (Da)(^a,b)</th>
<th>Receptor/globin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hb0.1(^c)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13,000 × g</td>
<td>64,212 ± 829</td>
<td></td>
</tr>
<tr>
<td><strong>Hb0.1 + α-IsdH(^c)(^y42A)(^c)(^d)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>140,070 ± 2096</td>
<td>2.0</td>
</tr>
<tr>
<td>13,000 × g</td>
<td>131,570 ± 894</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Hb0.1 + α-IsdH(^c)(^y42A)(^c)(^e)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>134,660 ± 1,872</td>
<td>1.8</td>
</tr>
<tr>
<td>13,000 × g</td>
<td>128,490 ± 1,115</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Hb0.1 + IsdH(^c)(^y42A)(^c)(^e)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>188,110 ± 1,518</td>
<td>3.2</td>
</tr>
<tr>
<td>13,000 × g</td>
<td>175,120 ± 1,045</td>
<td>2.9</td>
</tr>
<tr>
<td><strong>β-globin(^c)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>56,700 ± 724</td>
<td></td>
</tr>
<tr>
<td>13,000 × g</td>
<td>56,300 ± 508</td>
<td></td>
</tr>
<tr>
<td><strong>β-globin + α-IsdH(^c)(^y42A)(^c)(^e)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>54,500 ± 97</td>
<td>0.0</td>
</tr>
<tr>
<td>13,000 × g</td>
<td>54,500 ± 65</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>β-globin + IsdH(^c)(^y42A)(^c)(^e)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,000 × g</td>
<td>134,000 ± 1,321</td>
<td>2.0</td>
</tr>
<tr>
<td>13,000 × g</td>
<td>116,500 ± 617</td>
<td>1.6</td>
</tr>
</tbody>
</table>

\(^a\) Samples were centrifuged at 25°C in buffer containing 20 mM sodium phosphate, 150 mM NaCl, 450 mM sucrose, pH 7.5

\(^b\) Molecular masses were obtained by non-linear least squares exponential fitting of data using Beckman Origin-based software (Version 3.01)

\(^c\) Globin chains were held at a concentration of 5 μM (hemin units)

\(^d\) Receptor concentrations were in a 15-fold molar excess (75 μM)
Receptor concentrations were in a 30-fold molar excess (150 μM)
Table 4.2  Apparent rate constants of hemin transfer

<table>
<thead>
<tr>
<th>Hemin donor</th>
<th>Hemin acceptor</th>
<th>$k_{fast}$ (s$^{-1}$)$^a$</th>
<th>$k_{slow}$ (s$^{-1}$)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MetHb</td>
<td>α-IsdH$^b$</td>
<td>3.27 ± 0.07</td>
<td>0.038 ± 0.001</td>
</tr>
<tr>
<td></td>
<td>IsdH$^b$</td>
<td>3.12 ± 0.08</td>
<td>0.039 ± 0.001</td>
</tr>
<tr>
<td></td>
<td>apo-Mb$^{c,d}$</td>
<td>0.004$^c$ (0.01)$^f$</td>
<td>0.0002$^c$ (0.003)$^f$</td>
</tr>
<tr>
<td>MetHb0.1</td>
<td>α-IsdH$^b$</td>
<td>0.34 ± 0.01</td>
<td>0.026 ± 0.001</td>
</tr>
<tr>
<td></td>
<td>IsdH$^b$</td>
<td>0.73 ± 0.03</td>
<td>0.041 ± 0.001</td>
</tr>
<tr>
<td></td>
<td>apo-Mb$^{c,d}$</td>
<td>0.0004$^c$</td>
<td>0.00008$^c$</td>
</tr>
</tbody>
</table>

$^a$ $k_{fast}$ and $k_{slow}$ values were obtained by fitting to a double exponential expression.
$^b$ Rates were determined at a 30-fold excess [Receptor] pH 7.5, 25°C.
$^c$ Rates were determined at pH 7.0, 37°C.
$^d$ The data are from Ref. (35).
$^e$ $k_{fast}$ and $k_{slow}$ are from the β- and α-globin chains in dimeric Hb, respectively.
$^f$ $k_{fast}$ and $k_{slow}$ are from the β- and α-globin chains in monomeric Hb, respectively.
$^g$ $k_{fast}$ and $k_{slow}$ are from the β- and α-globin chains in tetrameric Hb, respectively. These values are also the same for tetrameric wild-type Hb.
Table 4.3  Thermodynamic parameters and affinity data for Hb binding

<table>
<thead>
<tr>
<th></th>
<th>( K_D )</th>
<th>( \Delta H^o )</th>
<th>( \Delta S^o )</th>
<th>( \Delta G^o )</th>
<th>( n^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \mu M )</td>
<td>kcal/mol</td>
<td>cal/mol ( \cdot K )</td>
<td>kcal/mol</td>
<td></td>
</tr>
<tr>
<td>Hb + ( \alpha )IsdH(^{Y642A})</td>
<td>4.0 ± 0.6</td>
<td>-6.0 ± 0.5</td>
<td>4.6 ± 1.9</td>
<td>-7.4 ± 0.1</td>
<td>0.59 ± 0.03</td>
</tr>
<tr>
<td>Hb + ( \alpha )IsdH(^{N2})</td>
<td>7.5 ± 1.1</td>
<td>-10.3 ± 0.5</td>
<td>-11.1 ± 2.0</td>
<td>-7.0 ± 0.1</td>
<td>0.62 ± 0.04</td>
</tr>
<tr>
<td>Hb + ( \alpha )IsdH(^{LN3(Y642A)})</td>
<td>≥ mM</td>
<td>Endothermic</td>
<td>positive</td>
<td>ND(^c)</td>
<td>ND</td>
</tr>
</tbody>
</table>

\(^a\)Samples were in buffer containing 20 mM sodium phosphate, 150 mM NaCl, 450 mM sucrose, pH 7.5, 298K

\(^b\)\( n \) refers to the molar ratio protein:Hb.

\(^c\)ND refers to “not determined”.

Errors represent the standard deviation of four replicates.
4.7 References


10.1017/CBO9781107415324.004


Chapter 5

Solution Structure of the PhoP DNA-Binding Domain from

*Mycobacterium tuberculosis*
I was the first author of the published work. I carried out the NMR experiments, analyzed the NMR data, determined the NMR structure of PhoPC, and carried out dynamics experiments. Professor Clubb and I wrote the manuscript.
5.1 Overview

Tuberculosis caused by *Mycobacterium tuberculosis* is a leading cause of death worldwide. The PhoP protein is required for virulence and is part of the PhoPR two-component system that regulates gene expression. The NMR-derived solution structure of the PhoP C-terminal DNA-binding domain is reported. Residues 150 to 246 form a structured domain that contains a winged helix-turn-helix motif. We provide evidence that the transactivation loop postulated to contact RNA polymerase is partially disordered in solution, and that the polypeptide that connects the DNA-binding domain to the regulatory domain is unstructured.
5.2 Introduction

The bacterial pathogen *Mycobacterium tuberculosis* (MTB) causes tuberculosis resulting annually in ~1.5 million fatalities world-wide (WHO 2014). During the initial stage of an infection, MTB is inhaled and transported to the lungs where it is believed to infect alveolar macrophages (Russell et al. 2010). Once inside the macrophage, MTB encounters new oxidative and acidic stressors, and alters its gene expression profile to cope with this new environment (Gonzalo-Asensio et al. 2008). The MTB PhoP-PhoR (PhoPR) two-component system plays a critical role in microbial adaptation. In this system, environmental signals received by the membrane associated PhoR protein trigger its autophosphorylation. PhoR then transfers the phosphate to the cytoplasmic PhoP response regulator, which in its phosphorylated state sequence specifically binds DNA with high affinity (Pathak et al. 2010). Several studies have shown that PhoP is required for MTB virulence (Perez et al. 2001; Walters et al. 2006; Gonzalo-Asensio et al. 2008). In particular, *phoP* mutant strains of MTB exhibit attenuated growth in murine and macrophage models of infection (Perez et al. 2001). Attenuated growth is due largely to the absence of certain complex lipids in the MTB cell envelope that protect the microbe against host defense mechanisms (Walters et al. 2006; Ryndak et al. 2008). Over 110 genes in MTB exhibit altered expression levels in the absence of PhoP (Walters et al. 2006). PhoP regulates genes involved in a variety of cellular functions, including: hypoxia response, respiratory metabolism, stress response, secretion of major T-cell antigen ESAT-6, synthesis of pathogenic lipids, and MTB persistence through transcriptional regulation of the enzyme isocitrate lyase (Gonzalo-Asensio et al. 2008). While anti-microbial therapies against MTB exist, new anti-tubercular therapeutics are needed as current approaches can be problematic to implement as they require the use of a sustained treatment regimen and new drug resistant forms of MTB have emerged (Ryndak et al. 2008).
The PhoP protein contains two autonomously folded domains, an N-terminal receiver domain that is phosphorylated by PhoR (PhoPN; spanning residues 1-138), and a C-terminal effector domain that binds DNA (PhoPC; comprising residues 150-247) (Pathak et al. 2010). It is unclear how phosphorylation of the PhoPN domain causes PhoP to bind DNA with high affinity via its PhoPC domain. Previously, we have shown that the linker connecting the PhoPN and PhoPC domains is required for phosphorylation-dependent DNA binding in vitro, and thus, the transmission of the phosphorylation signal between the domains (Pathak et al. 2010). To gain insight into the mechanism of DNA binding and the structural role of the interdomain linker, here we report the NMR structure and backbone dynamics of the PhoPC DNA-binding domain and several residues from the linker polypeptide that connects it to the receiver domain in the full-length protein.

5.3 Results

The structure of PhoPC is formed by two β-sheets that pack against a three-helix bundle (Fig. 5.2c). To facilitate comparisons with previously reported structures, the secondary structural elements of PhoPC are labeled based on their appearance in the full-length protein (Fig. 5.2c). Beginning at the N-terminus, the polypeptide adopts a four-stranded antiparallel β-sheet that is formed by residues that are contiguous in the primary sequence (strands β6-β9). The chain then forms helix α6, which packs all of the strands in the four-stranded sheet. After this two additional alpha helices (α7 and α8) occur that pack against helix α6 to form the three-helix bundle. The structure is completed by a β-hairpin constructed from strands β11 and β12, which are attached via strand β12 to the body of the protein by contacts to a short β-strand (β10) segment located between helices α6 and α7. Notably, in all the members of the ensemble,
residues at the end of helix α7 form a single turn of a 3_10-helix (residues 203-205). The N-terminal β-sheet forms a hydrophobic interface with α6 as evidenced by NOE’s between V165 (β8) and F179 (α6), L151 (β6) and I187 (α6), and F153 (β6) and V186 (α6). Hydrophobic contacts are also observed between helices α6, α7, and α8. This is supported by the observation of NOEs between the side chains of T177, T180, and L181 located in α6 and V202 in α7, between L199 (α7) and V214 (α8), as well as between the side chains of L181 (α6) and V218 (α8). A small interface joins the C-terminal sheet to the helical bundle which is defined by NOE’s between Y241 (β12) and residues V218 (α8) and I198 (α7), as well as NOE’s between L233 (β11) and the side chains of L221 (α8) and I225 (α8).

Extensive chemical shift and NOE assignments form the foundation for the structure determination of PhoPC. The chemical shifts of 95%, 100% and 50% of the backbone amide (excluding proline residues), side chain methyl, and aromatic side chain atoms were assigned, respectively. For the structured regions of the protein, an average of ~14 NOE-derived distance restraints were identified per residue. These included, ~5 long range (> 4 residues apart), ~3 medium range (2 ≤ residue separation ≤ 4), and ~5 sequential NOE-derived distance restraints per residue. From the backbone chemical shifts the programs TALOS+ and CSI 2.0 predicted the presence of similar secondary structural elements in the protein (Shen et al. 2009; Hafsa and Wishart 2014). However, some of the elements predicted by the programs differ subtly in their amino acid lengths and location. For example, strands β6 and β9 are predicted by both programs to be of similar length, but they differ by one residue in their location in the primary sequence.

Heteronuclear [^1H-^{15}N]-NOE data provides insight into the flexibility of the polypeptide backbone. A total of 71 out of 101 backbone amide resonances are well resolved enabling their[^1H-^{15}N]-NOE values to be reliably determined. Overall, the DNA-binding domain of PhoPC
adopts a rigid structure, as the majority of the residues spanning strands β6 to β12 exhibit [1H-15N]-NOE values greater than 0.6 (Fig. 5.3a). This finding is also compatible with order parameters predicted from the backbone chemical shift data (RCI-S², random coil index order parameter) as the RCI-S² values for residues within the domain generally range between 0.8-0.9 (Berjanskii and Wishart 2005) (Fig. 5.3b). Interestingly, both the [1H-15N]-NOE and RCI-S² data indicate that the loop connecting helices α7 and α8 (α7/α8 loop, residues 206-211) has elevated mobility; [1H-15N]-NOE and RCI-S² values for loop residues D206, F207, G208, and D210 are less than 0.6 (Fig. 5.3a) and 0.8, respectively. It is also compatible with the NMR structure, as the coordinates for residues in the α7/α8 loop are slightly more structurally disordered than the remainder of the domain (Fig. 5.2a). However, the coordinates of the loop are not entirely disordered because several long-range NOE distance restraints are observed between the loop and the remainder of the protein. The loops connecting α8 and β11, and the turn connecting the two strands in the C-terminal β-hairpin also exhibit slightly elevated mobility as indicated by [1H-15N]-NOE values less than 0.6 for K230 and V239 (Fig. 5.3a), along with the observation of slightly decreased RCI-S² values for these regions of the polypeptide (Fig. 5.3b). In addition to the DNA-binding domain, the PhoPC polypeptide studied here contains a portion of the interdomain linker segment that is important for phosphorylation-dependent DNA binding (Pathak et al. 2010). Three lines of evidence indicate that the linker is unstructured in PhoPC. First, the [1H-15N]-NOE values of residues K144 and R147 have negative [1H-15N]-NOE values, and residues N148 and V149 exhibit [1H-15N]-NOE values of only 0.18 ± 0.04 and 0.52 ± 0.03, respectively (Fig. 5.3a). Second, residues in the linker do not exhibit long-range NOE cross-peaks in the NOESY spectra. Finally, the RCI-S² values for residues spanning 142-149 have small magnitudes that progressively decrease towards the N-terminus.
5.4 Discussion

The NMR structure of PhoPC adopts a winged-helix fold similar to members of the OmpR/PhoB subfamily of proteins (Wang et al. 2007). *In vitro* studies have demonstrated that PhoPC binds DNA with similar affinity as the full-length protein, indicating that PhoPC is the primary determinant for DNA binding (Pathak et al. 2010). The helix-turn-helix motif in PhoPC is formed from helices α7 and α8. Helix α8 likely mediates sequence specific binding to the major groove, and binding is presumably further assisted by DNA contacts from the “wing” structure formed by residues in the C-terminal β-hairpin. This binding mode is compatible with the positively charged electrostatic surface formed by residues within the C-terminal β-hairpin and α8 (Fig. 5.2d) (Wang et al. 2007). Most of the remainder of the protein surface is negatively charged, which could function to orient the protein on the DNA duplex (Wang et al. 2007). The notion that α8 interacts with the major groove is also strongly supported by the results of *in vitro* binding studies as mutants containing alterations in the helix are impaired in DNA binding (Das et al. 2010).

In general, the NMR structure is similar to previously reported crystal structures of PhoP. In particular, the backbone coordinates of the NMR structure have a r.m.s.d. of 1.21 and 1.25 Å to the crystal structures of PhoPC (pdb accession: 2PMU) and full-length PhoP (pdb accession: 3R0J), respectively. The largest differences in the structures occur in the loop that connects α7 to α8 (loop 1), which has been postulated to interact with RNA polymerase to activate transcription (Wang et al. 2007) (Fig. 5.2b). As compared with the crystal structure of the full-length protein, the loop in the NMR structure is kinked forward, whereas in the crystal structure it is kinked backwards toward the side of helix α7 (Fig. 5.2b) This difference primarily originates at residues R204, Y205, D206, and G208 in the loop which have distinct φ and ψ angles in the NMR and crystal structures. However, these structural differences are likely to be due to loop motions as
many of its residues have $[^1H-^{15}N]-$NOE values less than 0.6. This notion is consistent with the crystal structure of PhoPC, since interpretable electron density for the loop is available for only four of the six protein molecules within the asymmetric unit. Another less substantial structural difference is the positioning of the C-terminal β-sheet which in the crystal structures is packed slightly closer to helix α8 than in the NMR structure.

Overall, most of the residues in the NMR structure of PhoPC adopt favorable or additionally allowed backbone conformations when assessed using programs PROCHECK and MolProbity (Table 5.1) (Laskowski et al. 1996; Chen et al. 2010). Interestingly, in all the conformers of the ensemble residues A154 and A168 adopt a conformation in which they have a positive φ angle and are located in the disallowed and generously allowed regions, respectively. This is likely the predominant conformation of these residues in solution as they are well defined by the NOE data (Fig. 5.3c) and they also exhibit positive φ angles in both crystal structures of PhoP (Wang et al. 2007; Menon and Wang 2011). Residue T227 also adopts a high-energy conformation that places it in the generously allowed regions of the Ramachandran plot. However, this residue is located in loop 2 and is poorly defined by the NMR data. In comparison to the latest structure reported of PhoP (Menon and Wang 2011), the same secondary structural elements are present in the NMR structure. Strands β6, β7, β9, α8, and β11 appear at identical positions in the primary sequence. However, there are subtle differences as the lengths of elements β8, α6, and β12 are extended by one residue in the crystal structure as compared to the NMR structure. In addition, helix α7 is two residues longer in the NMR structure as compared to the crystal structure. Finally, although strand β10 is the same length in both structures, in the NMR structure it is displaced one residue towards the C-terminus in the primary sequence.

The linker connecting the N- and C-terminal domains is important for phosphorylation-
coupled DNA binding (Pathak et al. 2010). This raised the possibility that it might interact with PhoPC to modulate its structure and ability to bind DNA, and that these interactions could be dependent upon phosphorylation of PhoPN. Our data indicate that the linker is unstructured in the absence of the PhoPN domain, as its residues have low magnitude [\(^{1}H-^{15}N\)]-NOE and RCI-S\(^2\) values, and are not defined by long-range NOEs in the NOESY spectra. This suggests that in solution there is no intrinsic propensity of the linker to interact with the PhoPC domain. This finding is compatible with the crystal structures of PhoP, as residues for the linker exhibited only partial electron density and could not be modeled in either structures of intact PhoP or PhoPC (Wang et al. 2007; Menon and Wang 2011). Studies suggest that the phosphorylation of PhoPN induces PhoP dimerization, consequently increasing DNA binding affinity in a cooperative manner (Sinha et al. 2008; He and Wang 2014). The presence of a long and flexible linker between the PhoPN and PhoPC domains could permit PhoPN dimerization, while still enabling large-scale domain rearrangements of PhoPC that are needed for it to productively engage DNA as a tandemly oriented dimer (Menon and Wang 2011). Insight into the mechanism of phosphorylation-coupled DNA binding by PhoP awaits the structure determination of the PhoP-DNA complex.

5.5 Experimental Procedures

5.5.1 Protein expression, purification, and NMR sample preparation

NMR was used to determine the structure of a polypeptide containing the C-terminal DNA-binding domain of PhoP (PhoPC, residues 142-247 of PhoP, Fig. 5.1a). The protein also contains at its N-terminus 30 amino acids that include a six residue histidine tag and residues that from the linker segment between the N- and C-terminal domains in the full length protein. Uniformly $^{13}C$-
and $^{15}$N-labeled PhoPC was expressed in *Escherichia coli* BL21(DE3) cells grown in M9 medium supplemented with $^{15}$NH$_4$Cl and [1$^{13}$C$_6$] glucose. Cultures were grown at 37º C to an A$_{600}$ of 0.4 before induction with isopropyl β-D-thiogalactoside to a final concentration of 1 mM. Induction proceeded at 18º C overnight before harvesting the cells by centrifugation at 5,400 × g for 20 min at 4º C. The cell pellet was then resuspended in lysis buffer consisting of: 50 mM NaPO$_4$, 500 mM NaCl, 10 mM CaCl$_2$, and 10 mM MgCl$_2$ (pH 8.0). Lysozyme was added to a final concentration of 0.1 mg/mL and the resuspended cells were then kept on ice for 30 min. EDTA and NaCl were then added to the final concentrations of 25 mM and 600 mM, respectively, along with a protease inhibitor cocktail (Calbiochem) and phenylmethanesulfonyl fluoride (Sigma). Cells were lysed by sonication and centrifuged at 11,300 × g for 40 min at 4º C. The supernatant was then incubated with 1.0 mL of pre-equilibrated Ni$^{2+}$ resin (Clontech) for 60 min at 4º C on a rotisserie before being transferred to a gravity column. The resin was then washed with 10 mL lysis buffer containing 1.0 M NaCl, 10 mL lysis buffer containing 15 mM imidazole, and the PhoPC protein was finally eluted by adding lysis buffer containing 300 mM imidazole. Fractions containing PhoPC were pooled and dialyzed against NMR buffer. NMR samples of PhoPC contained 1 mM U-$^{15}$N,$^{13}$C PhoPC dissolved in 50 mM NaPO$_4$, 300 mM NaCl, and 0.01% NaN$_3$ (pH 6.5). Two samples were studied by NMR that either contain 8% or 99.9% v/v of D$_2$O. Because of the high salt concentration in the sample buffer, a “Shaped Sample Tube” (Bruker) was employed to maximize sensitivity.

### 5.5.2 NMR spectroscopy, data collection, and chemical shift assignments

NMR experiments were performed at 298 K on Bruker Avance 500-, 600-, and 800-MHz spectrometers equipped with triple resonance cryogenic probes. NMR spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the programs CARA (Keller 2004) and
PIPP (Garrett et al. 1991). Backbone $^1$H, $^{13}$C, and $^{15}$N chemical shifts were assigned by analyzing data from the following experiments: 2D $[^1$H-$^{15}$N]-HSQC, 3D CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HNCA, HNHA, HBHA(CO)NH, HBHANH, CC(CO)NH, $^{15}$N-edited TOCSY-HSQC (40 ms mixing time), and $^{15}$N-edited NOESY-HSQC (120 ms mixing time) (Fig. 5.1b). Aliphatic side-chain assignments were obtained by analyzing 2D $[^1$H-$^{13}$C]-HSQC (aliphatic), 3D HC(C)H-COSY, HC(C)H-TOCSY, and (H)CCH-TOCSY spectra [For a review, see (Cavanagh et al. 1995)]. Aromatic side-chain assignments were determined using 2D $[^1$H-$^{13}$C]-HSQC (aromatic) and 3D $^{13}$C-edited NOESY-HSQC (aromatic, 150 ms mixing time) datasets. NOE distance restraints were obtained by analyzing 3D $^{15}$N-edited NOESY-HSQC (120 ms mixing time), $^{13}$C-edited NOESY-HSQC (aliphatic, 120 ms mixing time), and $^{13}$C-edited NOESY-HSQC (aromatic, 150 ms mixing time) spectra. $^{13}$C-edited NOESY experiments were optimized for either aliphatic or aromatic signal detection by positioning the $^{13}$C carrier at 40.0 ppm or 80.0 ppm, respectively. Dihedral angle restraints, $\phi$ and $\psi$, were obtained using the program TALOS+ (Shen et al. 2009). $^3$J$_{HN}^a$ couplings were measured using a 3D HNHA spectrum. To identify disordered regions in PhoPC, heteronuclear $[^1$H-$^{15}$N]-NOE values were measured (in triplicate) and analyzed using the program SPARKY (Goddard and Kneller 2006).

5.5.3 NMR structure determination

Simulated annealing with restrained molecular dynamics was used to calculate the structures of PhoPC using the program XPLOR-NIH (Schwieters et al. 2003). The structure was determined in an iterative manner. Initially, NOE cross-peaks in the NOESY spectra were assigned automatically using the programs ATNOS-CANDID (Herrmann et al. 2002b; Herrmann et al. 2002a) and UNIO (Herrmann 2010). To facilitate automatic assignments, the PhoPC crystal
structure (pdb accession: 2PMU) was used as input for UNIO. Hydrogen atoms were added to the crystal structure using the Bax laboratory web-server (http://spin.niddk.nih.gov/bax/nmrserver/pdbutil/sa.html). The first round of UNIO calculations correctly generated the global fold of the protein. The structure was then iteratively refined by adding new distance restraints that were manually identified in the NOESY spectra. In this procedure, additional side chain chemical shift assignments were obtained by inspecting all of the NMR spectra, and the 3D $^{15}\text{N}$-edited NOESY-HSQC and 3D $^{13}\text{C}$-edited NOESY-HSQC spectra were further analyzed to identify NOEs that were not originally assigned by UNIO. Importantly, in this procedure all of the NOEs assigned by UNIO were verified manually. Approximately 10 cycles of iterative refinement were performed. Fig. 5.3c shows a plot of the number of NOE derived distance restraints identified per residue. For the structured regions of the protein, an average of ~14 distance restraints define the conformation of each residue. As expected, fewer NOE distance restraints were obtained for residues located within surface loops and the polypeptide termini.

Towards the end of the refinement process, hydrogen bond restraints were employed that were identified from deuterium exchange data and characteristic NOE patterns present in the NOESY spectra (Wüthrich 1986). The final structures were also refined so as to agree with the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts of the protein, and $^3\text{J}_{\text{HN}^\alpha}$ coupling data. The range for the $\phi$ and $\psi$ dihedral angle restraints was set to two times the error defined by the program TALOS+, or ± 30°, whichever value was larger. A total of 200 structures were calculated, of which 187 were completely compatible with the NMR data; they had no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively. Chemical shift assignments have been deposited in the Biological Magnetic Resonance Bank (BMRB 11590), and the coordinates and

196
structure restraints have been deposited in the protein data bank (PDB 2RV8). The solution structure of PhoPC is represented by an ensemble of the 30 lowest energy conformers generated from this calculation (Fig. 5.2a, pdb accession 2RV8). Residues 150-205 and 211-246 in PhoPC are structured, and the root mean square deviation of their backbone and heavy atom coordinates to the mean structure is $0.48 \pm 0.10$ Å and $1.05 \pm 0.10$ Å, respectively. Table 1 contains the complete structural and restraint statistics.
5.6 Figures

Figure 5.1  NMR spectra of PhoPC.

(a) Schematic of full-length PhoP showing its component domains: N-terminal domain (blue), linker segment (green), and C-terminal domain (red). The bracket at the bottom illustrates the length of our construct. (b) The [$^1$H,$^{15}$N]-HSQC spectrum of PhoPC (142-247) [G190 not shown].
Figure 5.2 NMR structure of PhoPC and structural comparisons.

(a) Superposition of the 30 lowest energy structures with the N- and C-termini labelled. (b) Overlay of the lowest energy structure (blue) and the crystal structure (red) (pdb accession: 3R0J [N-terminal domain not shown]). (c) Lowest overall energy structure with the secondary structure labelled according to the crystal structure. (d) Electrostatic potential of the protein surface calculated in PYMOL (Delano 2002). The positively charged surface potential is shown in blue and the negatively charged surface potential is shown in red. For clarity, residues 149 to
246 are shown in the structures above and were used to calculate the electrostatic potential.

Figure was generated using MOLMOL (Koradi et al. 1996) and PYMOL.
Figure 5.3  NMR relaxation, chemical shift and distance restraint data.

(a) Plot of the heteronuclear $[^1\text{H}-^{15}\text{N}]$-NOE values versus residue number. Errors were calculated by taking the standard deviation of values taken from three separate experiments. (b) Order parameters predicted based on backbone chemical shifts using the RCI server (RCI-S$^2$). Areas shaded in grey represent residues belonging to the linker. (c) Plot showing the number of NOE-derived distance restraints as a function of residue number. Restraints are classified as long range (black, $|i-j| > 4$), medium (blue, $2 \leq |i-j| \leq 4$) and sequential (red, $|i-j| = 1$). Figure was generated using Graph Pad Prism (version 5.01).
Table 5.1  Statistics for the solution structure of PhoPC

<table>
<thead>
<tr>
<th>Root mean square deviations</th>
<th>&lt;SA&gt;^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE interproton distance restraints (Å) (793)^b</td>
<td>0.015 ± 0.003</td>
</tr>
<tr>
<td>Dihedral Angle Restraints (degrees)^c (185)</td>
<td>0.357 ± 0.044</td>
</tr>
<tr>
<td>^3J_{HN}^a coupling constants (Hz) (25)</td>
<td>0.214 ± 0.025</td>
</tr>
<tr>
<td>Secondary ^13C shifts (ppm)</td>
<td></td>
</tr>
<tr>
<td>^13C_α (ppm) (91)</td>
<td>1.070 ± 0.046</td>
</tr>
<tr>
<td>^13C_β (ppm) (91)</td>
<td>0.986 ± 0.046</td>
</tr>
</tbody>
</table>

| Deviations from idealized covalent geometry                    |                 |
| Bonds (Å)                                                     | 0.0014 ± 0.0001 |
| Angles (degrees)                                              | 0.392 ± 0.005   |
| Impropers (degrees)                                           | 0.245 ± 0.006   |

| PROCHECK results^d                                             |                 |
| Most favorable region (%)                                       | 80.4 ± 2.3      |
| Additionally allowed region (%)                                 | 16.6 ± 2.3      |
| Generously allowed region (%)                                   | 1.8 ± 1.0       |
| Disallowed region (%)                                          | 1.2 ± 0.2       |

| Coordinate precision^e                                          |                 |
| Protein backbone (Å)                                           | 0.48 ± 0.10     |
| Protein heavy atoms (Å)                                        | 1.05 ± 0.10     |

^a <SA> represents an ensemble of the 30 lowest energy structures calculated by simulated annealing. The number of terms for each restraint is given in parentheses. None of the structures exhibit distance violations greater than 0.5 Å, dihedral angle violations greater than 5°, or coupling constant violations greater than 2 Hz.

^b Distance restraints: 238 sequential, 125 medium (2 ≤ residue separation ≤ 4) and 211 long range (> 4 residues apart).

^c Experimental dihedral restraints comprised 93 φ and 92 ψ angles.

^d PROCHECK (Laskowski et al. 1996) data include residues 150 – 205 and 211 – 246. MolProbity (Chen et al. 2010) was also used to assess the quality of the structure. For the structured regions of the protein, 96% ± 1% of the residues were in the favored or allowed regions of the Ramachandran plot.

^e The coordinate precision is defined as the average root mean square deviation of the 30 individual simulated annealing structures and their mean coordinates. The reported values are for residues 150 to 205 and 211 to 246. The backbone value refers to the N, C_α, and C’ atoms.
5.7 References


