STRUCTURAL INVESTIGATION OF ISDB, THE HEMOGLOBIN RECEPTOR OF *STAPHYLOCOCCUS AUREUS*

by

Brittany Anne Fonner

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

August 2014

DEDICATION

To my mom and dad, thank you for your love and support. This work is for you.

ACKNOWLEDGEMENTS

I have truly enjoyed my time at Montana State University and would like to acknowledge so many wonderful people who have been with me throughout this incredible journey. First I would like to thank my graduate advisor, Dr. Valérie Copié, for her support throughout this challenging process. I would also like to thank Dr. Brian Tripet for his encouragement, valuable insight and scientific knowledge. He greatly contributed to this project and to my growth as an independent researcher. Furthermore, I would like to thank the members of my committee, Dr. Benfang Lei, Dr. Martin Lawrence, Dr. Brian Bothner, Dr. Joan Broderick, and Dr. Jennifer DuBois, for their knowledge and helpful discussions throughout my time here.

I would like to thank my professors at East Stroudsburg University of Pennsylvania for their support and encouragement in my undergraduate studies and throughout graduate school. They cultivated my love of chemistry and without them I may not have decided to pursue a graduate degree. In particular, I would like to thank my academic advisor, Dr. Michael Doherty, for sticking with me even when I changed my major, my biochemistry professors, Dr. Michelle Jones-Wilson and Dr. John Freeman, who told me how awesome proteins were, although I did not fully appreciate that fact until graduate school, and to my first general chemistry lab professor, Dr. Sharmaine Cady, for demanding excellence in my scientific writing.

And finally I would like to thank my many friends and family, here and back home, for their continued love and support.

TABLE OF CONTENTS

1. INTRODUCTION 1

Staphylococcus aureus 2
Role of Iron in Biology 4
Bacterial Iron Acquisition Systems 6
Gram-negative Bacteria 8
Gram-positive Bacteria 12
Iron-regulated Surface Determinant (Isd) Pathway 15
Regulation and Expression 15
Role of the Isd Proteins 16
Surface Proteins: IsdB, IsdH, IsdA, and IsdC 16
Membrane Transporter Proteins: IsdE, IsdD, and IsdF 18
Heme Degrading Enzymes: IsdG and IsdI 20
Near Iron Transporter (NEAT) Domains 20
Hemoglobin Receptors 23
Scope of Dissertation 23

2. MATERIALS AND METHODS 25

Introduction 25
Material and Methods for the Structure Determination of $IsdB^{N1}$ 26
Cloning, Protein Expression, and Purification 26
Nuclear Magnetic Resonance (NMR) Spectroscopy 28
Structure Calculation and Refinement 33
Hemoglobin-binding Studies of $IsdB^{N1}$ 36
Site-directed Mutagenesis for Hemoglobin-binding Studies 36
Hemoglobin-binding Studies 37
Heme Transfer Studies of IsdB Constructs 37
Preparation of Apo-IsdB Constructs 39
Heme Transfer Reactions 40
Material and Methods for the NMR Investigation of the Linker Domain 40
Cloning, Protein Expression, and Purification 40
NMR Spectroscopy 41
Circular Dichroism Spectroscopy 42
NMR Investigation of the Linker Region with the N1 and N2 Domains of IsdB 42
Material and Methods for $IsdB^{N2}$ and $IsdB^{Linker-N2}$ 42
Material and Methods for $IsdB^{N1-Linker}$ 43
NMR Experiments of IsdB Linker Constructs 43
Material and Methods for the Structure Determination of apo-NEAT 2 43
Cloning, Protein Expression, and Purification 44
NMR Spectroscopy 45

TABLE OF CONTENTS-CONTINUED

Structure Calculation 45

3. STRUCTURAL AND FUNCTIONAL STUDIES OF ISDB^{N1} 47

Introduction 47
Characterization of IsdB^{N1} 48
Backbone and Side Chain Resonance Assignments 49
NMR Solution Structure of IsdB^{N1} 53
Structural Homology 58
Hemoglobin Binding 64
Heme Transfer Reactions 67

4. INVESTIGATION OF THE LINKER REGION OF ISDB 77

Introduction 77
NMR Spectroscopy 77
Characterization of the Linker Region 79
Investigating Different IsdB Linker Constructs 81
Comparison to the Linker Region of the IsdH Homolog 86

5. THE SOLUTION STRUCTURE OF APO-ISDB^{N2} 88

Characterization of NEAT 2 88
Extent of NMR Assignments 90
Structure Determination 91
Comparison to the Published Holo-IsdB^{N2} Crystal Structure 96
Structural Homology 96

6. CONCLUDING REMARKS 99

REFERENCES CITED 107

APPENDICES 118

APPENDIX A: NMR Experiments and Parameters Used for Data Collection for IsdB^{N1} 119
APPENDIX B: Example Input Files for Structure Calculation of IsdB^{N1} in CYANA 121
APPENDIX C: Scripts Needed for Structure Calculations in CYANA 127

LIST OF TABLES

Table Page

2-1. Spectra acquired for backbone and side chain assignments of $IsdB^{N1}$ 32

3-1. NMR structure statistics for $IsdB^{N1}$ 56

3-2. Heme transfer rates from metHb to IsdB constructs at 25°C 73

5-1. NMR structure statistics for apo-$IsdB^{N2}$ M362A/I438A 94

LIST OF FIGURES

Figure Page

1-1. The structure of iron protoporphyrin IX 5

1-2. Comparison of heme import by Gram-negative and Gram-positive bacteria 7

1-3. The structure of the siderophore receptor, FecA 11

1-4. The structure of the hemophore HasA from *Serratia marcescens* 12

1-5. Structure of the peptidoglycan cell wall of *S. aureus* 13

1-6. The Iron-regulated Surface Determinant pathway of *S. aureus* 15

1-7. The Isd operon in *S. aureus*. 16

1-8. Sortase attachment of proteins to the cell wall of *S. aureus*. 17

1-9. The surface proteins of the Isd pathway in *S. aureus*. 19

1-10. A model of the Isd ATP-Binding Cassette transporter of *S. aureus* 21

1-11. The immunoglobulin-like fold of the NEAT domain. 22

2-1. Schematic depicting the domain arrangement of IsdB and IsdB fragments. 25

2-2. Amino acids with characteristic Cα and Cβ chemical shifts. 30

2-3. Example of sequential backbone assignment for $IsdB^{N1}$ using the HNCACB and CBCA(CO)NH experiments 31

2-4. Proton-proton distance restraints based on the NOESY assignment for $IsdB^{N1}$. 33

2-5. Schematic of the steps for the restrained molecular dynamics protocol in AMBER9 36

3-1. A schematic of IsdB and its domain arrangement according to residue number. 48

LIST OF FIGURES-CONTINUED

Figure Page

3-2. Labeled ^{1}H-^{15}N HSQC spectrum of $IsdB^{N1}$ 51

3-3. Secondary structural topology of $IsdB^{N1}$ derived from chemical shift indexing and dihedral angles predictions from TALOS+ analysis of chemical shift data 52

3-4. Ramachandran plot of $IsdB^{N1}$ 55

3-5. Solution structure of $IsdB^{N1}$. 57

3-6. Van der Waal interactions of residues in α1 and β sheet 1.....60

3-7. Greek key motif in $IsdB^{N1}$. 61

3-8. Comparison of the structure of $IsdB^{N1}$ with the solution NMR structure of $IsdH^{N1}$ and the crystal structure of $IsdH^{N2}$. 62

3-9. Comparison of the structure of $IsdB^{N1}$ with the NEAT domain of IsdA.. 63

3-10. Sequence alignment of $IsdB^{N1}$ to the NEAT domains of all Isd proteins from *S. aureus*.. 65

3-11. Hemoglobin binding studies of $IsdB^{N1}$ and $IsdB^{N1}$ protein variants.. 68

3-12. Structural alignment between $IsdB^{N1}$ and $IsdH^{N2}$.. 69

3-13. Heme transfer between methemoglobin and $IsdB^{N1LN2}$. 70

3-14. Heme transfer reactions between methemoglobin (metHb) and various IsdB constructs... 72

3-15. Comparison of $IsdB^{N1}$ to IsdX1 from *Bacillus anthracis*.....75

4-1. Labeled ^{1}H-^{15}N HSQC spectrum of $IsdB^{Linker}$. 79

4-2. Secondary structural topology of $IsdB^{Linker}$ 80

4-3. CD scan of $IsdB^{Linker}$ 81

LIST OF FIGURES-CONTINUED

Figure Page

4-4. Titration experiment of $IsdB^{Linker}$ and $IsdB^{N2}$.....82

4-5. Overlay of ^{1}H-^{15}N HSQC spectra of $IsdB^{Linker}$, $IsdB^{N2}$, and $IsdB^{Linker-N2}$.....83

4-6. Overlay of ^{1}H-^{15}N HSQC spectra of $IsdB^{35L-N2}$ and $IsdB^{20L-N2}$.....84

4-7. Overlay of ^{1}H-^{15}N HSQC spectra of $IsdB^{N1}$ and $IsdB^{N1-Linker}$. 85

4-8. Sequence alignment of $IsdB^{Linker}$ and $IsdH^{Linker}$.....87

5-1. Overlay of the ^{1}H-^{15}N HSQC spectra of wild type apo-$IsdB^{N2}$ and mutated apo-$IsdB^{N2}$ M362A/I438A. 89

5-2. Labeled ^{1}H-^{15}N HSQC spectrum of apo-$IsdB^{N2}$ M362A/I438A. ……..91

5-3. Ramachandran plot of apo-$IsdB^{N2}$ M362A/I438A.. 93

5-4. The NMR solution structure of apo-$IsdB^{N2}$ M362A/I438A. 95

5-5. Structure alignment between apo-$IsdB^{N2}$ M362A/I458A and holo-$IsdB^{N2}$... 97

5-6. Overlay of the ^{1}H-^{15}N HSQC spectra of apo-$IsdB^{N2}$ M362A/I438A and holo-$IsdB^{N2}$ M362A/I438.....97

5-7. Comparison of the structure of apo-$IsdB^{N2}$ M362A/I438A.....98

ABSTRACT

Staphylococcus aureus is an opportunistic pathogen which when left unchallenged can cause severe toxicity and death in mammals. Critical to *S. aureus* growth is the ability to scavenge iron from hemoglobin (Hb). To acquire iron *S. aureus* has evolved a sophisticated protein-mediated heme acquisition pathway, which comprises nine iron-regulated surface determinant (Isd) proteins involved in heme capture, transport and degradation. A key protein of the acquisition pathway is the hemoglobin receptor protein IsdB, which comprises two NEAr transporter (NEAT) domains that act in concert to bind Hb and extract heme for subsequent transfer to downstream acquisition pathway proteins. Despite significant advances in the structural knowledge of other Isd proteins, the mechanisms and molecular basis of the IsdB-mediated heme acquisition process is not well understood.

In order to provide more insights into the mode of function of IsdB, structural studies via nuclear magnetic resonance (NMR) spectroscopy were employed on different domains of IsdB. The three-dimensional solution structure of $IsdB^{N1}$ revealed an immunoglobulin-like fold that is consistent with other NEAT domain proteins. Site directed mutagenesis studies revealed two key aromatic residues, F164 and Y167, involved in methemoglobin (metHb) interactions with IsdB. The protein variant F164D did not bind to metHb under NMR conditions. In heme transfer studies between metHb and IsdB constructs containing the two NEAT domains and the linker region, the amino acid substitution of F164D diminished but did not knock out the ability of IsdB to remove heme from metHb. A double amino acid substation of F164D and Y167D did abolish heme transfer from metHb to IsdB, therefore identifying key residues of $IsdB^{N1}$ interaction with metHb. Studies of the linker region revealed an overall α-helical propensity and an interaction between the linker region and the second NEAT domain, $IsdB^{N2}$. Solving the apo-$IsdB^{N2}$ structure revealed slight differences in the heme-binding pocket, specifically in β-strands 7 and 8 that interact with the heme moiety, when compared to the published crystal structure of holo-$IsdB^{N2}$. The findings in this thesis provide a structural role for $IsdB^{N1}$ enhancing the rate of extraction of heme from metHb by $IsdB^{N2}$ and interactions between the domains of IsdB.

CHAPTER ONE

INTRODUCTION

Defense against the opportunistic pathogen, *Staphylococcus aureus*, is a constant challenge for the human body. Approximately 30% of individuals are carriers of *S. aureus*, which predominantly colonizes on the skin and inside the nose. Given the opportunity, however, *S. aureus* invades the body and employs elaborate strategies to avoid the innate immune system. To aid the immune system in its defense against invasion, antibiotics such as methicillin and vancomycin are administered to infected persons. Unfortunately, many strains of *S. aureus* resist these antibiotics and new strategies are needed. In war, blocking the enemy's supply lines effectively causes starvation. As a means to fight *S. aureus* infection, understanding its essential nutrient pathways and creating new therapeutics that could block the import of these nutrients is of great importance. An essential nutrient to most living organisms is iron. The human body tightly regulates iron, because of its toxicity and as a defense against pathogenic organisms such as *S. aureus*. The development of elaborate systems to steal iron from its host allows *S. aureus* to survive even in low iron environments. Understanding these iron acquisition pathways is the first step into developing small molecules or peptides that could block the import of iron, thereby starving the organism. The preferred source of iron for *S. aureus* is in the form of heme, which is found predominantly in red blood cells bound to hemoglobin. The structure/function relationship of proteins is well established and thus elucidating the structure of a protein at the molecular level can reveal details

about the mechanism of function (1). To gain insight into how *S. aureus* utilizes host hemoglobin as its iron source, the structure of the hemoglobin receptor, IsdB, was investigated by nuclear magnetic resonance (NMR) spectroscopy. The IsdB protein is part of an elaborate pathway that sequesters, transports, and internalizes heme from hemoglobin. This chapter discusses the current knowledge of iron acquisition by bacterial pathogens from their host, with emphasis on the Isd pathway utilized by *Staphylococcus aureus.*

Staphylococcus aureus

Dr. Alex Ogston first described (around 1880) a form of *Micrococcus* that grew in grape-like clusters, and termed it *Staphylococcus* (2). There are over 40 species of *staphylococci*, and most of them live as part of the natural flora of mammals. However a few species do cause diseases in humans. Of those few *staphylococci* species that cause infection, *S. aureus* is by far the most devastating. At least one third of the human population asymptomatically carries *S. aureus* as part of their normal flora either on the skin or in the nasal carriage. However, *S. aureus* is an opportunistic pathogen and will invade its unsuspecting host if barriers into the body are broken. Once inside, *S. aureus* can infect the bloodstream and deep tissues, causing a plethora of infections, including skin infections: cellulitis, impetigo, scalding skin syndrome, and abscesses, and more invasive infections such as bacteremia, endocarditis, and osteomyelitis, which can easily become life threatening. *S. aureus* infection is the most prevalent nosocomial (hospital-

acquired) infection and is responsible for approximately 20% of blood infections in US hospitals since the 1960's (3,4).

The severity of invasive *S. aureus* infections have increased with the emergence of antibiotic resistant strains. In the 1950's, strains of *S. aureus* developed resistance to penicillin, so methicillin was introduced in 1959. However, within two years, methicillin-resistant strains of *S. aureus* (MRSA) were isolated in the United Kingdom, followed by a host of other countries, including Japan and the United States (5). There is now a growing concern for both hospital-acquired and community-acquired MRSA. In order to treat MRSA, the antibiotic vancomycin is generally used, but vancomycin resistant strains are emerging as well. This iterates the need for new therapeutics.

The pathogenicity of *S. aureus* can be attributed to a multitude of strategies employed by the organism to evade the innate immune system, and avoid phagocytosis (6). A thick cell wall along with expression of a polysaccharide capsule can inhibit incorporation of antimicrobial peptides (7). Recognition by neutrophils can be impeded by the expression of proteins, such as Protein A that binds to IgG in the incorrect orientation for recognition by the Fc receptor (8) or Clumping factor A (ClfA) that binds fibrinogen (6). *S. aureus* excretes a number of cytolytic toxins that lyse host cells, including leukocytes (6). Various adhesion proteins that are able to attach to almost any human tissue cover the surface of *S. aureus*. The secreted staphylokinase proteins can affect host blood clotting which allows the bacteria to spread through the blood stream more efficiently (6). Evidence for the internalization of *S. aureus* by host cells that do not normally phagocytize is increasingly pointing to *S. aureus* as an intracellular pathogen as

well (9). The damage to human tissue and extraction of nutrients is essential to the virulence of *S. aureus* infection, and the most limiting resource is iron.

Role of Iron in Biology

Iron is an essential nutrient for most of life and is required for a variety of biological processes including DNA biosynthesis, photosynthesis, N_2 fixation, H_2 production and consumption, respiration, oxygen transport, and gene regulation (10). Iron is commonly found in one of two redox states, the reduced Fe^{2+} ferrous form and the oxidized Fe^{3+} ferric form. However, at physiological conditions (pH 7) Fe^{3+} is insoluble and forms iron hydroxide aggregates with a solubility constant around 10^{-18} M (10). At physiological conditions, Fe^{2+} is soluble but can activate the Fenton reaction, $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH\bullet + OH^-$, causing the reduction of hydrogen peroxide into hydroxyl radicals, which are detrimental to macromolecules. Iron is therefore tightly regulated by organisms and is usually bound to proteins either alone, in iron-sulfur clusters, or as heme. The ability of iron to switch between the two major redox states Fe^{2+}/Fe^{3+} allows it to serve as the catalytic center of many enzymes involved in redox reactions. The ligands of the iron can greatly change the redox potential (10).

The most prevalent form of iron in the human body is bound to the protoporhyrin IX molecule, heme (11). Heme is hydrophobic and can intercalate into lipid membranes and cause oxidative damage, therefore it is usually found bound to hemoproteins. It is used in oxygen transport (hemoglobin) and electron transport (cytochrome c) proteins (12).

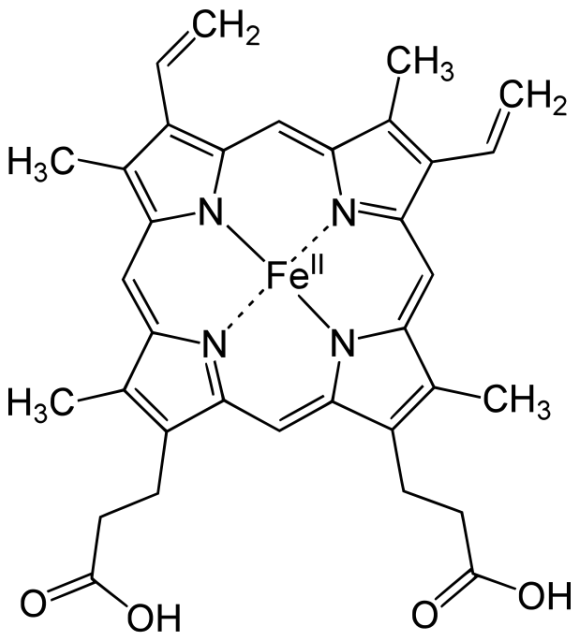


Figure 1-1. The structure of iron protoporphyrin IX. Heme is the largest source of iron in the human body. Iron is bound to the center of the ring system by four nitrogen atoms.

For pathogenic bacteria, iron is a difficult commodity to acquire in the host. Fe^{2+} is extremely limited and Fe^{3+} is highly insoluble. Therefore the largest sources of iron in the host for bacteria are iron-binding proteins. Ferric iron is sequestered and transported by the glycoproteins, transferrin and lactoferrin. Both proteins comprise two domains capable of binding one Fe^{3+} each. Transferrin (Tf), predominantly found in blood serum, binds to Fe^{3+} with a $K_d \sim 10^{-23}$ M. At normal physiological conditions, transferrin in only 30% saturated with iron (13) and acts as a buffer in the blood. Lactoferrin, a similar glycoprotein, is found in milk and most other mucosal secretions (14). In addition to iron sequestration, lactoferrin plays an important role in the innate immune system and is a promising antimicrobial (14). Both proteins have an extremely high affinity for Fe^{3+} (12) and therefore creates a challenge for iron acquisition by bacteria.

Heme-binding proteins, the largest storeroom of iron in the host, can be a source for bacteria of either iron or the heme molecule itself. The majority of heme iron (>2 g of iron) (13) is located in red blood cells bound to the oxygen transport protein, hemoglobin (Hb). Hemoglobin comprises two α- and two β-chains (~16 kDa each) forming a tetrameric protein. Each subunit binds to one heme molecule. The four nitrogen atoms of the protoporphyrin IX moiety and an axial histidine ligate the iron. In oxyhemoglobin, divalent oxygen is bound to the sixth ligand position, while in the oxidized form, methemoglobin (metHb) this axial position is free. When red blood cells are lysed and hemoglobin is released into the blood stream, hemoglobin can quickly oxidize to form the dimeric metHb. (12). In the event that red blood cells are lysed, the body releases the large glycoprotein, haptoglobin (Hp) to bind and clear hemoglobin from the bloodstream (12) and hemopexin to bind free heme (15). The Hp-Hb complex must be degraded upon clearance to the liver because the binding affinity of Hp to hemoglobin is so high that the complex cannot be broken (12). Because free heme is toxic to cells, hemopexin is released to bind heme that may have dissociated from Hb in the event of red cell lysis and transport the heme to the liver. Hemopexin has an extremely high affinity for heme with a $K_d < 1pM$ (15).

Bacterial Iron Acquisition Systems

Acquiring the tightly regulated supply of iron from the human body is a challenge for bacteria. Bacteria, nevertheless, have developed elaborate systems to aid in this endeavor. Both Gram-negative and Gram-positive bacteria have evolved to acquire this iron either freely from the extracellular milieu, remove it from iron-binding proteins, or

to extract it in the form of heme from the heme-binding proteins of the host.

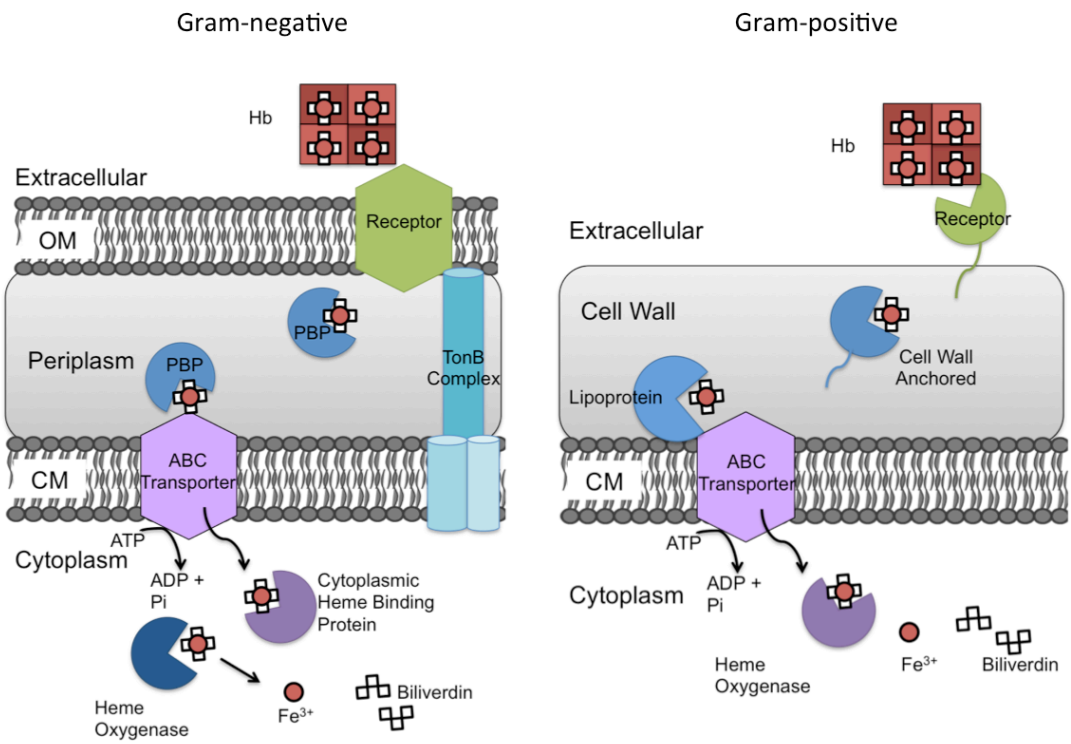


Figure 1-2. Comparison of heme import by Gram-negative and Gram-positive bacteria. In Gram-negative bacteria, the receptor binds and abstracts heme from hemoglobin (Hb) and transports the heme across the outer membrane (OM) via energy coupled by the interaction with the TonB system. The heme is bound by a periplasmic binding protein (PBP) and transported through the periplasm to the ATP-binding cassette (ABC) transporter. The heme is transported through the membrane by hydrolysis of ATP and once in the cytoplasm is degraded by heme oxygenases (16). In Gram-positive bacteria, the receptor, attached to the peptidoglycan cell wall, binds and abstracts the heme. The heme is transported to the cellular membrane by cell wall anchored proteins. The membrane associated lipoprotein then transports the heme to the ABC transporter where the heme molecule is driven through the membrane by hydrolysis of ATP. The heme molecule can then be degraded by heme oxygenases.

Gram-negative Bacteria

Although ferrous iron is tightly regulated in the host, it is more accessible in the anaerobic conditions of the gut-intestinal tract. As such the gram-negative organisms *Escherichia coli*, *Helicobacter pylori*, and *Porphyromonas gingivalis* that populate the gut-intestinal tract do have a specific transport system for ferrous iron, the Feo system (17). The Fe^{2+} is believed to diffuse through porins in the outer membrane, into the periplasmic space. The ferrous iron is then shuttled through the cytoplasmic membrane by an actively driven ATP/GTP process via the FeoB protein, a permease that contains an N-terminal hydrophilic domain and a C-terminal integral-membrane domain with two gate motifs (18).

Gram-negative bacteria employ to methods to acquire iron/heme from iron containing host proteins. The first is to have direct contact with the iron containing host proteins via receptors located on the outer membrane (OM) (16). The second is to secrete iron scavengers to bind iron/heme from the environment and then recognize them when they have returned to the cell surface. Most gram-negative bacteria have two different types of secreted molecules, siderophores, which are low molecular weight compounds with a high affinity for Fe^{3+}, or hemophores, proteins with a high affinity for heme (19).

The systems for both host protein interaction and uptake of bacterial secreted iron scavengers are similar (16). There is an outer membrane (OM) receptor that recognizes and binds the host protein and/or the iron scavenger (see Fig. 1-2). The iron/heme is transported through the receptor by energy coupled with the potential of the cytoplasmic membrane by the TonB-ExbBD complex. The Ton complex comprises ExbB, an integral

membrane protein, and ExbD and TonB, the membrane-anchored periplasmic proteins. The iron/heme is either released in the periplasm or picked up by a periplasmic binding protein (PBP) and shuttled to the ATP-binding cassette (ABC) transporter, comprising the cytoplasmic ATPase/permease for transport through the cytoplasmic membrane into the cytoplasm (16).

There are many examples of outer membrane receptors that recognize host iron containing proteins. Transferrin and lactoferrin receptors have been reported for *Neisseria* and *Haemophilus* bacteria. The systems are comprised of transferrin binding protein (Tbp) TbpB/TbpA or lactoferrin binding protein (Lbp) LbpB/LbpA (20). The TbpB/LbpB proteins are attached to the outer membrane by a lipid anchor and possibly facilitate the binding of transferrin/lactoferrin to TbpA/LbpA respectively. TbpA/LbpA are integral membrane proteins that are predicted based on sequence homology to use surface loops to bind transferrin/lactorferrin, releasing the tightly bound ferric iron by forcing the domains apart. After iron is transported through the OM, it is picked up by the ferric ion binding protein FbpA and shuttled to the FbpB/FbpC transporter to be incorporated through the cytoplasmic membrane (20).

An example of the heme transport system is found in *Shigella dysenteriae*. Shigella heme uptake (Shu) system comprises the proteins ShuA, ShuT, and ShuUV. ShuA is the heme binding protein receptor, which was found to be specific for methemoglobin. ShuA binds heme with a conserved bis-histidine ligation of His-86 and His-420. These histidine residues are also found in the homologous protein of *Yersinia enterocolitica*, HemR (21), and HasR of *Serratia marcescens*. It was shown that ShuA

formed a complex with the full length TonB. The heme is then transported to ShuT, the heme perisplasmic binding protein, which includes the conserved glutamine residues that are presumably required for docking with the conserved lysine residues of the permease ShuUV. ShuT has an extremely high affinity for heme, and upon protein interactions with the cytoplasmic ABC transporter, along with hydrolysis of ATP, there is enough energy to release the heme from ShuT and transport it through a cytoplasmic membrane (22).

In order to scavenge iron from the environment, Gram-negative bacteria secrete siderophores. Siderophores are low molecular weight compounds that have a high affinity for iron, specifically Fe^{3+}. They are usually negatively charged so their affinity for Fe^{2+} is relatively low. In *E. coli* the siderophores, ferrichrome, enterobactin, and ferric citrate are recognized and transported by the receptors FhuA, FepA, and FecA, respectively (19). The receptors share the same overall fold of a 22-stranded β-barrel comprising the C-terminal residues, while approximately 150 N-terminal residues form a plug domain comprising a four β-strand sheet surrounded by loops and helices inside the barrel (Fig. 1-3). There are 11 extracellular loops that extend anywhere from 30 to 40 Å from the barrel. Some of these loops interact with the siderophore and close upon binding. The unwinding, or partial unwinding, of the switch helix of the plug domain upon siderophore binding, causes a conformational and positional change of the Ton box, an N-terminal sequence that interacts with TonB. The energy derived from the protein interactions and the TonB complex drive the siderophore complex through the receptor into the periplasm. Their respective periplasmic binding protein binds and then shuttles the siderophores to the ABC transporter for incorporation into the cytoplasm (20). Understanding the

mechanism of siderophore release and capture by bacteria has increased the overall knowledge of iron acquisition for other systems as well.

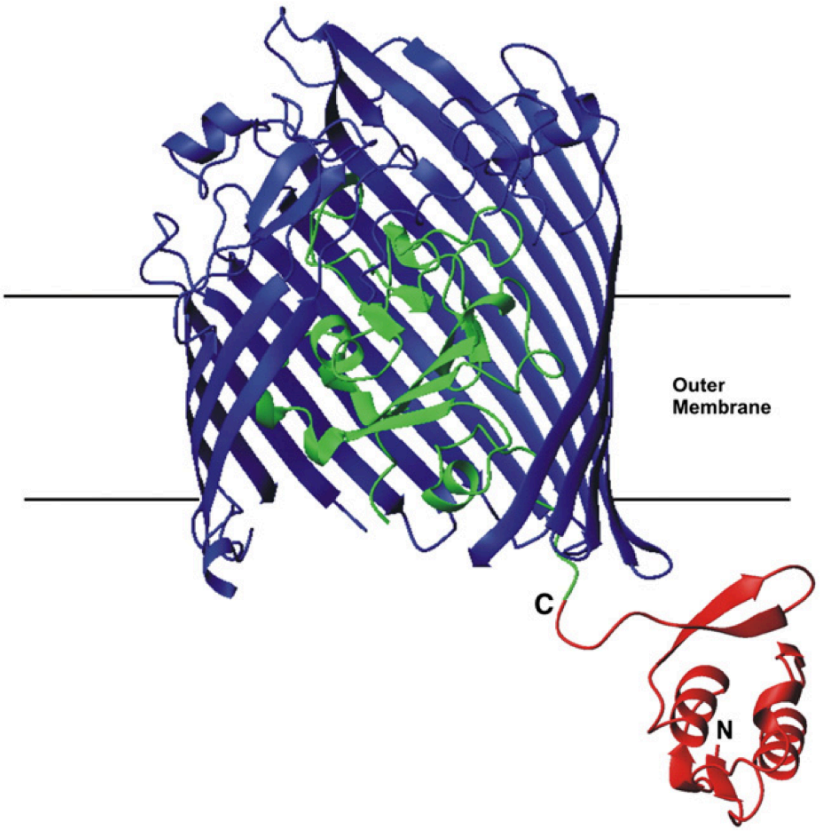


Figure 1-3. The structure of the siderophore receptor, FecA. The 22 stranded β-barrel in blue (the front section has been removed for clarity). In red, the N-terminal signaling domain and in green the plug domain. PDB ID: 2DIU and 1KMO. Reproduced from Krewulak and Vogel, 2008 (20).

The second method employed by Gram-negative bacteria is the secretion of hemophores, proteins with a high affinity for heme. HasA, a small hemophore secreted by *Serratia marcescens* has an extremely high affinity for heme, K_d ~18 pM (23). In order to bind heme, loop L1 undergoes a substantial structural change of ~30 Å for His-32 to ligate the iron, as shown in Figure 1-4. The heme is then passed to the outer

membrane receptor HasR, a β-barrel formed by 22 anti-parallel β-strands. At a K_d equal to 0.2 μM, the affinity of HasR for heme is substantially lower than the affinity of HasA. To overcome this, one of the two axial ligands of the heme in HasA is disrupted upon binding with HasR and Ile-671 of HasR sterically displaces the second ligand (24). In *Haemophilus influenza*, HxuA is a hemophore that acquires heme from hemopexin (25).

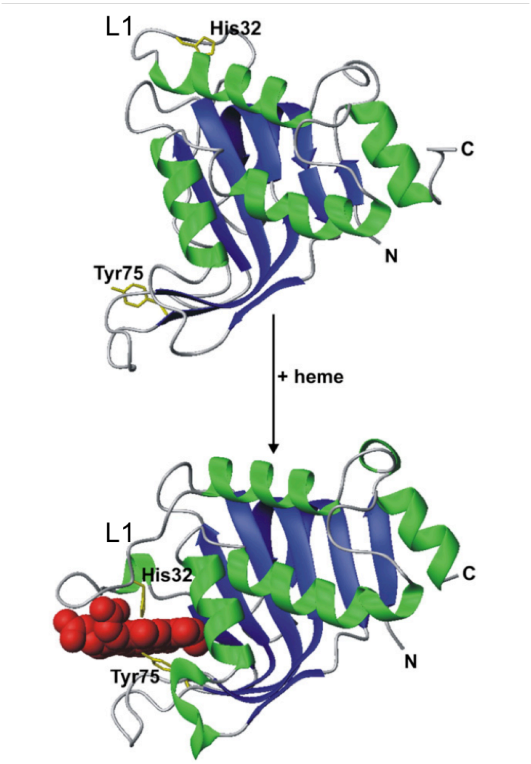


Figure 1-4. The structure of the hemophore HasA from *Serratia marcescens.* The apo- and holo- forms of HasA are shown in cartoon representation with β-strands in blue and α-helices in green. The two axial ligands are shown as stick representations. Reproduced from Krewulak and Vogel, 2008 (20).

Gram-positive Bacteria

For Gram-positive bacteria, the difficulty arises from having to pass nutrients across the thick outer cell wall. They do not contain an outer membrane and periplasm

space, so they do not utilize a TonB complex like in Gram-negative bacteria. They do, however, use ABC transporters to move iron/heme across the cytoplasmic membrane (26). Attached to the cytoplasmic membrane is a cell wall that is approximately 15 to 30 nm thick, which is a substantial barrier to moving ligands bound at the surface to the membrane. The cell wall comprises disaccharide *N*-acetylmuramic acid-(β1-4)-*N*-acetylglucosamine (MurNAc-GlcNAc) linked to the tetrapeptide (L-Ala-D-isoGln-L-Lys-D-Ala) (Fig. 1-5). The chains of MurNAc-GlcNAc are bound to the cell wall peptide via an amide bond between the L-Ala and the MurNAc. The chains are then cross-linked via a pentaglycine cross bridge producing the continuous layer of peptidoglycan surrounding the cytoplasmic membrane (27).

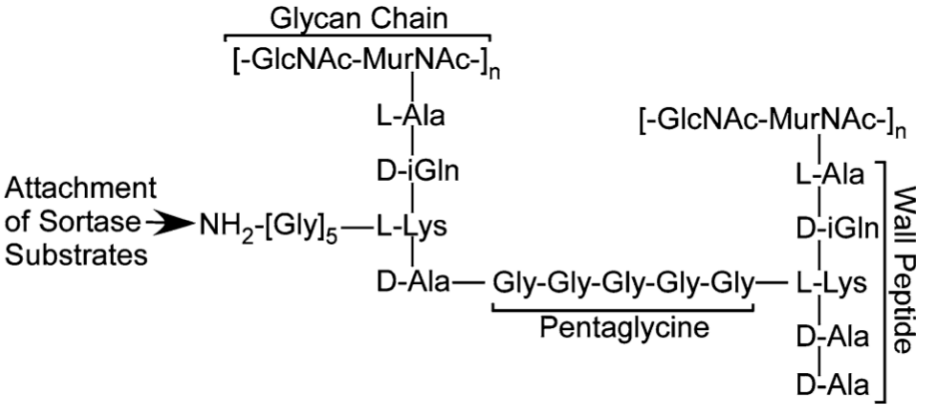


Figure 1-5. Structure of the peptidoglycan cell wall of *S. aureus*. Reproduced from Maresso and Schneewind 2006 (28).

In order to carry heme across this barrier, Gram-positive bacteria utilize similar systems as those utilized by Gram-negative bacteria, including siderophores. Additionally, they possess protein pathways that are attached to the cell wall and move the heme via protein relay to the cytoplasmic membrane. *S. aureus* acquires iron from

transferrin via siderophores (27,29). It produces two siderophores, staphyloferrin A and staphyloferrin B and can utilize other bacterial siderophores known as xenophores (27). The genes involved in siderophore synthesis are *sfnABCD* for staphyloferrin A and *sbnABCDEFGHI* for staphyloferrin B. Both operons have an ABC transporter gene adjacent to them, *htsABC* and *sirABC*. Both encode for a lipoprotein for siderophore binding and the transporter. Neither system, however, has a dedicated ATPase expressed from the operon, but FhuC is required for siderophore transport so is presumably the ATPase for multiple ABC transporter systems (27).

Originally hemophores were thought to be present only in Gram-negative bacteria, until the discovery of the iron-regulated surface determinant pathway (Isd) in *Bacillus anthracis* provided the first evidence of hemophores in Gram-positive bacteria (30). *B. anthracis* excretes two hemophores, a small ~15 kDa protein, IsdX1 and a much larger ~100 kDa protein, IsdX2. Although IsdX1 and IsdX2 are found in the extracellular milieu, a small fraction of IsdX2 associates with the cellular envelope (30). IsdX1 abstracts heme from hemoglobin and transfers the heme to either IsdX2 or IsdC, a protein anchored to the cell surface of *B. anthracis (31)*. In *Streptococcus pyogenes,* a heme relay pathway is embedded throughout the cell wall to capture heme on the surface and move it to the cytoplasmic membrane (32). The hemoglobin receptor, Shr, is exposed on *S. pyogenes* cell surface where it can interact with methemoglobin and abstract heme (33). Shr then passes the heme to the cell wall embedded protein, Shp, that binds heme with a unique two methionine ligand coordination. The heme is then passed to the lipoprotein, HtsA for incorporation into the cytoplasm by an ABC transporter (33,34).

Iron-regulated Surface Determinant (Isd) Pathway

Heme is the preferential source of iron for *S. aureus (35).* The heme capture and acquisition by *S. aureus* is mediated by a sophisticated network of surface-accessible iron-regulated surface determinant (Isd) proteins, which interact in a concerted way to abstract and transfer heme from hemoglobin to downstream protein effectors (36).

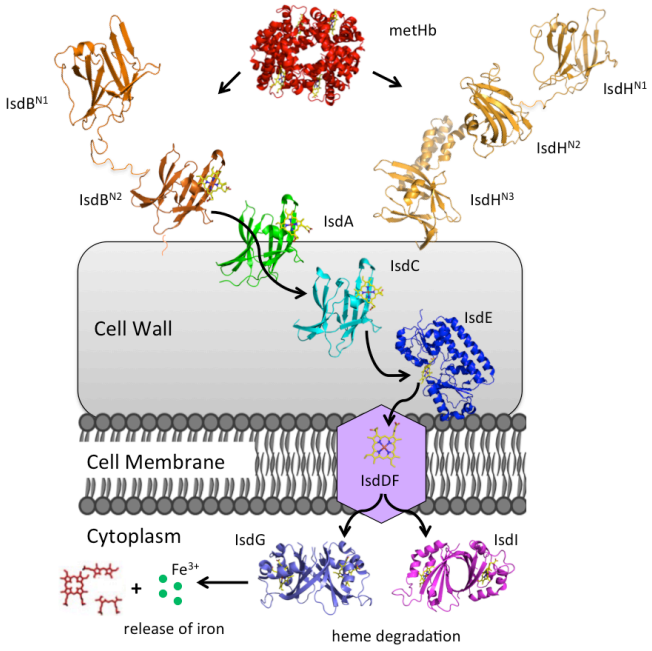


Figure 1-6. The Iron-regulated Surface Determinant pathway of *S. aureus*. Protein structures reproduced from PDB structural depositions, IsdB^{N2} (3RTL), IsdH^{N1}(2H3K), IsdH^{N2LN3} (4IJ2), IsdA (2ITE), IsdC (2O6P), IsdE (2Q8Q), IsdG (2ZDO), IsdI (3LGN).

Regulation and Expression

Iron is essential but can be toxic to the cell, so the proteins involved in iron acquisition are tightly regulated and expressed only when iron concentrations are deficient. The ferric uptake regulator protein, Fur, regulates expression of the Isd proteins

(36). Fur binds to Fe^{2+} when iron levels are sufficient (Figure 1-7). The dimeric form of Fur recognizes and binds a specific sequence of nucleotides (termed the Fur box) upstream of the promoter region inhibiting transcription of the *isd* operon.

When iron levels are low, Fur releases its Fe^{2+} causing conformational changes that release the protein from the promoter region, allowing RNA polymerase to initiate transcription of the operon (37). Seven of the *isd* genes are located in the Isd operon, *isdABCDEFG* but *isdH* and *isdI* are located in different genomic areas, but their expression is also controlled by Fur (36,38).

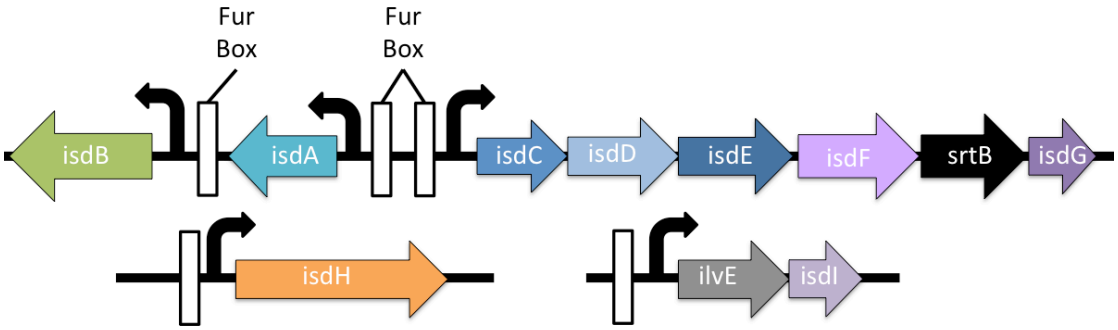


Figure 1-7. The Isd operon in *S. aureus*. Each black line represents a different area of the *S. aureus* genome. The white boxes represent the Fur binding site for transcriptional repression in times of high iron concentrations. (Figure adapted from Skaar, E. and Schneewind, O. 2004.) (39)

Role of the Isd Proteins

Surface Proteins: IsdB, IsdH, IsdA, and IsdC. The surface proteins are directed to the cell wall by an N-terminal secretion signal sequence. IsdB and IsdH are fully exposed on the staphylococcal cell surface as determined by proteolytic studies with proteinase K (38). IsdA is partially exposed on the surface and IsdC is embedded within the cell wall

(38). The four proteins are attached to the cell wall by a family of sortase enzymes. Sortase A (SrtA) attaches the majority of extracellular proteins in *S. aureus,* including IsdA, IsdB, and IsdH (40). The C-terminal tail includes an LPXTG motif where SrtA cleaves between Thr and Gly, and attaches the carboxyl group of Thr to the peptide crossbridges via a peptide bond. However the *isd* operon includes a different sortase enzyme, Sortase B (SrtB) (36). SrtB is responsible for attaching IsdC to the cell wall as IsdC has a different motif of NPQTN in its C-terminal tail.

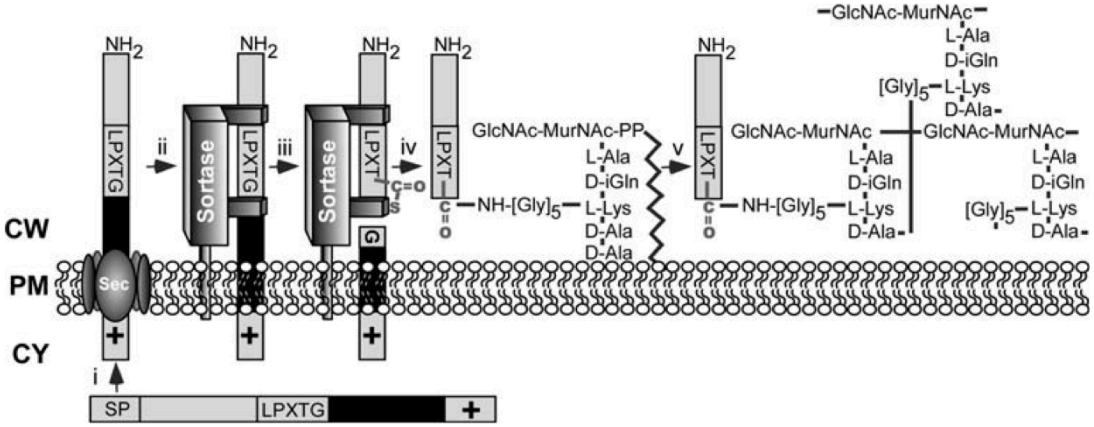


Figure 1-8. Sortase attachment of proteins to the cell wall of *S. aureus*. Proteins with the N-terminal secretion signal are exported through the membrane by the Sec pathway. Reproduced from Maresso and Schneewind 2006 (28)

In addition to their iron transport functions, IsdH, IsdB, and IsdA function as protein receptors and adhesion molecules. IsdB and IsdH bind to hemoglobin (35,41,42). IsdH has also been shown to bind to the larger haptoglobin-hemoglobin complex (43). IsdA does not bind hemoglobin, but acts as a human transferrin (hTf) receptor and a broad-spectrum adhesion protein shown to bind fibronectin (Fn), fibrinogen (Fg) and other human ligands. The dissociation constants for IsdA binding to Fg, Fn, and hTf are

all in the micromolar range, which suggests that IsdA functions as an adhesion molecule (44,45). IsdB has been shown to promote adherence of *S. aureus* to platelets and directly interacts with the platelet integrin GPIIb/IIIa (46). IsdB also interacts with other human cell lines, 293T and HeLa, and promotes *S. aureus* invasion of those cells in the absence of fibronectin (9). Vaccine trials have shown that antibodies to IsdA and IsdB can provide protective immunity in mouse models (41).

IsdC has been termed the central conduit in the Isd pathway (47). Iron is shuttled from the outer receptor proteins to IsdC, which has higher affinity for heme than holo-IsdA, *in vitro* (48,49). Homologous IsdC proteins have been found in *B. anthracis* (50) and *L. monocytogenes* (51), which suggest the importance of this central protein in the pathway.

Membrane Transporter Proteins: IsdE, IsdD, and IsdF. Heme transport through the membrane requires energy provided by ATP. Similar to the ATP-binding cassette (ABC) transporters of Gram-negative bacteria, *S. aureus* moves heme through the cell membrane via the transporter proteins IsdDEF and an ATPase (possibly FhuC). IsdE, the cognate substrate-binding lipoprotein, associates with the cell membrane. It belongs to the helical backbone metal receptor superfamily. The solved crystal structure of heme

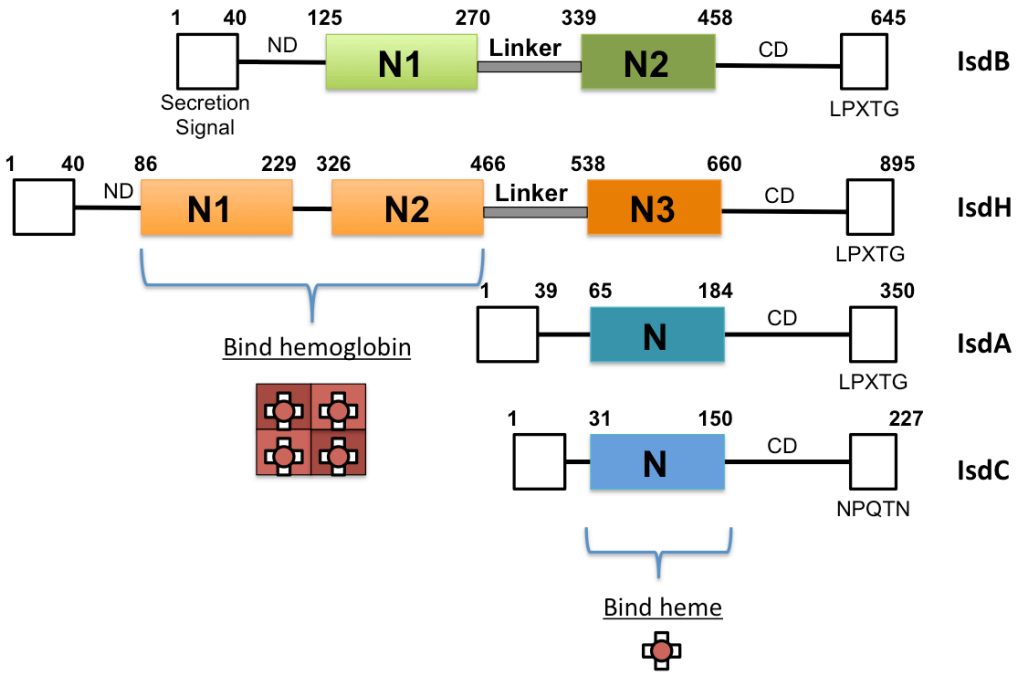


Figure 1-9. The surface proteins of the Isd pathway in *S. aureus*. A schematic of the cell surface Isd proteins, IsdA, IsdB, IsdC, and IsdH. The numbers at the top correspond to amino acid number. The first boxes represent the N-terminal secretion signal, the NEAT domains are labeled as N1-N3, and below the C-terminal boxes are the sortase anchoring motif.

bound IsdE revealed a bilobe two-domain protein with each lobe containing a β-sheet surrounded by α-helices and a long α-helix connecting the two lobes. The heme binds in a deep groove, where it is only 19% solvent exposed, located between the two lobes and unlike in the NEAT domains is 6-coordinate. The axial ligand arrangement may be important for both the acquisition of heme from IsdC and the subsequent transfer of heme to the permease, IsdF. The heme is ligated by a methionine and histidine residue, which is a unique ligand arrangement for heme transport proteins, but several *c* type cytochromes use this combination for heme coordination. (52). IsdD seems to be associated with the membrane, although its function remains to be elucidated. The sequence of IsdF predicts

a permease structure, a homodimeric integral membrane protein that allows the transport of heme through the membrane and into the cytoplasm. The Isd system is missing an obvious ATPase, but there are several iron-compound ABC transporters from *S. aureus* that lack a corresponding ATPase including HtsABC and SirABC. In those cases, FhuC provides the ATPase function.

Heme Degrading Enzymes: IsdG and IsdI. Once inside the cytoplasm, heme has two fates, either it is transferred to other heme proteins or the heme molecule is degraded, thereby releasing the iron for use in other cellular processes. The two proteins of the Isd system responsible for heme degradation are IsdG and IsdI. IsdG is expressed as part of the *isd* operon, while IsdI is found elsewhere in the genome. These two proteins have unique structures compared to other known monooxygenases, and are part of the ferrodoxin fold superfamily. IsdG and IsdI form a β-barrel at the dimer interface and each subunit binds a heme molecule with a His-77. These proteins ruffle the porphyrin ring, which has not been previously observed (53). The heme is degraded to release iron in the presence of NADPH-cytochrome P450 oxidoreductase or ascorbate (53,54).

Near Iron Transporter (NEAT) Domains. All of the cell surface Isd proteins contain at least one copy of a NEAT domain. Genomic analysis of pathogenic bacteria revealed the presence of a recurring domain located nearby known Fe^{3+} transporters (55). Based on sequence analysis, the NEAr iron Transporter (NEAT) domain adopts a predominantly β strand secondary structure and contained an N-terminal secretion signal.

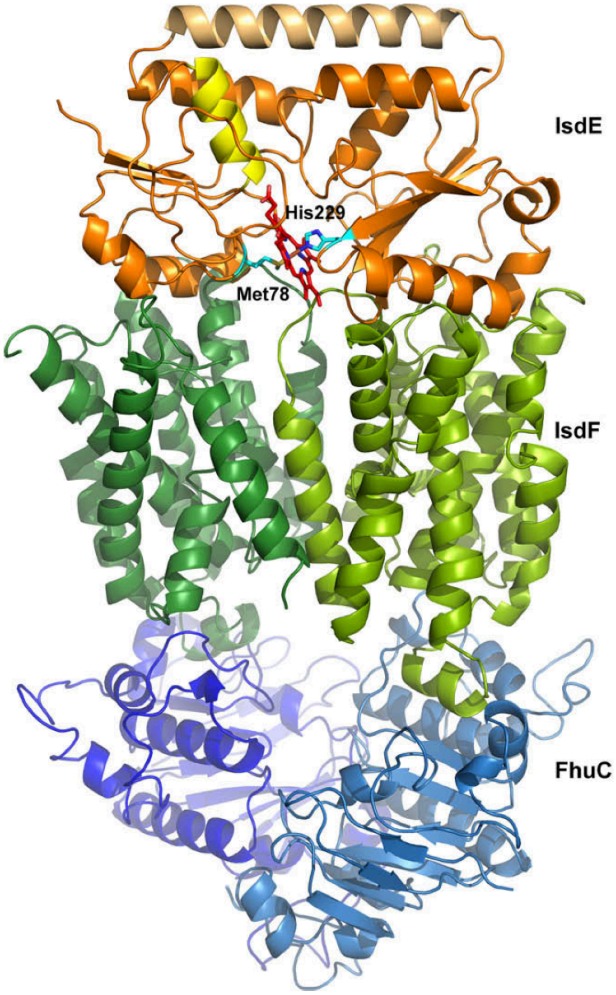


Figure 1-10. A model of the Isd ATP-Binding Cassette transporter of *S. aureus*. The crystal structure of IsdE (2Q8Q) was docked onto a model of the predicted IsdF structure along with FhuC, the probable ATPase for this system. Reproduced from Grigg, J. et. al. 2011 (56).

The presence of NEAT domains was confirmed in other pathogenic Gram-positive bacteria, including *Listeria monocytogenes* and *Clostridium perfringens (55).*

There has been a great interest in examining the structure of the NEAT domain in order to help understand its role in heme binding and transfer within the Isd proteins. To date, the NEAT domains of IsdH, IsdA, IsdC, and the C-terminal NEAT domain of IsdB

have been solved (47,57-61). These structural studies have shown that the NEAT domain adopts an immunoglobulin–like (Ig) fold, a seven β-strand antiparallel β sandwich with an eighth strand that connects the two sides. The heme-binding pocket has conserved hydrophobic residues and an YXXY motif where tyrosine is the axial ligand. Unlike catalytic heme proteins, such as peroxidases, whose heme molecules are buried away from the solvent, heme transport proteins appear to have a more exposed heme molecule. The NEAT domains of IsdA, IsdB, IsdC, and IsdH all have heme groups with 30-40% solvent exposure which functions to efficiently acquire heme and then pass it to the next protein in the chain (47,58,59,61).

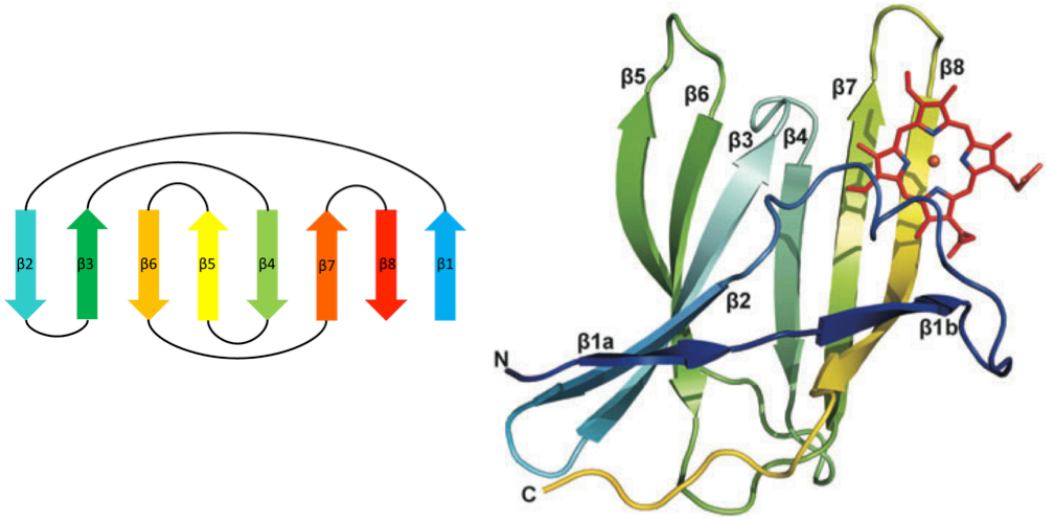


Figure 1-11. The immunoglobulin-like fold of the NEAT domain. A 2D schematic of the β-strand architecture of the NEAT domain on the left. The crystal structure of IsdA (PDB 2ITF) with β-strands labeled on the right. The NEAT domain figure was reproduced from Grigg, J. et. al. 2007 (52).

Hemoglobin Receptors

Many pathogenic bacteria utilize hemoglobin as an iron source. In order to capture the heme from hemoglobin, bacteria employ receptors on the outside of the cell. In the Gram-negative bacteria, *Shigella dysenteriae,* uses a large 73 kDa protein, ShuA, which binds and abstracts heme from methemoglobin via two conserved histidine residues His-86 and His-420 (22). These histidine residues are conserved in HemR (21) and HasR, the hemoglobin receptors from *Yersinia entercolitica* and *Serratia marcescens (22).* The β-barrel architecture forms a pore through the outer membrane and is blocked by a plug domain. In Gram-positive bacteria, the hemoglobin receptors do not form β-barrel structures. Instead, Shr from *Streptococcus pyogenes* contains two NEAT domains connected by a leucine-rich region. Unlike the hemoglobin receptors in *S. aureus*, IsdB and IsdH, both of the NEAT domains in Shr bind heme. Additionally, the NEAT domains do not bind metHb but instead the undefined N-terminal domain interacts with metHb. This N-terminal domain contains a unique amino acid sequence and appears only in other *S. pyogenes* orthologues and a few Duf1533 proteins. Therefore the mechanism of Hb recognition and subsequent binding is not currently known (33).

Scope of Dissertation

S. aureus is a devastating infectious bacterium that is a threat to global health. Its ability to obtain iron increases virulence during infection. The focus of this research thesis is to understand the unique functions of the IsdB domains in heme acquisition from metHemoglobin. The structure of the individual domains of IsdB and their interaction

with each other were investigated by nuclear magnetic resonance (NMR) spectroscopy. The materials and methods employed in this study can be found in Chapter 2. In Chapter 3, the three-dimensional (3D) solution structure of the N-terminal NEAT domain of IsdB, referred to as $IsdB^{N1}$, is presented. Mutagenesis studies revealed amino acid residues involved in methemoglobin binding. In Chapter 4, investigations into the structural and functional role of the connecting region between the two NEAT domains of IsdB, termed the linker region, will be discussed. The linker region adopts a predominantly alpha helical structure based on the analysis of the backbone chemical shifts and circular dichroism (CD) spectroscopy. *In vitro*, the linker region affects the ability to acquire heme from metHemoglobin by the C-terminal NEAT domain, $IsdB^{N2}$. The solution structure of the apo- form of $IsdB^{N2}$ is examined in Chapter 5. Structural features of the apo-$IsdB^{N2}$ NMR structure and the published crystal structure of holo-$IsdB^{N2}$ are compared. The final chapter summarizes the findings in this thesis and elaborates on how these findings contribute to our understanding of the mechanism by which IsdB interacts with hemoglobin and abstracts heme.

CHAPTER TWO

MATERIALS AND METHODS

Introduction

In order to explore the structure and function of the various domains of IsdB, IsdB constructs of different lengths were expressed recombinantly in *E. coli*. The recombinant fragments of IsdB are shown in Fig. 2-1. The following chapter describes the cloning and expression of these fragments. The materials and methods used in the determination of the three-dimensional (3D) NMR solution structures of $IsdB^{N1}$ and apo-$IsdB^{N2}$ are discussed, along with NMR experiments of the various IsdB constructs. Methemoglobin binding and heme transfer studies are included as well.

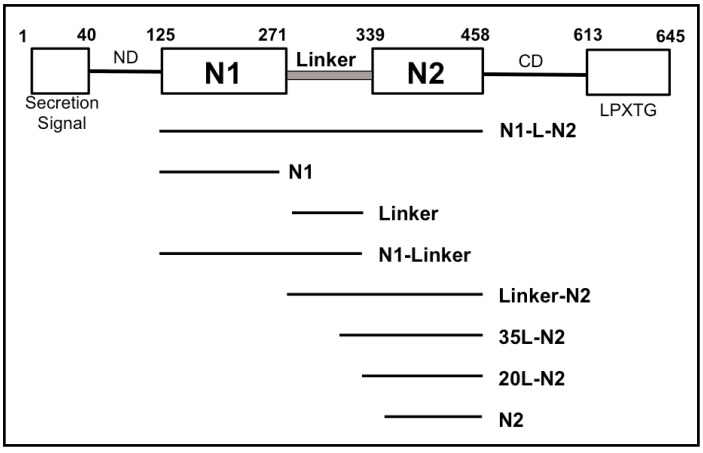


Figure 2-1. Schematic depicting the domain arrangement of IsdB and IsdB fragments. IsdB is 645 amino acids, but the N-terminus (amino acid residues 1-40) secretion signal and the C-terminal tail (amino acid residues 613-645) are cleaved . The N-terminal NEAT domain, $IsdB^{N1}$, and the C-terminal NEAT domain, $IsdB^{N2}$, are connected by a ~70 residue linker region, $IsdB^{Linker}$. The IsdB constructs discussed in this chapter contain one or more of these domains and are depicted above by a line that corresponds to the amino acids they encompass.

Material and Methods for the Structure Determination of $IsdB^{N1}$

Cloning, Protein Expression, and Purification

The gene sequence encoding the first NEAT domain of IsdB (residues Leu 125–Asp 272 of IsdB) was amplified by PCR from a pET-21d plasmid containing the *isdb* gene encoding residues 40-613 of the mature IsdB protein as described in (49). Forward and reverse primers used to amplify the gene sequence coding for $IsdB^{N1}$ were as follows: 5'-GACGACGACAAGATGTTGAATCAGGAACTTAGAGAAGCGAT-3' and 5'-GAGGAGAAGCCCGGTTCAATCTTCTTCAGTTTTGAATTTATCTGCA-3', respectively. The resulting PCR product was gel purified and cloned into the pET-46 Ek/LIC vector (*Novagen*). Sequencing of the cloned fragment, performed at the Iowa State University DNA Facility, confirmed the proper sequence and found no spurious mutations. The resulting $IsdB^{N1}$ protein encoded by the cloned IsdB gene fragment contained an N-terminal 6xHis tag with the following amino acids MAHHHHHHVDDDDKM added at the N-terminus of the native $IsdB^{N1}$ sequence. *Escherichia coli* BL21(DE3) cells transformed with the $IsdB^{N1}$ encoding plasmid were grown overnight into 5 mL of LB medium and saved as glycerol 1 mL cell stocks stored at -80 °C. To produce uniformly labeled ^{15}N, and/or ^{15}N and ^{13}C- labeled $IsdB^{N1}$, a 40 mL starter culture was grown overnight in LB medium containing 100 μg/mL ampicillin at 37 °C. This culture was used to inoculate a 1L culture of M9 minimal media containing 100 μg/mL of ampicillin and supplemented with $^{15}NH_4Cl$ (1.5 g/L), or $^{15}NH_4Cl$ and ^{13}C-labeled glucose (3.0 g/L) (i.e. D-glucose-$^{13}C_6$, 99% ^{13}C-enriched, *Cambridge Isotopes*) as the sole nitrogen and carbon sources respectively, to an initial optical density reading of

0.1 at 600 nm. Protein production was subsequently induced at OD_{600nm} of 0.6 by adding 1 mM IPTG (isopropyl β-D-thiogalactoside) to the resulting cell cultures, cells were allowed to grow for another 8 hours prior to harvest by centrifugation at 4,000xg (*Sorvall, RC-5*), and the resulting cell pellets stored at -20 °C until further use.

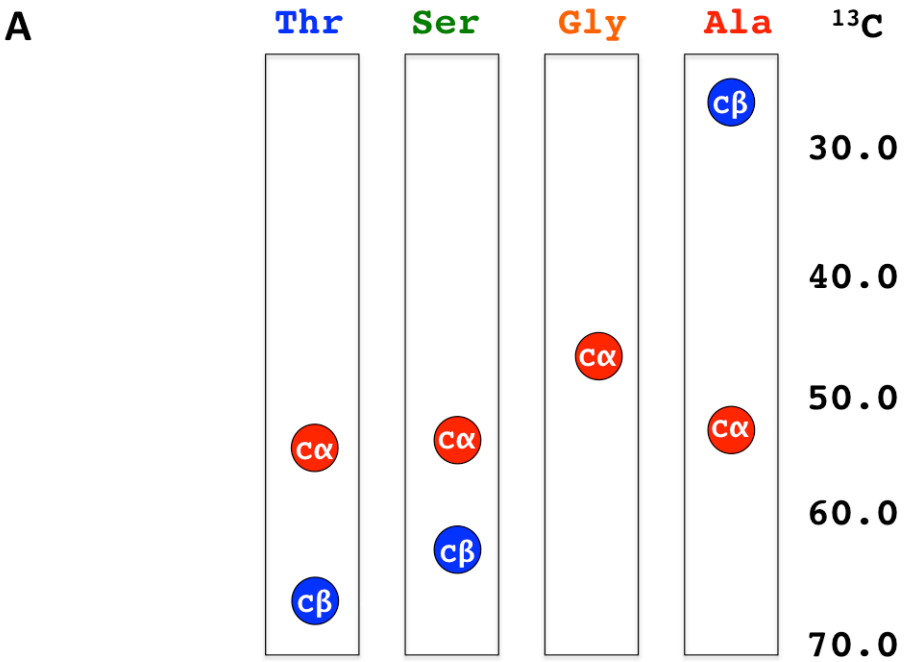
Cells were thawed and resuspended in 5 ml/g of lysis buffer (20 mM Tris, 500 mM NaCl, 50 mM Na_2HPO_4/ NaH_2PO_4, 10 mM imidazole, pH 8) with freshly prepared 0.1 mM PMSF. Cells were then lysed using an M-110L microfluidizer (*Microfluidics*). The resulting cell lysate was kept at 4 °C and clarified from resulting cell debris by centrifugation at 12,000xg for 20 minutes. The resulting supernatant was applied to a nickel affinity chromatography column containing 5 mL bed volume HisPur™ Ni-NTA Resin (*Thermo Scientific*). The column was washed with 2 times the bed volume of lysis buffer and eluted with lysis buffer containing 250 mM imidazole. The $IsdB^{N1}$ protein-containing fractions were pooled and dialyzed against NMR buffer (50 mM Na_2HPO_4/NaH_2PO_4, 400 mM sodium chloride, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, pH 6.8), in either 95 % H_2O/5 % D_2O or 100 % D_2O (for acquisition of 3D ^{13}C-edited ^{1}H-^{1}H TOCSY and NOESY NMR spectra), and concentrated to 1 mM protein concentration as determined by OD_{280nm} readings using an centrifugal filter with 10,000 Daltons molecular weight cutoff (*Millipore*). The purity of the protein (>90 % pure) was determined with SDS-PAGE and the molecular weight of the protein confirmed with mass spectroscopy at the Proteomics and Metabolomics Mass Spectrometry Research Facility of Montana State University. The monomeric state of $IsdB^{N1}$ was confirmed using size exclusion chromatography.

Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectra of 1mM $IsdB^{N1}$ protein samples in 50mM sodium phosphate, 400mM sodium chloride, 1mM EDTA, 0.1mM PMSF, 0.01% sodium azide, pH 6.8, in either 95% H_2O/5% D_2O or 100% D_2O (for acquisition of 3D ^{13}C-edited ^{1}H-^{1}H TOCSY and NOESY NMR spectra) were acquired on a four-channel Bruker AVANCE III 600 spectrometer equipped with a TCI NMR cryoprobe^{TM} at 298K (25 °C). Backbone and side chain resonance assignments were extracted from the standard heteronuclear (^{1}H, ^{15}N, ^{13}C) multidimensional NMR experiments: HNCA (62), HNCACB (63), CBCA(CO)NH (64), C(CO)NH (65), HBHA(CO)NH (66), HCC(CO)NH (65), HN-TOCSY (67),^{13}C-edited ^{1}H-^{1}H TOCSY (68), ^{13}C-^{1}H HSQC (69), and ^{13}C-edited and ^{15}N-edited ^{1}H-^{1}H NOESY (70) as reported in (71) see Appendix A. ^{1}H, ^{13}C, and ^{15}N chemical shifts were indirectly referenced to DSS. In addition, hydrogen/deuterium ($^{1}H/^{2}H$) solvent exchange experiments were performed using 500µL of a 1 mM ^{15}N-labeled protein sample of $IsdB^{N1}$ previously lyophilized and resuspended in 500uL of 100% D_2O. Acquisition of a series of 2D ^{1}H-^{15}N HSQC spectra at subsequent time points immediately followed the resolubilization of protein into D_2O. To assure that the protein structure of $IsdB^{N1}$ was unaffected by the lyophilization process, a control experiment was conducted whereby a 2D ^{1}H-^{15}N HSQC spectrum of the protein was recorded following lyophilization and resolubilization in H_2O and was found to be identical to the 2D ^{1}H–^{15}N HSQC spectrum recorded of unlyophilized $IsdB^{N1}$. Additionally, a 2D CLEANEX-PM NMR experiment (72) was performed to detect amide protons participating in fast chemical exchange with water. All NMR data were processed and

analyzed with NMRPipe (73) and Sparky (74) software.

The 2D ^{1}H-^{15}N Heteronuclear Quantum Correlation (HSQC) experiment provides a "fingerprint" of the protein backbone as each N-H pair gives rise to a resonance except for proline. This spectrum is used as a chemical shift map that correlates the N-H resonances to its backbone and side chain carbon atoms and protons. The two 3D heteronuclear experiments that are primarily used to sequentially assign the backbone resonances to their corresponding amino acid in the protein are the CBCA(CO)NH and the HNCACB. In the CBCA(CO)NH experiment, an NH group is correlated to the $^{13}C\alpha$ and $^{13}C\beta$ of the preceding residue. The HNCACB experiment correlates an NH group to the $^{13}C\alpha$ and $^{13}C\beta$ of itself and the proceeding residue. Magnetization is transferred from the $^{1}H\beta$ and $^{1}H\alpha$ to the $^{13}C\beta$ and $^{13}C\alpha$, through the carbonyl to the $^{15}N_H$ and finally to $^{1}H_N$ for detection. Using these two experiments, stretches of amide resonances were matched using the chemical shifts of Cα and Cβ atoms to interresidue Cα and Cβ chemical shifts. The Cα and Cβ chemical shifts were assigned an amino acid type based on empirically determined Cα and Cβ chemical shift ranges (75). Threonine and serine residues have Cβ resonances shifted downfield of their Cα resonances thereby distinguishing them from all other amino acids (Fig. 2-2A). Glycine exhibits upfield-shifted Cα resonances at 43-46 ppm and possesses no Cβ, therefore making it another characteristic amino acid. Other unique amino acid spin systems are A, I, V, and P. Walking through the backbone of the protein and connecting the amide resonances to sequential Cα and Cβ chemical shifts, along with identification of amino acid type, allowed these sequential sequences to be matched to the primary sequence of $IsdB^{N1}$ (Fig. 2-2B and Fig. 2-3).



B

```
125 LNQELREAIK NPAIKDKDHS APNSRPIDFE MKKKDGTQQF YHYASSVKPA
175 RVIFTDSKPE IELGLQSGQF WRKFEVYEGD KKLPIKLVSY DTVKDYAYIR
225 FSVSNGTKAV KIVSSTHFNN KEEKYDYTLM EFAQPIYNSA DKFKTEED
```

Figure 2-2. Amino acids with characteristic Cα and Cβ chemical shifts. (A) The unique chemical shift patterns for the Cα and Cβ atoms of threonine, serine, glycine, and alanine. (B) The primary sequence of $IsdB^{N1}$. The amino acids with unique chemical shift patterns are highlighted. These unique amino acids were used to help in the sequential assignment of the backbone atoms of $IsdB^{N1}$.

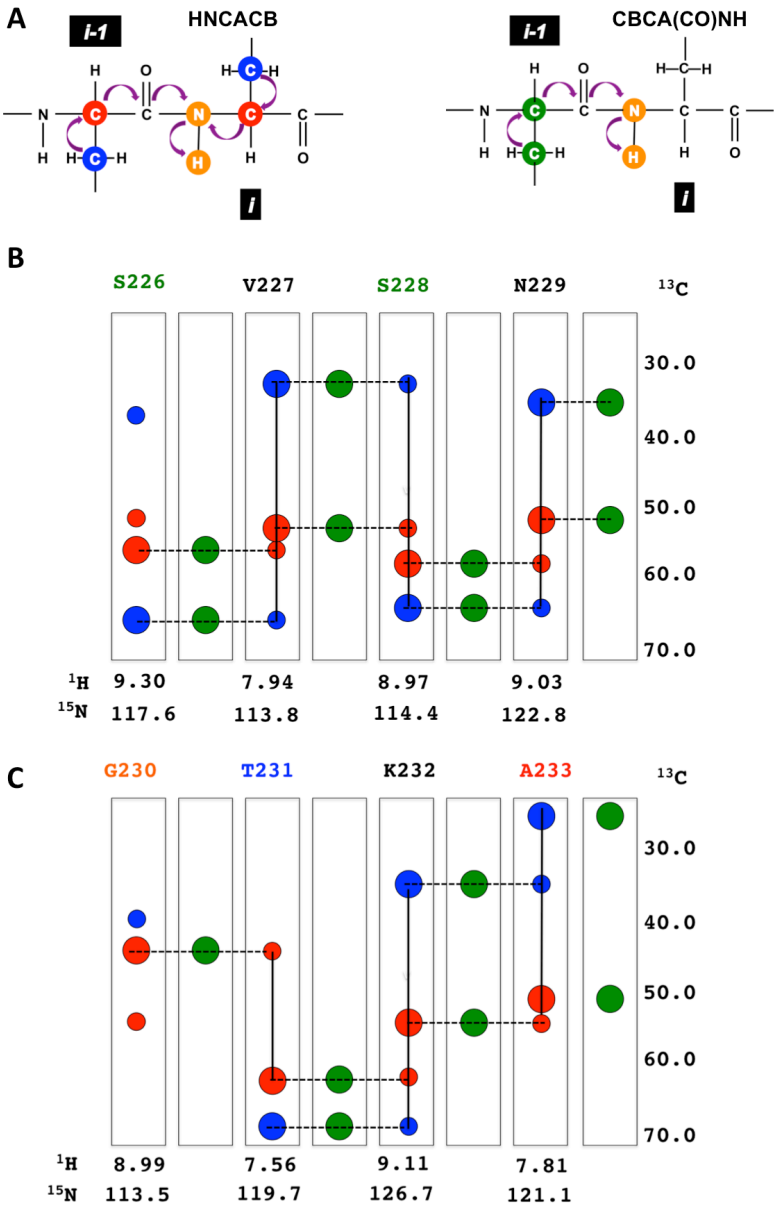


Figure 2-3. Example of sequential backbone assignment for IsdB^{N1} using the HNCACB and CBCA(CO)NH experiments. (A) The magnetization transfer between the residue of interest (*i*) and the preceding residue (*i*-1) for the HNCACB and CBCA(CO)NH experiments. (B) Sequential assignment of residues S226 to N229 and (C) residues G230 and to A233 using the data acquired from the HNCACB experiment where the large circles represent the resonances that belong to *i* and the smaller circles represent the resonances that belong to *i-1* (Cα resonances are in red and the Cβ resonances are in blue) and the CBCA(CO)NH experiment (both Cα and Cβ resonances in green). Each strip represents a slice of the 3D experiment at the corresponding ^{1}H and ^{15}N chemical shift coordinates.

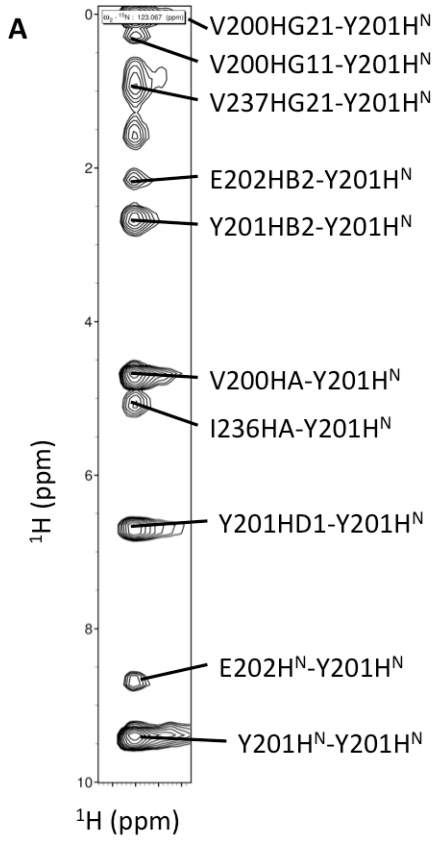
Once the initial stretches of the backbone resonances were assigned, the side chain carbons and protons were assigned (Table 2-1). The C(CO)NH correlates the side chain carbons with the NH of the succeeding residue. In order to assign the proton resonances, the HBHA(CO)NH and HC(CO)NH correlate the Hβ and Hα and the side chain protons to the NH of the succeeding residue, respectively. Other useful experiments in assigning protons are the HN-TOCSY and the HCCH-TOCSY. Proton assignment is important for the subsequent assignment of NOESY data, which is needed for the protein structure calculations.

Table 2-1. Spectra acquired for backbone and side chain assignments of $IsdB^{N1}$

Experiment	**Result**
HNCACB	Cβ and Cα chemical shifts of ***i*** and ***i-1*** correlated to NH of ***i***
CBCA(CO)NH	Cβ and Cα chemical shifts of ***i-1*** correlated to NH of ***i***
C(CO)NH	> Cβ chemical shifts of ***i-1*** correlated to NH of ***i***
HBHA(CO)NH	Hβ and Hα chemical shifts of ***i-1*** correlated to NH of ***i***
HC(CO)NH	> Hβ chemical shifts of ***i-1*** correlated to NH of ***i***
HN-TOCSY	Side chain hydrogens of ***i*** are correlated to the NH of ***i***
HCCH-TOCSY	Side chain hydrogens and their attached carbons correlated to the hydrogen and carbon resonances of other C-H pairs in the side chain

The nuclear Overhauser effect (NOE) is based on dipolar couplings between spin systems that are close in space. Cross peaks in a NOESY spectrum are generated for all proton pairs that are less than 5Å in real space. The intensity of the resonance is related to the physical distance by the 6^{th} power ($NOE \sim 1/r^6$). This intensity is translated into a

range of proton-proton distances for the protein structure calculations (Fig. 2-4). Interproton NOE assignments were determined from analysis of 3D ^{15}N-edited ^{1}H-^{1}H NOESY (τ_m = 100 ms) and 3D ^{13}C-edited ^{1}H-^{1}H NOESY (τ_m = 100 ms) spectra.



B

Res # Atom	Res # Atom	Upper Limit (Å)
201 H	202 H	5.47
201 H	201 QD1	4.64
201 H	236 HA	5.33
201 H	200 HA	No Restriction
201 H	201 QB	No Restriction
201 H	202 QB	5.23
201 H	237 QQG	5.50
201 H	200 QG1	5.38
201 H	200 QG2	5.38

Figure 2-4. Proton-proton distance restraints based on the NOESY assignment for IsdB^{N1}. (A) A strip from the ^{15}N-edited NOESY at 123.07 ppm $^{15}N_H$ resonance and 9.40 ppm $^{1}H_N$ resonance with NOE resonance assignments labeled for the HN of Y201. (B) The NOE assignments from A translated into proton-proton distance restraints by CYANA 2.1 for use in the structure calculations of IsdB^{N1}.

Structure Calculation and Refinement

To aid in the assignment of NOE resonances, a combination of manual inspection and automated NOE assignment was performed with UNIO'10 Version 2.0.2 (76,77).

Hydrogen bond distance restraints were determined from the NMR data acquired by the $^{1}H/^{2}H$ solvent exchange experiments and protons not involved in hydrogen bonding were further confirmed by the results of the CLEANEX experiment. Two distance restraints between the amide and carbonyl group atoms (NH-O = 1.8-2.8 Å, N-O =2.5-3.8 Å) were used for the hydrogen bond restraints. Backbone Φ and Ψ dihedral angles constraints were generated by TALOS+ (78) based on the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ chemical shifts. Structures of $IsdB^{N1}$ were calculated with CYANA 2.1 using CYANA's standard simulated annealing protocol (see example script in Appendix C) (77,79). The standard simulated annealing protocol utilized in CYANA is torsion angle molecular dynamics simulation. Unlike Cartesian coordinate molecular dynamic simulation, bond lengths and bond angles are fixed at optimal values and only torsion angles are allowed to move. The standard simulated (SA) annealing procedure is comprised of five steps including an initial minimization, the first SA stage, the second SA stage, a low temperature phase, and a final minimization. The initial minimization where the first 100 steps use distance restraints between atoms up to 3 residues within the sequence followed by another 100 steps using all of the restraints. Initial calculations by CYANA derived proton-proton distance restraints by calibrating the peak heights of NOE data derived from the NOESY spectra. From this initial set of structure calculations, a distance restraint file, all.upl was generated and used in subsequent structure calculations (Appendix B).

For the final set of calculations, 200 initial structures of $IsdB^{N1}$ were generated using CYANA. From this initial ensemble, 40 $IsdB^{N1}$ structures with the lowest residual target functions were selected for further refinement in AMBER9 (Assisted Model

Building with Energy Refinement) (80) using the AMPS-NMR (AMBER-based Portal Server for NMR structures) web portal (81), and a standard restrained molecular dynamics protocol (restrainedMD.xml) implemented within the AMBER99SB force field, a generalized Born model, and a 10Å TIP3P water box (81). The NMR restraints remain active for the entire restrainedMD.xml protocol. The six phases of restrainedMD.xml are shown in Figure 2-5. In brief, the 10Å hydration shell is minimized, followed by a minimization of the entire system. The system is heated to 300K in 10,000 restrained molecular dynamic steps, and then remains at constant temperature and pressure (1.0 bar) for 15,000 steps (30.0 ps of 2.0fs per step). The system cools down to 0K in 10,000 steps and followed by a final energy minimization. From this set of 40, twenty $IsdB^{N1}$ structures with lowest conformational energy and no experimental distance and angular constraints violations were selected to generate a final ensemble of $IsdB^{N1}$ protein structures. The overall quality of the final set of protein structures was analyzed with PROCHECK-NMR (82), MolProbity (83), Verify3D (84), and the protein structure validation suite (PSVS) software (85). The final ensemble of $IsdB^{N1}$ is deposited in the Protein Data Bank as entry 2MOQ. Figures were prepared with PYMOL (86) and structural homology identified with the DALI web server (87).

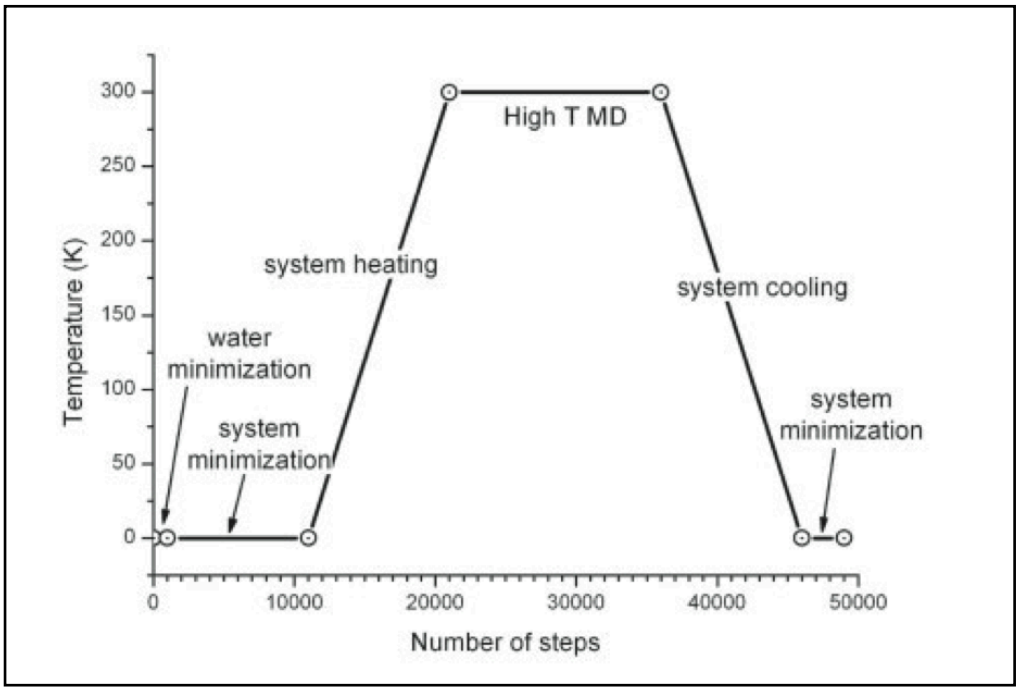


Figure 2-5. Schematic of the steps for the restrained molecular dynamics protocol in AMBER9. The restrainedMD.xml protocol was used to further refine the top 40 calculated structures of IsdB^{N1} from CYANA 2.1.

Hemoglobin-binding Studies of IsdB^{N1}

Site-directed Mutagenesis for Hemoglobin-binding Studies

DNA vectors encoding protein variants of IsdB^{N1} containing the mutations of Phe164 to an aspartic acid and/or Tyr167 to an aspartic acid (IsdB^{N1}-F164D, IsdB^{N1}-Y167D, and IsdB^{N1}-F164D/Y167D) were generated using a QuikChange Lightening (*Agilent*) site directed mutagenesis kit, cloned into pET46 plasmids, and the resulting DNA sequences confirmed by sequencing. Protein expression and purification of ^{15}N-labeled IsdB^{N1} protein variants was conducted as described above for recombinant IsdB^{N1} wild type (WT) protein.

Hemoglobin-binding Studies

NMR studies of ^{15}N-labeled IsdB^{N1} binding to human methemoglobin were performed at 298K (25 °C) in NMR buffer (50mM sodium phosphate, 100mM sodium chloride, 1mM EDTA, 0.1mM PMSF, 0.01% sodium azide at pH 6.8 in 95% H_2O/5 % D_2O). An 8 mg sample of lyophilized human hemoglobin (*Sigma*) was dissolved in 10 mL of NMR buffer and further dialyzed into NMR buffer overnight with two buffer changes. The human metHb stock solution was subsequently concentrated using a 10,000 molecular weight cutoff centrifugal spin concentrator to a concentration of 0.5 mM. To produce the metHb:IsdB^{N1} NMR samples, 500 μL of the 0.5 mM stock solution of human metHb was mixed with 500 μL of a 0.5 mM stock solution of ^{15}N-labeled IsdB^{N1} in NMR buffer. The sample was concentrated to 500 μL volume and the ^{1}H-^{15}N HSQC spectrum was recorded to monitor any changes in chemical shifts. The ^{1}H-^{15}N HSQC spectra were recorded at metHb (αβ dimer):IsdB^{N1} ratios of 1:0.25, 1:0.5, 1:1, 1:2, and 1:4. The titration experiment was repeated as stated above for all three of the ^{15}N-labeled IsdB^{N1} mutants: IsdB^{N1}-F164D, IsdB^{N1}-Y167D and IsdB^{N1}-F164D/Y167D. A competitive assay was conducted with unlabeled IsdB^{N1} was titrated in excess to a 1:1 metHb:^{15}N-IsdB^{N1} sample. The increased signal intensities of displaced ^{15}N-labeled Hb-bound IsdB^{N1} was monitored.

Heme Transfer Studies of IsdB Constructs

Six additional IsdB protein constructs encompassing various domains of the protein were generated to test their heme transfer capabilities upon hemoglobin binding to IsdB^{N1}. DNA vectors encoding IsdB amino acid stretch Leu125-Asn458 (referred to as

$IsdB^{N1LN2}$), Leu125-Asp272 ($IsdB^{N1\text{-}L}$), Glu271-Asn458 ($IsdB^{LinkerN2}$), Glu304-Asn458 ($IsdB^{35LN2}$), Lys319-Asn458 ($IsdB^{20LN2}$), and Asn339-Asn458 ($IsdB^{N2}$) were generated by PCR and amplified from a pET-21d plasmid containing the *isdb* gene encoding residues 40-613 of full-length IsdB as described in (49). Resulting PCR products were gel purified and ligated into pET-46 Ek/LIC (*Novagen*) plasmids similar to what was done for $IsdB^{N1}$ (88). In addition, the $IsdB^{N1}$ mutations were generated in the $IsdB^{N1LN2}$ construct. DNA vectors encoding the $IsdB^{N1LN2}$-F164D, $IsdB^{N1LN2}$-F164A, $IsdB^{N1LN2}$-Y167D, and $IsdB^{N1LN2}$-F164D/Y167D protein variants were generated using the QuikChange Lightening (*Quigen*) site directed mutagenesis kit and resulting DNA sequences verified by Sanger sequencing. $IsdB^{N1LN2}$ protein variants were expressed and purified as described for wild type $IsdB^{N1}$ (88). In brief, *E. coli* BL21(DE3) cells transformed with various IsdB protein construct-encoding pET46 plasmids were grown overnight into 5 mL of LB medium and resulting 1mL cell stocks stored in glycerol at -80 °C. For the heme transfer experiments, a small 40 mL starter culture was grown overnight in LB at 37°C and used to inoculate a 1L culture of LB medium to an initial $OD_{600\ nm}$ reading of 0.1. At an $OD_{600\ nm}$ reading of 0.6, 1mM IPTG was added to induce protein expression. Resulting cells were allowed to grow for an additional 8 hours prior to harvest via centrifugation at 4,000 x g (Sorvall, RC-5), and resulting cell pellets stored at -20 °C until further use.

Cells were thawed and resuspended in 5 ml per gram of cells of lysis buffer (20 mM Tris, 500 mM NaCl, 50 mM Na_2HPO_4/NaH_2PO_4, 10 mM imidazole, pH 8) with freshly prepared 0.1 mM PMSF. Cells were then lyzed using an M-110L microfluidizer

instrument (*Microfluidics*). Resulting cell lysates were kept at 4 °C and clarified from resulting cell debris by centrifugation at 12,000 x g for 20 minutes, and resulting supernatant applied to a nickel affinity chromatography column containing 5mL bed volume of HisPur™ Ni-NTA Resin (*Thermo Scientific*). The column was washed with 2x the bed volume of lysis buffer and eluted with lysis buffer containing 250mM imidazole. Protein-containing fractions were pooled and dialysed against 20mM Tris-HCl buffer, pH 8.0. Protein concentration was established by measuring absorbance at 280 nm (OD_{280nm}), protein purity assessed by SDS-PAGE, and protein molecular mass confirmed with mass spectrometry using MS instruments of the Proteomics and Metabolomics Mass Spectrometry Facility of Montana State University.

Preparation of Apo-IsdB Constructs

Many of the IsdB constructs are able to bind heme from *E.coli*, so in order to have apo- protein for the heme transfer experiments, the heme moiety was removed from IsdB constructs by the methyl ethyl ketone extraction method (89). Briefly, a 5 mL sample of 0.2 mg/mL of protein in 20mM Tris-HCl buffer, pH 8, was adjusted to a pH of ~2.8 with 1 M HCl. An equal volume (5 mL) of methyl ethyl ketone was added to the protein solution and vortexed. After the organic and aqueous phases separated, the aqueous phase was dialyzed against 6 L of water overnight with two exchanges. The proteins were then dialyzed into the heme transfer buffer, 20mM sodium phosphate at pH 7.5 and 150mM sodium chloride.

Heme Transfer Reactions

Heme transfer reactions from holo-metHb to $IsdB^{N1LN2}$ and other IsdB constructs were performed on a conventional UV/Vis spectrometer (*Thermo Evolution 60s*) at 25°C by monitoring spectral changes as previously described (49). In brief, 4μM holo-metHb (*Sigma*) was mixed with 25μM apo- acceptor IsdB protein constructs in a total volume of 120 μL of 20mM sodium phosphate, 150 mM sodium chloride, at pH 7.5. The absorbance changes of the Soret band at 406nm were recorded for each reaction over a course of 5 minutes and repeated in triplicate. The time courses of the absorbance changes were plotted and analyzed using Prism (*Graphpad*) software with a single phase exponential expression to determine the relative rates of heme transfer.

Materials and Methods for the NMR Investigation of the Linker Domain

Cloning, Protein Expression, and Purification

A pET-21d plasmid encoding the gene sequence of the linker region of IsdB (amino acid residues Glu271 to Thr338) was obtained from the laboratory of Dr. Lei. Isotopically labeled $IsdB^{Linker}$ protein was expressed recombinantly in *E. coli* BL21(DE3) cells as described previously for $IsdB^{N1}$.

The cell pellets were resuspended in 5 mL/g of lysis buffer (20mM Tris-HCl, pH 8.0) with freshly prepared 0.1mM PMSF and lysed using an M-110L microfluidizer (*Microfluidics*). The resulting cell lysate was kept at 4 °C and clarified from resulting cell debris by centrifugation at 12,000 xg for 20 minutes. The resulting supernatant was applied to a 5mL HiTrap SP sepharose Fast Flow ion-exchange column (*GE Healthcare*). The column was washed with 3x the column volume of lysis buffer and the protein was

eluded with a stepwise gradient of 50mM to 200mM NaCl in 20mM Tris-HCl, pH 8.0. The fractions containing protein were pooled and ammonium sulfate ($(NH_4)_2SO_4$) was added to a concentration of 2M. The protein was then applied to a 5mL HiTrap Phenyl Fast Flow (High Sub) affinity column (*GE Healthcare*) and washed with 20mM Tris, 2M $(NH_4)_2SO_4$, pH 8.0 and then eluded with a stepwise gradient of 1.5M to 0M $(NH_4)_2SO_4$, in 20mM Tris-HCl, pH 8.0. The protein containing fractions were pooled and dialyzed against NMR buffer (50 mM K_2HPO_4/KH_2PO_4, 100 mM potassium chloride, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, pH 6.8), in either 95 % H_2O/5 % D_2O or 100 % D_2O (for acquisition of 3D ^{13}C-edited ^{1}H-^{1}H TOCSY and NOESY NMR spectra), and concentrated to 2.4 mM protein concentration as determined by OD_{280nm} readings using an centrifugal filter with 3,500 Daltons molecular weight cutoff (*Millipore*). The purity of the protein (>95 % pure) was determined with SDS-PAGE and the molecular weight of the protein confirmed with mass spectroscopy at the Proteomics and Metabolomics Mass Spectrometry Research Facility of Montana State University.

NMR Spectroscopy

NMR spectra of 2.4mM protein samples of $IsdB^{Linker}$ in 50mM potassium phosphate, 100mM potassium chloride, 1mM EDTA, 0.1mM PMSF, 0.01% sodium azide, pH 6.8, in either 95% H_2O/5% D_2O or 100% D_2O were acquired at 292K (19 °C) on a four-channel Bruker AVANCE III 600 spectrometer with an inverse detection triple resonance (^{15}N, ^{13}C, ^{1}H) conventional NMR probe equipped with triple axis gradients. Backbone and side chain resonance assignments were extracted from the standard heteronuclear (^{1}H, ^{15}N, ^{13}C) multidimensional NMR experiments as previously described

for $IsdB^{N1}$.

Circular Dichroism (CD)

To determine the secondary structure of $IsdB^{Linker}$, circular dichroism (CD) spectra were measured on a JASCO J-710 spectropolarimeter. A 180 μM (1.5mg/mL) sample of $IsdB^{Linker}$ protein was scanned from 190-250 nm (50mM $KHPO_4$, 100mM KCl, pH 6.8) at a rate of 100 nm/min at 25°C in a 0.1 mm path length cell.

NMR Investigation of the Linker Region with the N1 and N2 Domains of IsdB

Material and Methods for $IsdB^{N2}$

The pET-21d plasmids encoding the gene sequence of the N2 domain of IsdB (amino acid residues Asn339 to Asn458) and the linker with the N2 domain (amino acid residues Glu271 to Asn458) were obtained from the laboratory of Dr. Lei. The $IsdB^{N2}$ and $IsdB^{Linker-N2}$ proteins were expressed recombinantly as described above. The pellets were resuspended and lysed by an M-110L microfluidizer in 20mM Tris-HCl buffer, pH 8.0 with 0.1mM PMSF. The lysed cells were centrifuged at 12,000xg to remove cell debris and the supernatant was applied to a 3x5mL stacked HiTrap SP sepharose Fast Flow ion-exchange column (*GE Healthcare*). The column was washed with lysis buffer and the protein was eluded with a stepwise gradient of 50mM to 300mM NaCl in 20mM Tris-HCl, pH 8.0. The fractions containing protein were pooled and ammonium sulfate ($(NH_4)_2SO_4$) was added to a concentration of 2M. The protein was then applied to a 5mL HiTrap Phenyl Fast Flow (High Sub) affinity column (*GE Healthcare*) and washed with

20mM Tris, 2M $(NH_4)_2SO_4$, pH 8.0 and then eluded with a stepwise gradient of 1.5M to 0M $(NH_4)_2SO_4$, in 20mM Tris-HCl, pH 8.0. The protein containing fractions were pooled and dialyzed against NMR buffer (50 mM K_2HPO_4/KH_2PO_4, 100 mM potassium chloride, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, pH 6.8), in either 95 % H_2O/5 % D_2O

Material and Methods for $IsdB^{N1\text{-}Linker}$

A DNA vectors encoding IsdB amino acid stretch Leu125-Thr338 (referred to as $IsdB^{N1\text{-}Linker}$) was generated by PCR and amplified from a pET-21d plasmid containing the *isdb* gene encoding residues 40-613 of full-length IsdB as described in (49). The resulting PCR product was gel purified and ligated into pET-46 Ek/LIC (*Novagen*) plasmid similar to what was done for $IsdB^{N1}$ (88). Protein expression and purification of $IsdB^{N1\text{-}Linker}$ was performed as described above for $IsdB^{N1}$.

NMR Experiments of IsdB Linker Constructs

$^{1}H\text{-}^{15}N$ HSQC NMR spectra of ^{15}N-labeled $IsdB^{Linker}$, $IsdB^{N2}$, and $IsdB^{LinkerN2}$ were performed at 298K (25 °C) in NMR buffer (50mM potassium phosphate, 100mM potassium chloride, 1mM EDTA, 0.1mM PMSF, 0.01% sodium azide at pH 6.8 in 95% H_2O/5 % D_2O) with a protein concentration of 0.28 mM.

Materials and Methods for the Structure Determination of apo-NEAT 2

Initially, the structure determination of apo-$IsdB^{N2}$ was performed on the wild type sequence. However, initial NMR studies showed an issue with soluble aggregates.

The backbone amide resonances in the ^{1}H-^{15}N HSQC appeared intense, however in further investigations of the ^{15}N and ^{13}C labeled samples, the NMR signals were broadened or missing. The hypothesis was that there were hydrophobic residues that were not buried and were capable of making contacts between domains. The first mutation was Ile438 to an alanine, which still displayed signs of aggregation. and it still aggregated. An additional mutant was incorporated into the $IsdB^{N2}$ I438A protein variant, Met 362 to alanine or Val356 to a glycine. Both the $IsdB^{N2}$ M362A/I438A and $IsdB^{N2}$ V356G/I438A were equivalent in NMR signal intensity pointing to the protein not aggregating. The following describes the NMR structure determination of apo-$IsdB^{N2}$ M362A/I438A.

Cloning, Protein Expression, and Purification

The gene sequence encoding the second NEAT domain of IsdB (residues Asn 338 to Asn 458 of IsdB) was amplified by PCR from a pET-21d plasmid containing the *isdb* gene encoding residues 40-613 of the mature IsdB protein as described in (49). Forward and reverse primers used to amplify the gene sequence coding for $IsdB^{N2}$ were as follows: 5'-GACGACGACAAGATGAATGAAAAAATGACTGATTTACAAGA-3' and 5'-GAGGAGAAGCCCGGTCTAATTGGCTTTTGTAAATGCTTC-3', respectively. The resulting PCR product was gel purified and cloned into the pET-46 Ek/LIC vector (*Novagen*). Sequencing of the cloned fragment, performed at the Iowa State University DNA Facility, confirmed the proper sequence and found no spurious mutations. The DNA vector encoding the $IsdB^{N2}$ V356G /I438A protein variant was generated using a QuikChange Lightening (*Agilent*) site directed mutagenesis kit, cloned into pET46 plasmid, and resulting DNA sequences confirmed by sequencing. The

resulting $IsdB^{N2}$ protein encoded by the cloned IsdB gene fragment contained an N-terminal 6xHis tag with the following amino acids MAHHHHHHVDDDDKM added at the N-terminus of the native $IsdB^{N2}$ sequence. Isotopically labeled ^{15}N and/or ^{15}N and ^{13}C $IsdB^{N2}$ V356G/I438A protein was expressed and purified by the same method as described previously for $IsdB^{N1}$ with the addition of 8 M urea to the lysis buffer. The protein was refolded on the nickel-nta column by washing and eluting with buffer that did not contain urea.

NMR Spectroscopy

NMR spectra of 1mM $IsdB^{N2}$ V356G /I438A protein samples in 50mM sodium phosphate, 1mM EDTA, 0.1mM PMSF, 0.01% sodium azide, pH 6.8, in either 95% H_2O/5% D_2O or 100% D_2O (for acquisition of 3D ^{13}C-edited ^{1}H-^{1}H TOCSY and NOESY NMR spectra) were acquired on a four-channel Bruker AVANCE III 600 spectrometer equipped with a TCI NMR cryoprobe^{TM} at 306K (33 °C). Backbone and side chain resonance assignments were extracted from the standard heteronuclear (^{1}H, ^{15}N, ^{13}C) multidimensional NMR experiments as described above for $IsdB^{N1}$.

Structure Calculation

Interproton NOE assignments were determined from analysis of 3D ^{15}N-edited ^{1}H-^{1}H NOESY (τ_m = 100 ms) and 3D ^{13}C-edited ^{1}H-^{1}H NOESY (τ_m = 100 ms) spectra. Hydrogen bond restraints were determined from the NMR data acquired by the ^{1}H-^{2}H solvent exchange experiments and protons not involved in hydrogen bonding were further confirmed by the results of the CLEANEX experiment. Backbone Φ and Ψ

dihedral angles constraints were generated by TALOS+ (78) based on the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ chemical shifts. For the final set of calculations, 200 initial structures of $IsdB^{N2}$ were generated with CYANA 2.1 using CYANA's standard simulated annealing protocol (77,79).

CHAPTER THREE

STRUCTURAL AND FUNCTIONAL STUDIES OF ISDB^{N1}

Introduction

S. aureus expresses a mature IsdB protein possessing 645 amino acids. Residues 1-40 include the secretion signal for protein export through the cytoplasmic membrane via the secretory (sec) pathway (28). The C-terminal tail contains primarily positively charged residues, preceded by a segment of hydrophobic residues and the LPXTG motif (where X is any amino acid) (28). Sortase A (SrtA) cleaves the C-terminal tail in the LPXTG motif, attaching Thr613 to the crosslinking glycine of the peptidoglycan cell wall (28). The resulting IsdB protein attached to the cell wall encompasses residues Ala41 through Thr613, and, due to its size, would be challenging to solve by traditional heteronuclear nuclear magnetic resonance (NMR) spectroscopy. IsdB contains two near iron transporter (NEAT) domains connected by a ~70 residue linker region. Previous kinetic and thermodynamic studies by Lei and colleagues (Zhu, et. al. 2014, manuscript under review) indicated that the NEAT domains of IsdB performed different functions, but worked together to abstract heme from methemoglobin. In order to gain insight into how the individual domains function, structure determination of the individual NEAT domains and the linker region by NMR was employed to characterize recombinant IsdB domains. This chapter examines the solved NMR solution structure of the N-terminal NEAT domain, IsdB^{N1}, its interactions with metHb, and its role in heme acquisition from metHb by IsdB.

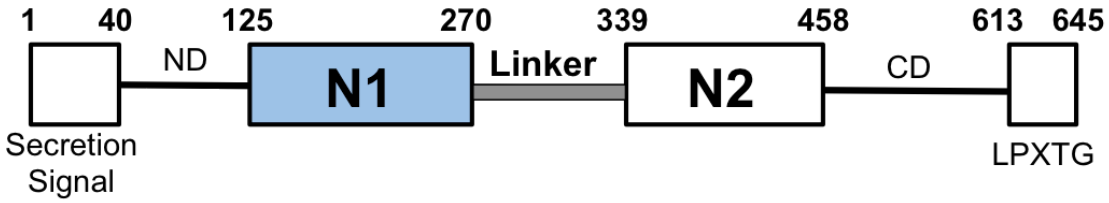


Figure 3-1. A schematic of IsdB and its domain arrangement according to residue number. IsdB contains five domains after the N-terminal secretion signal is cleaved and the C-terminus is covalently attached to the cell wall at Thr613. The N-terminal NEAT domain is labeled as N1 and encompasses residues 125-270.

Characterization of $IsdB^{N1}$

To gain insight into the interaction of IsdB with methemoglobin, the structure of $IsdB^{N1}$ was undertaken by nuclear magnetic resonance (NMR) spectroscopy. The initial inspection of the amino acid sequence of IsdB conducted by Andrade and colleagues (55) revealed the presence of two NEAr iron Transporter (NEAT) domains, approximately 120 amino acid residues each. A pET-21d plasmid containing the gene sequence for the N-terminal NEAT domain, IsdB residues Ala140 to Glu270, was obtained from Dr. Lei's group. Preliminary protein expression of $IsdB^{N1}$ in *E. coli*, however, produced low yields of soluble protein, with the majority of $IsdB^{N1}$ located in the inclusion bodies. Following cell lysis with 8M urea and protein purification via ion-exchange chromatography, the protein exhibited signs of aggregation at the millimolar concentration range needed for multidimensional NMR experiments. The protein construct of $IsdB^{N1}$, residues 140 to 270, was either unstable or not refolded properly.

Further inspection of the amino acid sequence and predicted secondary structure performed by GOR IV secondary structure prediction method of IsdB suggested the presence of a small α-helix, approximately 10 residues upstream of Ala140. To determine

if this secondary structural element was important for $IsdB^{N1}$ stability, the gene sequence encoding IsdB residues Leu125 to Asp272 was cloned into a pET-46 vector and the resulting protein contained an N-terminal His_6 tag for Ni-NTA purification. The new $IsdB^{N1}$ construct produced high yields of soluble monomeric protein in minimal media and the purified protein could be concentrated to 1 mM in NMR buffer.

Backbone and Side Chain Resonance Assignments

The ^{1}H-^{15}N HSQC spectrum correlates the amide proton and the nitrogen of the protein backbone and ideally a well-folded protein should exhibit one peak for each bonded N-H pair. The initial ^{1}H-^{15}N HSQC spectrum of $IsdB^{N1}$ displayed very large dispersion of the amide proton chemical shifts ($\Delta^{1}H_N$ of ~4.5ppm) consistent with a well-folded structure rich in β-strands. Although most of the peaks in an ^{1}H-^{15}N HSQC correspond to the nitrogen and the amide proton in the protein backbone, the NH_2 side chains of glutamine and asparagine will appear as doublets in the top right corner of the spectra, and the indole NH of tryptophan appears around 10.5 ppm. The NMR data collected on the recombinant $IsdB^{N1}$ protein enabled resonance assignments of 93 % all assignable $^{1}H_N$ and ^{15}N resonances (132 out 142, and excluding all 7 prolines of the protein sequence, which do not contain an amide proton) (Fig. 3-2). Ten backbone amide resonances could not be assigned due to significant resonance broadening or extensive spectral overlap. These residues could be in more dynamic regions of the protein. Four of the missing assignments correspond to aromatic residues F164, Y165, H166, and Y167. Based on sequence homology, these residues are in equivalent positions to a stretch of

contiguous aromatic residues (Y125, Y126, H127, F128, F129) of $IsdH^{N1}$ (a close homolog to $IsdB^{N1}$) that are shown to be located in a disordered region and also unassigned in the NMR solution structure of $IsdH^{N1}$ (57). The completeness of the remaining backbone resonance assignments corresponded to: 97 % of $^{13}C_{\alpha}$ (144/149), 95 % of $^{1}H_{\alpha}$ (142/149), and 97 % of $^{13}C_{\beta}$ (139/ 144). Side chain resonances were assigned to greater than 87 % completeness. Several aromatic side-chain proton resonances were assigned from the ^{15}N-edited NOESY HSQC experiment.

The secondary structure of $IsdB^{N1}$ was elucidated from the chemical shift data of the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ by the chemical shift index (CSI) method (90) and by TALOS+. The H_{α} and ^{13}C chemical shifts of the protein backbone are directly influenced by the secondary structure they have adopted, and, therefore, those shifts deviate either upfield or downfield from amino acids in a random coil conformation. The CSI method calculates independent predictions based on the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ chemical shifts, and then combines that data to form a consensus. The TALOS+ (78) program predicts the backbone ϕ and ψ dihedral angles based on the chemical shifts of each residue type and taking into account the neighboring residues in the sequence. This data is compared to a database of empirically determined ϕ and ψ angles obtained from 200 X-ray or NMR protein structures that have near complete backbone NMR assignment data in the BMRB along with high-resolution X-ray crystal structures deposited in the PDB.

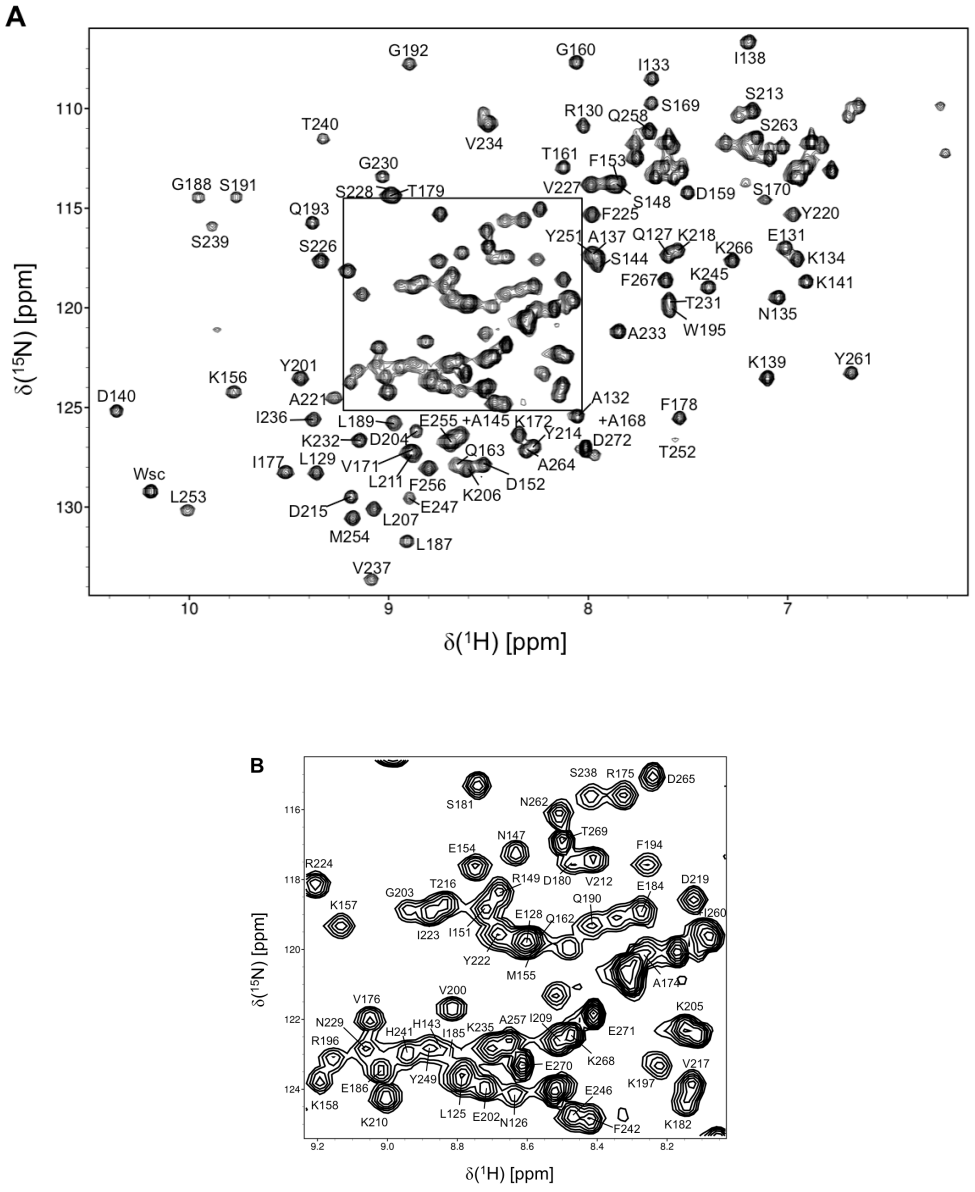


Figure 3-2. Labeled ^{1}H-^{15}N HSQC spectrum of $IsdB^{N1}$. (A)2D ^{1}H-^{15}N HSQC spectrum of 1 mM $IsdB^{N1}$ in 50 mM $NaHPO_4/NaH_2PO_4$, pH 6.8, 400 mM NaCl, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectrum was recorded at 298 K on a Bruker DRX 600 MHz spectrometer. Backbone resonance assignments are indicated with the *one-letter* amino acid code and residue number. The tryptophan indole NH side chain signal is marked with ‘‘Wsc.’’ A ‘‘+’’ in front of residue numbers marks $^{1}H/^{15}N$ resonances that are too weak to be observed at the current intensity level of the displayed spectrum. Resonances from side chain NH_2 are not labeled. (B) *Insert* a close-up view of the congested middle region of the HSQC spectrum with assigned NH resonances labeled.

The chemical shift data of the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ atoms delineated the presence of 8 β-strands spanning residues 150–156 (β1); 175–178 (β2); 183–189 (β3); 198–202 (β4); 208–215 (β5); 221–227 (β6); 223–242 (β7); 252–256 (β8), and a single α-helical segment spanning residues 128–134 (Fig. 3-3). The 8 β-strands point to a similar overall fold as that of other published NEAT domains, including $IsdH^{N1}$ (57).Backbone and side-chain (^{1}H, ^{15}N, ^{13}C) NMR resonance assignments for $IsdB^{N1}$ have been deposited in the BioMagResBank (http://www.bmrb.wise.edu) under the BMRB accession number 19056.

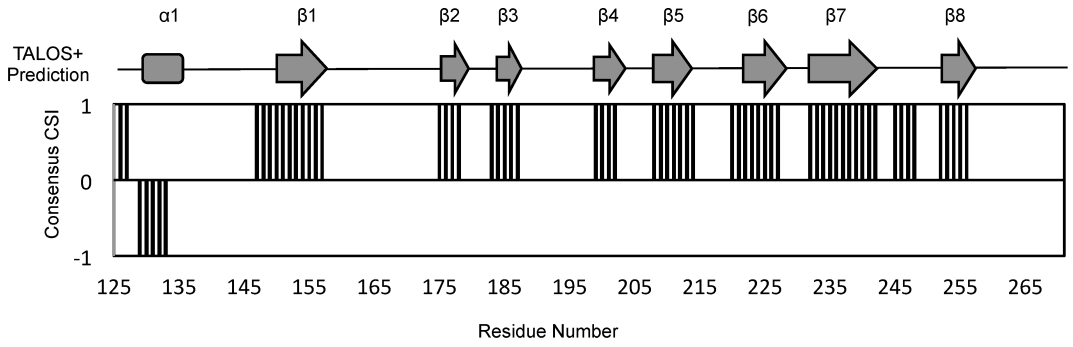


Figure 3-3. Secondary structural topology of $IsdB^{N1}$ derived from chemical shift indexing and dihedral angles predictions from TALOS+ analysis of chemical shift data. The α-helix and 8 β-strands of $IsdB^{N1}$ are shown above the CSI and TALOS+ predictions as a rectangle and arrows, respectively. The consensus chemical shift index (CSI) plot shown below $IsdB^{N1}$ secondary structural elements is based on the CSI analysis of all H_{α}, C_{α}, and C_{β} chemical shifts

NMR Solution Structure of IsdB^{N1}

A total of 1535 NOE assignments were extracted from the ^{15}N- and ^{13}C-edited NOESY spectra. Of these, 825 structurally significant proton-proton distance restraints were used in the final structure calculation as determined by the *distances modify* command in CYANA. The *distances modify* command inserts pseudo atoms if there is no stereospecific assignments available, modifies distances based on known distance restraints between intra residue atoms and adjusts distances for pseudo atoms, and removes redundant and/or meaningless distance restraints from the data set. Examples of meaningless distance restraints include duplicate restraints between the same two protons (the more restrictive restraint is kept), protons with fixed distances such as protons in the same aromatic ring, and protons that have no restriction. These restraints are usually found as intra-residue protons where there is no structural conformation that could violate the set distance, therefore this information would not add to the overall three-dimensional fold of the protein. The structure calculations additionally included 254 dihedral angles derived from chemical shift data using the program TALOS+ and 90 hydrogen-bond restraints (2 restraints per NH pair). The chosen 20 conformers of the final ensemble for IsdB^{N1} have no NOE or dihedral angle violations greater than 0.5 Å or 5°, respectively. The root mean square deviation (RMSD), which is the average distance between the Cα positions of the backbone atoms from backbone position of the mean structure, is 0.65 Å respectively for the secondary structural elements. Analysis of the dihedral angles (Figure 3-4) revealed that 85% of residues are in the most favored conformations on the Ramachadran plot and 13% are in additionally allowed regions, which correspond to a

reasonable structure. Further structural analysis using Verify3D (91), ProsaII, and Procheck revealed global raw scores of 0.35, 0.51. and -0.69, respectively. Verify3D assesses each amino acid in the 3D structure and compares it to the 1D protein sequence. The criterion that determines how well it fits into the model is based on the burial area, the fraction of side-chain area covered by polar residues, and the secondary structure. The global raw score ranges from -1 to +1 with +1 being the best score. The calculated NMR structural model falls into the acceptable range for Verify3D. The final restraint and structural statistics for $IsdB^{N1}$ are listed in Table 3-1 and the final structural NMR ensemble of $IsdB^{N1}$ is shown in Figure 3-5.

As suggested by the solved structures of other NEAT domains from Isd proteins, $IsdB^{N1}$ is an immunoglobulin-like fold with a twisted anti-parallel β sandwich (Figure 3-5B). The N-terminus begins with a small α-helix (L129-K134) and is followed by a 13-residue stretch of random coil (N135-N147). The coil wraps around the first β-sheet and ends with the start of strand β1a (S148-P150). This helix/loop region was not included in the first $IsdB^{N1}$ construct. There are, however, van der Waal interactions between L129, I133, and I138 and the second β-sheet, which probably added to the stability of the overall structure (Figure 3-6). Typical of β -sandwiches, there are many cross connections between the two sheets. The first β-sheet comprises strands β1a-β2-β3-β6-β5b and forms a twisted antiparallel β-sheet (Figure 3-5B). The opposite β-sheet comprises β1b-β8-β7-β4-β5a. The arrangement of β strands 3 through 7 forms a Greek key motif (Figure 3-7). This topology, found in the majority of known β-sandwiches, intertwines between the two sheets.

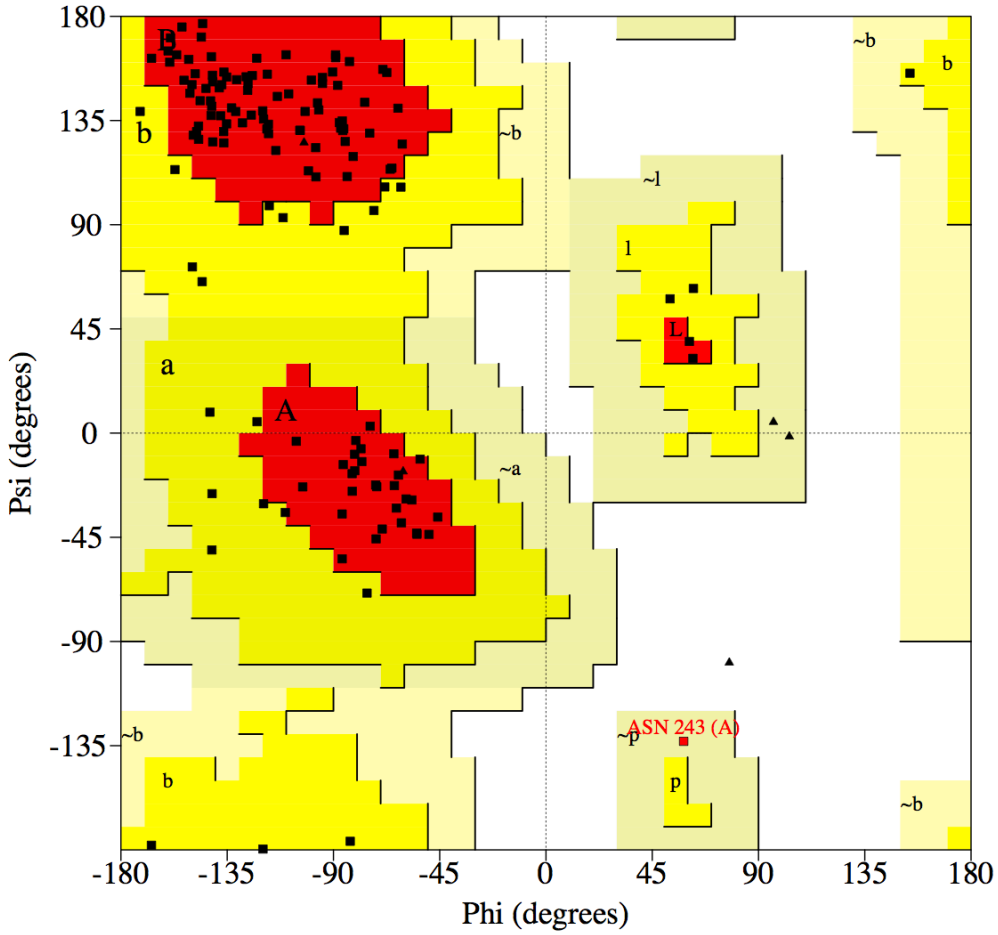


Figure 3-4. Ramachandran plot of IsdB^{N1}. The core regions in red correspond to most favored, yellow regions to additionally allowed, tan regions to generously allowed, and white regions to disallowed regions. Residues are represented by black squares in favored and additionally allowed regions, and residues in generously allowed or disallowed regions in red. Asn243 is in a generously allowed region and in the structure is found in the tight β-hairpin between strands β7 and β8. The Ramachandran plot was generated by PROCHECK-NMR for the lowest energy conformers of the NMR ensemble of IsdB^{N1} structures.

Table 3-1. NMR structure statistics for $IsdB^{N1}$

Constraints for final structure calculations	
Total NOE distance restraints	825
Intraresidue	132
Sequential (| *i -j* |= 1)	315
Medium range (1 < | *i -j* | < 5)	97
Long range (| *i -j* | ≥ 5)	281
Dihedral angle restraints[a]	
ϕ angles	136
ψ angles	141
Hydrogen bonds restraints[b]	46
Structure statistics (20 conformers)	
CYANA target function ($Å^2$)[c]	1.68
Residual distance violations	
Number > 0.5 Å	0
Ramachandran plot statistics (%)[d]	
Residues in most favored regions	84.6
Residues in additionally allowed regions	13.6
Residues in generously allowed regions	1.8
Residues in disallowed regions	0
Average rmsd to mean structure (Å)[e]	
Protein backbone	0.65 ± 0.11
Protein heavy atoms	1.41 ± 0.17
Global Raw Scores[f]	
Verify3D	0.35
ProsaII	0.51
Procheck (ϕ-ψ)	-0.69

[a] ϕ and ψ angles were derived from the program TALOS+, based on the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$ and ^{15}N chemical shifts. [b] Two distance restraints between the amide and carbonyl group atoms (NH-O = 1.8-2.8 Å, N-O =2.5-3.8 Å) were used for the hydrogen bond restraints. [c] CYANA target function calculated after water refinement in AMBER with the NOE restraints. [d] Ramachandran plot analysis performed with PROCHECK. [e] Residues 129-133, 151-156, 175-178, 184-189, 198-202, 209-215, 221-227, 233-242, 251-256 were used. [f] Global raw scores derived from Protein Structure Validation Suite.

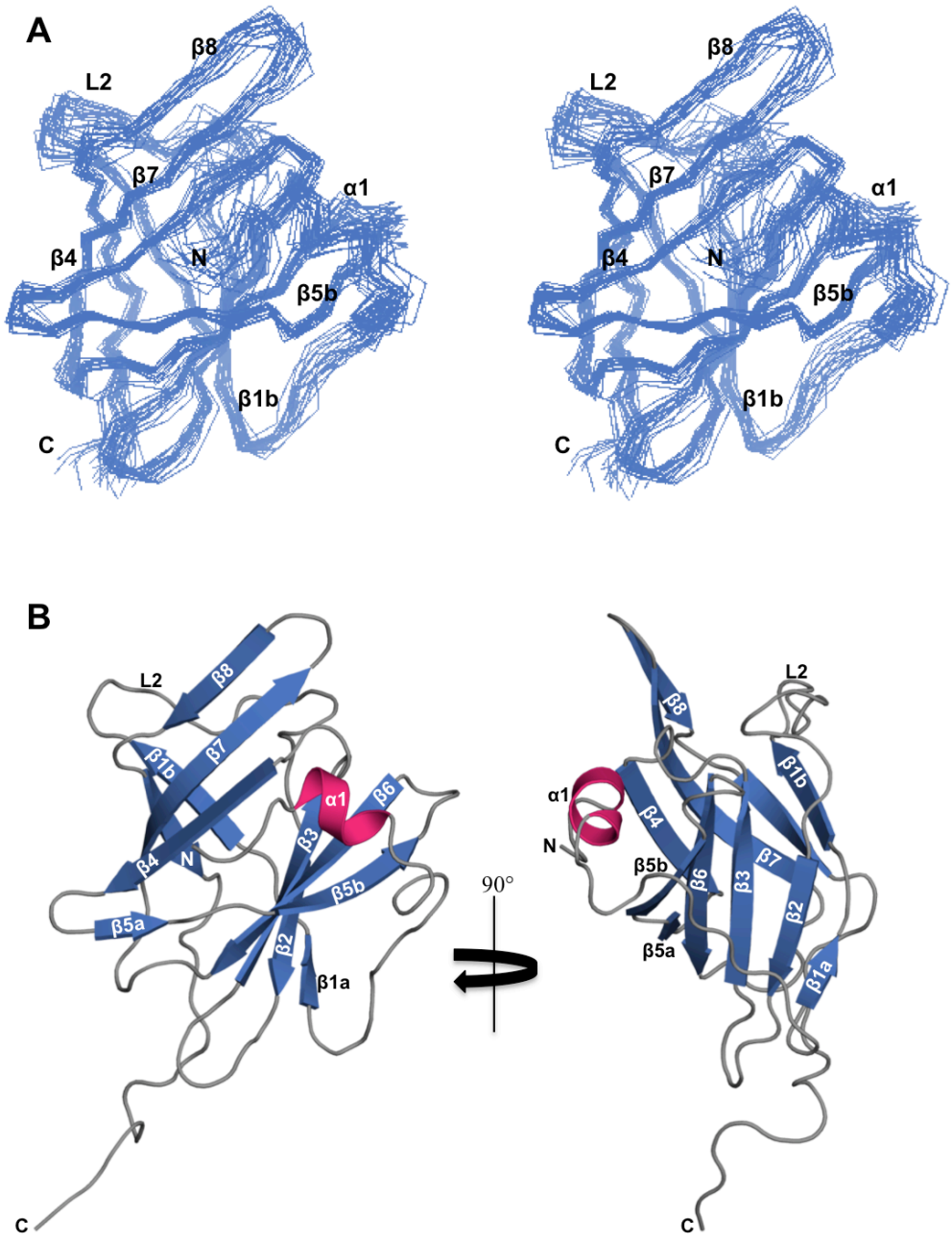


Figure 3-5. Solution structure of IsdB^{N1}. (A) A cross-eyed stereo view of NMR ensemble of IsdB^{N1}. The calculated and water refined 20 lowest energy conformers. The disordered C-terminus has been removed from this view. The backbone heavy atoms (N, C_α, and C') of the secondary structural elements were superimposed using the program Theseus version 2.0.6. (B) Ribbon representation of the lowest energy structure of IsdB^{N1} selected from the final NMR ensemble. The secondary structure elements are labeled and color-coded with the β-strands in marine blue, the α-helix in magenta, and the loops in gray.

The Greek key motif in IsdB^{N1} adopts the B form of the typical β-sandwich Greek key motifs (92), where the third strand (β5) is hydrogen bonded to the fourth strand (β6) and is therefore part of the sheet that contains the first and fourth strands (β3- β6- β5). Although β5 does hydrogen bond to the sheet (β4- β7) before it breaks to cross over to the opposite face, the motif is still considered the B form (rather than the A form) because it forms more contacts with β6 (Figure 3-7). A few backbone assignments are missing in β8 between residues N243 and Y251 in the NMR spectra, causing a lack of inter-proton distance restraints for that region. In addition, a "break" in the hydrogen-bonding pattern from residues Y249 to Y251 causes a small bulge in β8 (can be seen as a break in the ribbon cartoon representation in Figure 3-5B). The hydrophobic core of the β-sandwich contains the bulky side chains of Phe153, Ile185, Leu187, Leu189, Trp195, Phe198, Val200, Ile223, Phe225, Val234, and Ile236. Although the NEAT domain is a well fold immunoglobulin-like fold, L2, the loop region between β strands 1b and 2 comprising the residues Q163-A168, is disordered. This loop segment contains a series of consecutive aromatic residues, F164, Y165, H166, and Y167, whose backbone chemical shifts were unassigned in the NMR spectra.

Structural Homology

The characteristic immunoglobulin-like fold of the NEAT domain is highly conserved and therefore it is no surprise that a DALI search of structural homologs revealed NEAT domains from other *S. aureus* Isd proteins as the closest structural homologs. The most similar NEAT domains included both of the non-heme binding NEAT domains of IsdH, IsdH^{N1} (PDB entry 2H3K) (57) and IsdH^{N2} (PDB entry 4IJ2)

(93) with 2.1 Å rmsd over 139 aligned residues with 43% sequence identity and 2.0 Å rmsd over 133 aligned residues with 65% sequence identity, respectively. The closest heme binding homolog is the NEAT domain of IsdA (PDB 3QZN) (58) with 2.4 Å rmsd over 115 residues with 23% sequence identity. The structural similarities between $IsdB^{N1}$ and $IsdH^{N1}$ and $IsdH^{N2}$ can be seen in the overlays in Fig. 3-8. As seen in the first overlay of $IsdB^{N1}$ with the solution NMR structure of $IsdH^{N1}$, L2 is disordered. This disordered loop region in $IsdH^{N1}$ contains a similar sequence of aromatic residues, YYHFF (57). In the crystal structure of $IsdH^{N1}$ complexed to metHb, L2 adopts an α-helical conformation. There are many examples of proteins, or regions of proteins, adopting structural order upon binding with their target (94). Although thermodynamically switching to a more ordered structure has an entropic cost, i.e. a large reduction in external entropy by restricting the translational and rotational freedom of the unstructured portion of the protein, the favorable enthalpic contributions from the protein-protein interactions (van der Waals, hydrogen bonding) can compensate for the reduction in entropy. Adopting structural order upon binding is one way to ensure binding specificity (94). Upon comparing $IsdB^{N1}$ to the crystal structure of $IsdH^{N2}$ there is an RMSD of 2 Å. From a structural perspective the greatest area of difference is also in L2, where the aromatic motif of $IsdH^{N2}$ forms an α-helix in the metHb bound structure. This conserved aromatic sequence is found in all of the metHb binding NEAT domains of Isd proteins in *S. aureus* (Fig. 3-10), with $IsdB^{N1}$ and $IsdH^{N2}$ possessing the same aromatic residues FYHY.

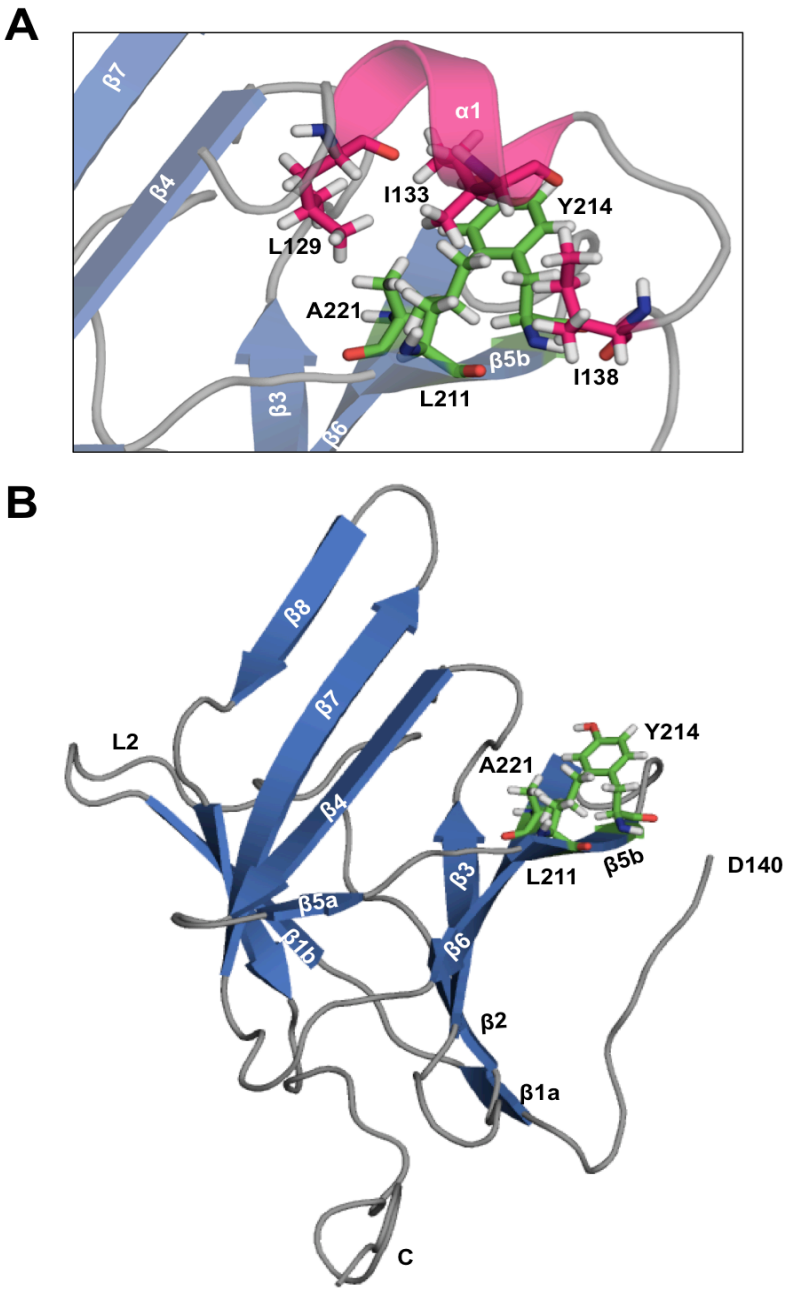


Figure 3-6. Van der Waal interactions of residues in α1 and β sheet 1. (A) A close up view of the van der Waal interactions between L129, I133, and I138 with residues in the first β sheet, L211, Y214, and A221. (B) A ribbon representation of the original $IsdB^{N1}$ construct that did not contain residues 125-139. The hydrophobic residues of the first β sheet would be exposed in this truncated construct. This might have led to the instability of the initial protein sample. The secondary structure elements are labeled and color-coded with the β-strands in marine blue, the α-helix in magenta, and the loops in gray.

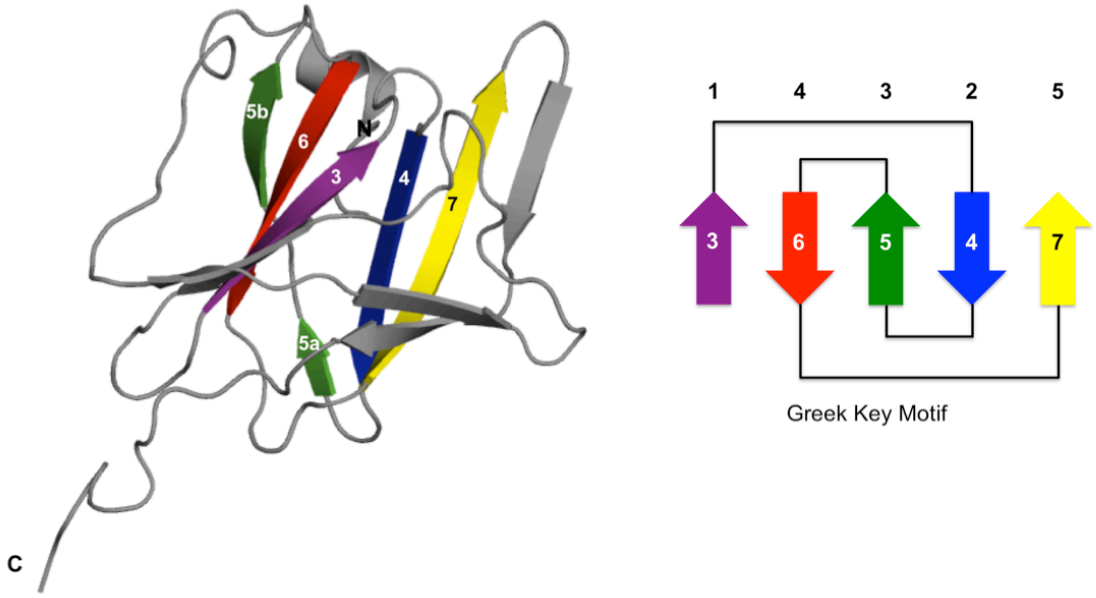


Figure 3-7. Greek key motif in IsdB^{N1}. On the left is a ribbon representation of IsdB^{N1} in gray with the N- and C-termini labeled and the five β strands (3,4, 5a-b, 6, and 7) comprising the Greek key motif colored in purple, blue, green, red, and yellow. On the right is a corresponding color-coded schematic of the Greek key motif. The numbers above the schematic 1-4-3-2-5 correspond to the *G* topology of a Greek key motif.

Another interesting feature of IsdB^{N1} is the small α-helix at the N-terminus followed by the long region of random coil that wraps around the outside of the β sheet. The solved structures of NEAT domains involved in heme binding do not have this feature and instead originate at the first β-strand. This same helix/loop region, however, appears in the NEAT domains structures that do not participate in heme binding (IsdH^{N1} and IsdH^{N2}). Moreover, this segment seems to be important for the overall folding of the domain, given that the first construct of IsdB^{N1} was unstable without this extended N-terminal region. The important van der Waals interactions between residues in the α-helix located in this N-terminal region and the underlying β-sheet is shown in Fig. 3-6.

The NEAT domain of IsdA is the closest heme-binding structural homolog to $IsdB^{N1}$. Although, the overall Ig-like fold is present in both structures, there are some notable differences between the two NEAT domains. The first difference is the lack of a heme-binding pocket in $IsdB^{N1}$. In addition to $IsdB^{N1}$ not possessing the conserved YXXXY motif of the heme-binding NEAT domains of S. aureus Isd proteins, the lack of a fully folded α-helix prevents a true pocket from forming. Moreover, β8 of IsdA forms more of a curve in order to sandwich the heme group, whereas β8 of $IsdB^{N1}$ contains a β-bulge in the middle of the strand that causes a unique twist of the sheet that is also seen in the structures of $IsdH^{N1}$ and $IsdH^{N2}$.

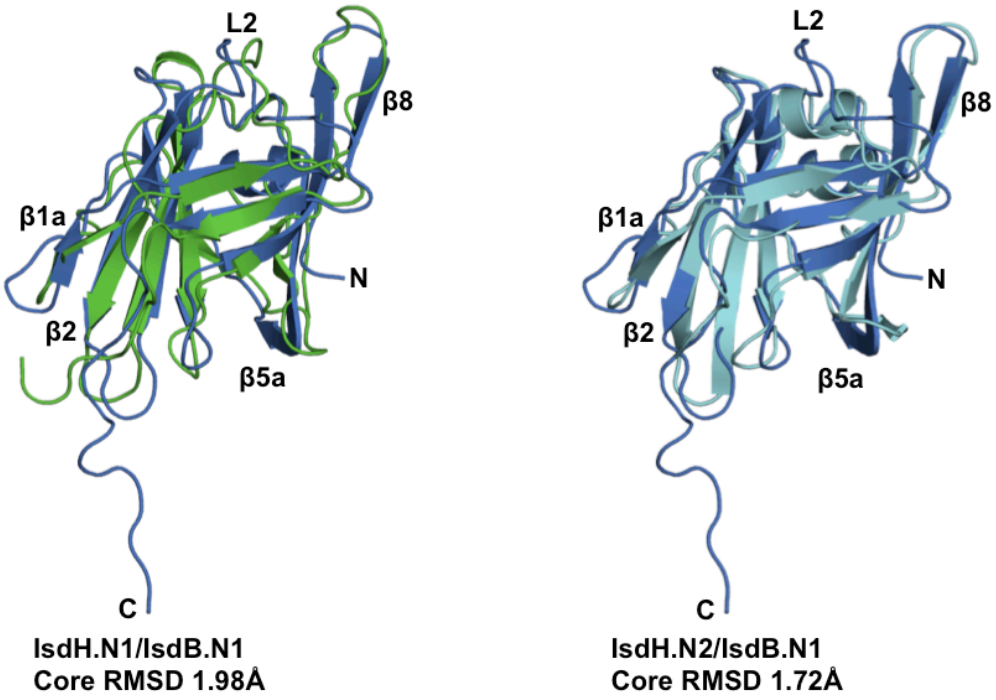


Figure 3-8. Comparison of the structure of $IsdB^{N1}$ with (Left) the solution NMR structure of $IsdH^{N1}$ (PDB 2H3K) (Right) the crystal structure of $IsdH^{N2}$ (PDB 4IJ2). $IsdB^{N1}$ is colored marine blue. The core RMSD value was calculated with WinCoot (95).

The structure of IsdB^{N1} was used to compare the primary sequences of all the NEAT domains from *S. aureus* using the structural alignments from Dali (Fig 3-10). The conserved residues between all of the NEAT domains can predominantly be found in the secondary structural elements, with the loop regions being areas of less sequence homology. The best example of this is the NEAT domain of IsdC, which has an additional 8 residues in the loop region that connects β7 and β8, which has been shown to be essential in its interactions and heme transfer to the lipoprotein, IsdE (47). Of the most conserved residues, one tryptophan is found in the hydrophobic core of all of the NEAT domains and makes contacts to the heme moiety in the heme-binding NEAT domains,

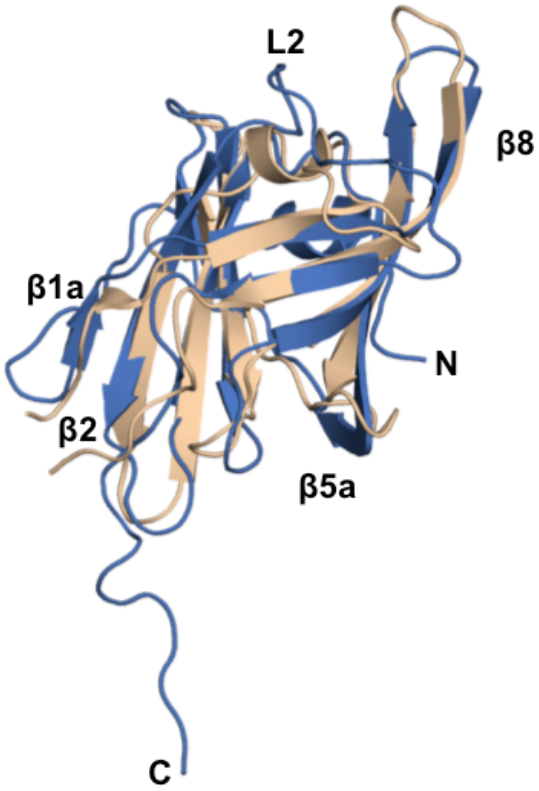


Figure 3-9. Comparison of the structure of IsdB^{N1} with the NEAT domain of IsdA. IsdB^{N1} is colored marine blue and IsdA is colored in wheat (PDB 3QZN).

$IsdB^{N2}$, $IsdH^{N3}$, IsdA, and IsdC (47,58,59,61). The Hb-binding domains possess a conserved aromatic sequence in the loop region (L2) connecting strands β1 and β2. This aromatic motif in $IsdH^{N1}$ (YYHFF) and $IsdH^{N2}$ (FYHY) is the major Hb-binding interface as shown in the bound crystal structures (93,96). In contrast, the heme-binding NEAT domains do not contain aromatic residues in this region, and instead have a conserved serine and methionine. The conserved serine makes contact with a propionate group of the heme. A unique feature of the Hb-binding NEAT domains, $IsdB^{N1}$, $IsdH^{N1}$, and $IsdH^{N2}$ is the N-terminal α-helix. The sequence for this region is highly conserved, with Leu and Ile available to form van der Waal interactions with the underlying β-sheet. Furthermore, the loop region between β6 and β7 contains a highly conserved sequence of VSXGT. This loop region is used to crossover from the first β sheet back to the second β sheet.

Hemoglobin Binding

To investigate the binding of $IsdB^{N1}$ to metHb, a series of titration experiments were performed using ^{15}N-labeled $IsdB^{N1}$ and non-isotopically labeled metHb. The control ^{1}H-^{15}N HSQC spectrum was recorded on a 0.5 mM sample of ^{15}N-labeled $IsdB^{N1}$ in 50mM sodium phosphate at pH 6.8 with 100 mM sodium chloride (Figure 3-11A). A 0.5 mM concentration of non-isotopically labeled metHb (αβ dimer, MW 32,000 Da) in the same buffer was mixed with the ^{15}N-labeled $IsdB^{N1}$ for a protein ratio of 1:1 metHb:$IsdB^{N1}$. As shown in Fig. 3-8B, the resulting HSQC spectrum shows a near

complete disappearance of signal intensity and substantial linebroadening of the amide proton cross-peaks, indicative of a large complex formation of IsdB^{N1} with metHb.

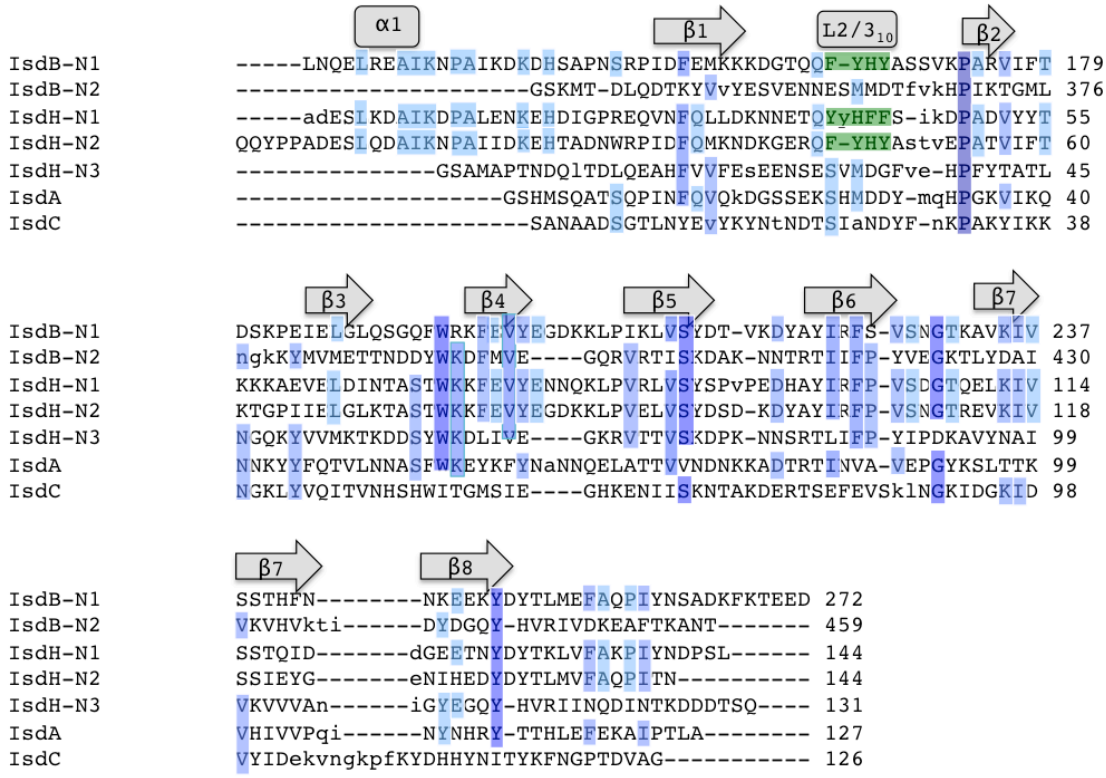


Figure 3-10. Sequence alignment of IsdB^{N1} to the NEAT domains of all Isd proteins from *S. aureus*. PDB ID of IsdB^{N2} (3RTL), IsdH^{N1} (2H3K), IsdH^{N2} (4IJ2), IsdH^{N3} (2Z6F), IsdA (3QZN), and IsdC (2O6P). The uppercase letters are the amino acids that have structural equivalence, while the lowercase amino acids are structurally nonequivalent. Secondary structure elements are denoted with arrows for β-strands and rectangles for helices. The sequence alignment was generated using DaliLite version 3.1 using IsdB^{N1} as the subject (87).

The same experiment was carried out with protein ratios of metHb:IsdB^{N1} of 1:0.5, 1:0.75, 1:2, and 1:3. In order to assure that the observed effect was not caused by the dimeric state of metHb, the experiment was repeated on E73, a DNA binding protein from the Sulfolobus Turreted Icosahedral Virus (97), which should not interact with

metHb. The resulting HSQC spectrum displayed no broadening or diminished signal intensities of cross peaks verifying that $IsdB^{N1}$ is indeed interacting with metHb. Interestingly, the intense peaks that still appear in the HSQC spectrum belong to the C-termini, Thr269-Asp272, indicating that the C-terminus of $IsdB^{N1}$ does not experience an altered chemical environment upon $IsdB^{N1}$ binding to metHb. Adding additional ^{15}N-labeled $IsdB^{N1}$ until the ratio was ~3:1 $IsdB^{N1}$:metHb revealed an increase signal intensity of the amide proton crosspeaks indicating unbound ^{15}N-labeled $IsdB^{N1}$ in solution (Figure 3-11C). In order to determine whether the previous results were a result of aggregation, a competitive binding study was performed. Increasing amounts of unlabeled $IsdB^{N1}$ sample was added to a sample of 1:1 ratio of ^{15}N-labeled $IsdB^{N1}$:metHb. After the addition of 4 equivalents of unlabeled protein, the signal intensities of ^{15}N-labeled $IsdB^{N1}$ began to increase (Figure 3-11D). This indicated that the interaction between $IsdB^{N1}$ and metHb was due to complex formation and not aggregation; because as more unlabeled $IsdB^{N1}$ protein was added, it outcompeted the binding of the ^{15}N-labeled $IsdB^{N1}$.

Based on the bound crystal structures of $IsdH^{N1}$ and $IsdH^{N2}$ to metHb where the aromatic motif located in L2 is the predominant binding area, the proposed metHb binding site of $IsdB^{N1}$ would also be the aromatic motif, FYHY, located in L2 (Figure 3-12). To test whether the aromatic residues in this region are indeed responsible for $IsdB^{N1}$ bidning to metHb, three $IsdB^{N1}$ variants were generated, single amino acid substitutions, $IsdB^{N1}$-F164D and $IsdB^{N1}$-Y167D, and a double amino acid substitution $IsdB^{N1}$-F164D/Y167D. These amino acids, Phe164 and Tyr167, were selected based on a comparison of the structure of IsdBN1 overlaid on the metHb bound crystal structures of

$IsdH^{N1}$ and $IsdH^{N2}$ (Fig. 3-12). When the first variant, $IsdB^{N1}$-Y167D, was titrated with a 1:1 ratio of metHb, the resulting HSQC spectrum indicated the same result as wild type $IsdB^{N1}$ (Figure 3-11E), the signal intensities were greatly diminished. However, a lower ^{15}N-labeled $IsdB^{N1}$-Y167D variant protein concentration was required to increase the signal intensity and show unbound variant protein, indicating that this mutation did slightly affect the binding capability of $IsdB^{N1}$ to metHb. The second variant, $IsdB^{N1}$ F164D/Y167D, did not show any indication of interaction with metHb. In view of this result, the single variant $IsdB^{N1}$-F164D was engineered to test whether both substitutions were needed to diminish metHb binding. The subsequent HSQC of a 1:1 ratio of the $IsdB^{N1}$-F164D variant with metHb exhibited no signs of interaction (Figure 3-11F). This indicated that substituting just one amino acid, Phe164 to an Asp, was enough to knockout metHb binding within the NMR conditions.

Heme Transfer Reactions

The amino acid substitutions of F164D and Y167D indicated a disruption of metHb binding in the NMR conditions. In order to investigate how these substitutions would effect the heme transfer reaction from metHb to IsdB, the F164D, Y167D, and F164D/Y167D amino acid substitutions were engineered into the construct, $IsdB^{N1LN2}$, comprising residues L125-N458 of full length IsdB. It was previously reported that heme transfer from metHb to IsdB is rapid (49) and that only the N1 domain covalently bound to the N2 domain with the Linker region is needed for the transfer reaction (98). The transfer of heme from metHb to IsdB can be monitored via UV-Vis spectroscopy by measuring the decrease in absorbance of the Soret band at 406nm. As seen in Figure 3-

13, the Soret band of heme-bound metHb has a sharper profile than that of the Soret band of heme-bound IsdB.

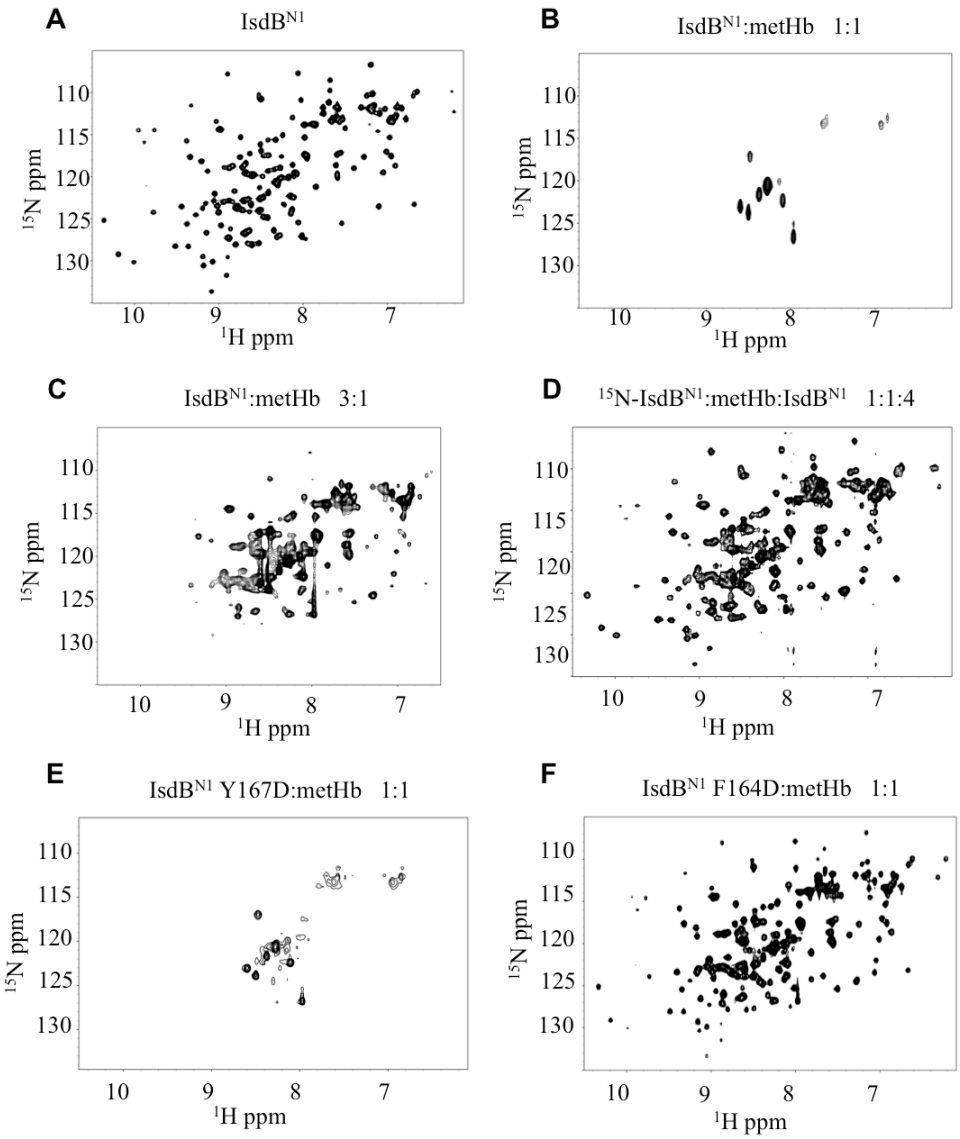


Figure 3-11. Hemoglobin binding studies of IsdB^{N1} and IsdB^{N1} protein variants. All 2D ^{1}H-^{15}N HSQC spectra were recorded at 298 K in 50 mM sodium phosphate at pH 6.8 with 100 mM sodium chloride. (A) A control sample of ^{15}N-labeled IsdB^{N1} (B) A 1:1 ratio of ^{15}N-labeled IsdB^{N1} with metHb (αβ dimer) (C) Titration of 3:1 ratio of ^{15}N-labeled IsdB^{N1} with metHb (D) Competitive binding study where non-isotopically labeled IsdB^{N1} was added to a 1:1 ratio of ^{15}N-labeled IsdB^{N1} with metHb (E) A 1:1 ratio of IsdB^{N1}-Y167D mutant with metHb (F) A 1:1 ratio of IsdB^{N1}-F164D mutant.

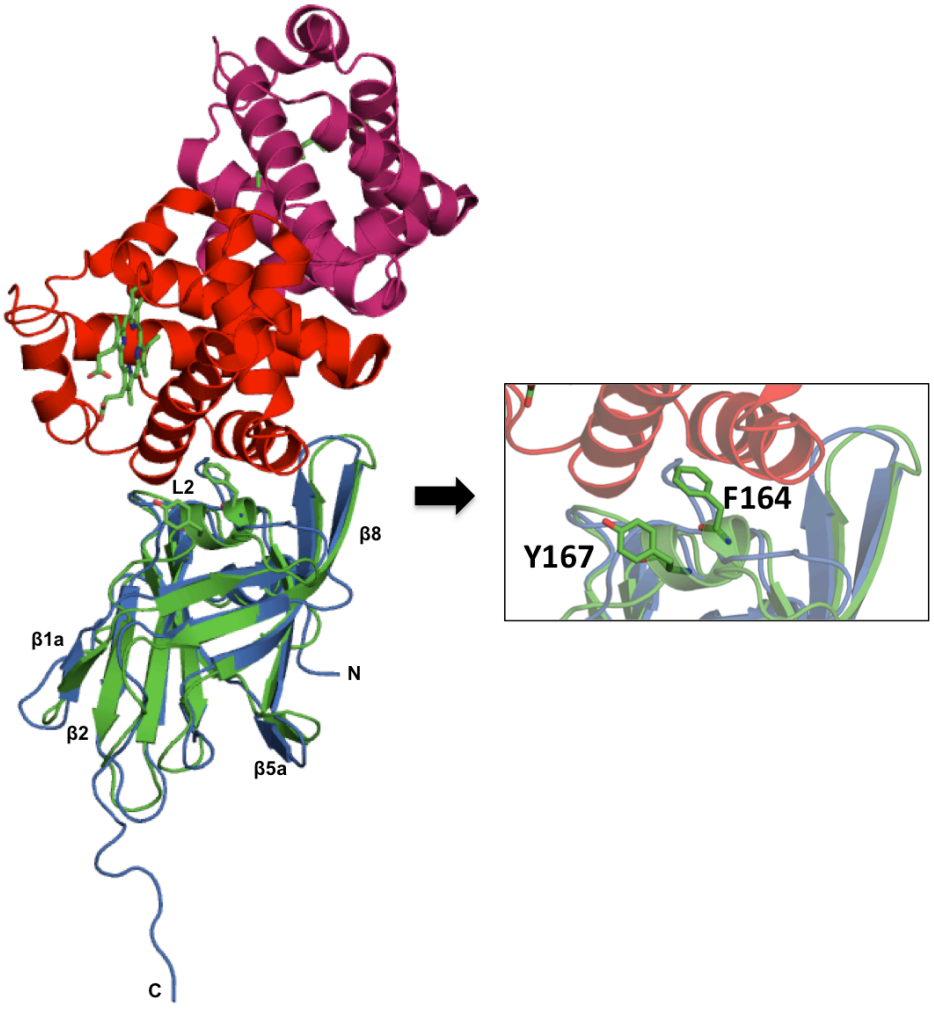


Figure 3-12. Structural alignment between IsdB^{N1} and IsdH^{N2}. (a) A superimposition from COOT of IsdB^{N1} with IsdH^{N2} (PDB:4IJ2) bound to methemoglobin. IsdB^{N1} is in marine blue, IsdH^{N2} is in green, and the α-subunit of methemoglobin is in red and the β-subunit in magenta with the heme groups colored green. (b) A close-up of the metHb-binding site of IsdH^{N2}. In the solution structure of IsdB^{N1}, loop 2 is disorder, whereas the equivalent loop region in IsdH^{N2} forms an α-helix when bound to metHb. The aromatic residues F164 and Y167 make contacts with metHb in the crystal structure of IsdH^{N2}, which prompted the substitution of the equivalent residues in IsdB^{N1} to investigate their possible role in metHb binding and heme transfer.

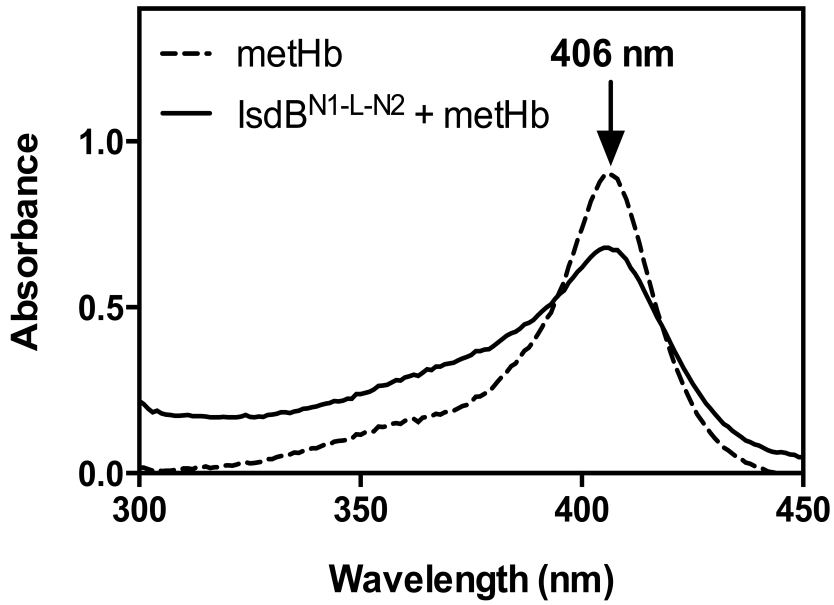


Figure 3-13. Heme transfer between methemoglobin and $IsdB^{N1LN2}$. A UV-Vis scan showing the Soret band of metHb with 4 μM heme concentration at 406 nm (dashed line) and the decrease in absorbance once heme has been transferred from metHb to $IsdB^{N1LN2}$ (25 μM protein concentration) (solid line).

The relative rate of heme transfer from metHb to wild type $IsdB^{N1LN2}$ was determined to be $0.040 \pm 0.001\ s^{-1}$ (Figure 3-14A, Table 3-2). MetHb transferred its heme to the single variant $IsdB^{N1LN2}$-Y167D a little faster than wild type with a relative rate of $0.062 \pm 0.001\ s^{-1}$ but still within the same realm of the wild type rate. While the single substitution, F164D, appears to knock out $IsdB^{N1}$ binding to metHb in the NMR, this same substitution in $IsdB^{N1LN2}$ does not completely diminish the ability of $IsdB^{N1LN2}$ to steal heme from metHb, although the rate is significantly decreased to $0.010 \pm 0.001\ s^{-1}$. Although this result was surprising, the change of phenylalanine to an aspartic acid could cause interaction with Lys-11 on the α-chain of metHb, which is located near the binding site of $IsdH^{N1}$ and $IsdH^{N2}$. To test this, another amino acid substitution, F164A, was

engineered. As shown in Fig. 3-14A, the change of Phe164 to an alanine decreased the reaction rate even further to 0.004 ± 0.001 s^{-1}. This was the same reaction rate as the double amino acid substitution, F164D and Y167D, $IsdB^{N1LN2}$ variant, which is comparable to the rate of heme transfer from metHb to a construct of IsdB lacking the N1 domain, $IsdB^{LinkerN2}$ (residues L271-N458) (Figure 3-14B). Interestingly, when the linker region is covalently attached to N1, $IsdB^{N1\text{-}L}$, (residues L125-T337) the heme transfer from metHb to the N2 domain does proceed albeit at a slower rate than wild type $IsdB^{N1LN2}$, 0.013 ± 0.001 s^{-1} (Figure 3-14B).

The rate by which N2 can abstract heme from metHb seems to be influenced by the presence of the linker region. In order to investigate if the whole linker region is needed for heme abstraction, constructs of truncated $IsdB^{Linker\text{-}N2}$ were engineered with either 35 or 20 C-terminal amino acid residues of the linker ($IsdB^{35LN2}$, residues E304-N458; $IsdB^{20LN2}$, residues K319-N458). Monitoring heme transfer from metHb to these truncated constructs, it appears that there are residues in the N-terminal section of the linker region that are needed for heme transfer, as $IsdB^{35LN2}$ and $IsdB^{20LN2}$ did not have a significant difference in rate from $IsdB^{N2}$. Nevertheless, the presence of N1 covalently attached with the linker and N2 still has the greatest influence on the transfer rate from metHb. In the short time period used for the reactions, all of the trials fit a single-phase exponential expression. The heme transfer reaction between metHb and IsdB is fast and fits a single phase exponential, indicating direct heme transfer from metHb to IsdB (49). A two-phase reaction would suggest the dissociation of heme from β and α subunits, respectively, followed by uptake by the truncated IsdB constructs. In this study, only the

"faster" first phase was observed (i.e. the dissociation of heme from the β subunit of metHb which has a much weaker affinity for the heme than the α subunit), and the time course would need to be increased to observe the slower second phase (99).

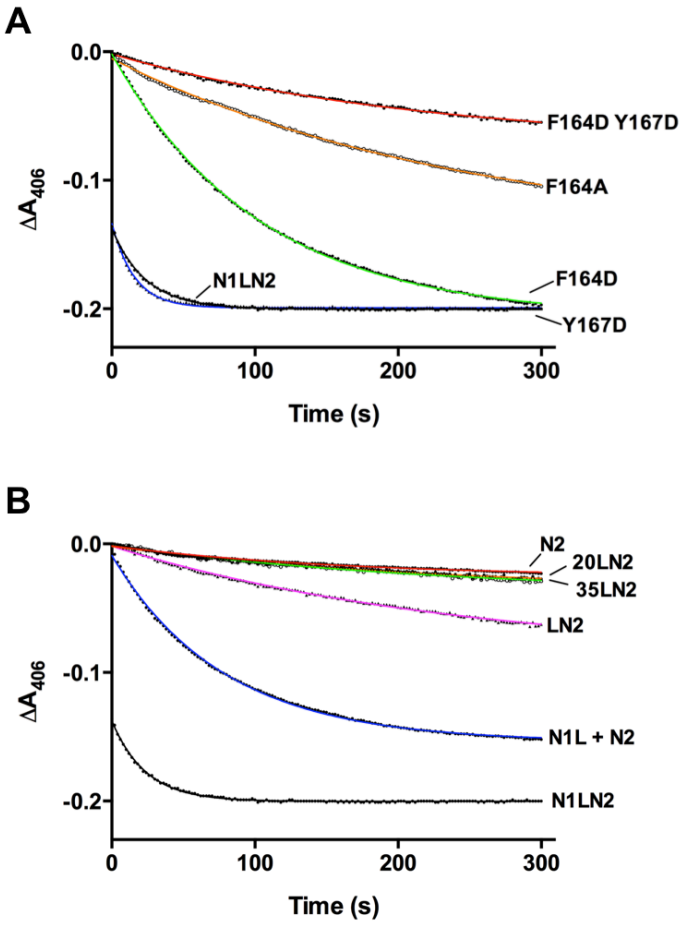


Figure 3-14. Heme transfer reactions between methemoglobin (metHb) and various IsdB constructs. (A) The heme transfer reactions from metHb to $IsdB^{N1LN2}$ and the engineered $IsdB^{N1LN2}$ protein variants. (B) The heme transfer reactions from metHb to truncated IsdB constructs. There is approximately a 10 second delay from the start of the reaction and the time point 0s. Heme transfer was monitored by the relative change in absorbance at 406 nm over a period of 5 minutes. Holo-metHb (4μM) was mixed with apo-acceptor IsdB protein variants (25 μM) in 20mM $NaHPO_4$ 150mM NaCl at pH 7.5. The reactions were performed in triplicate and the average of two or three trials were graphed respectively. The data points are in black or gray, and the data was fit to a single-phase exponential expression (the colored lines) using Prism (*Graphpad*) software.

Table 3-2. Rates of heme transfer between metHb and IsdB constructs at 25°C

N1-L-N2 Constructs	k s^{-1}
WT N1-L-N2	0.040 ± 0.001
F164D	0.010 ± 0.001
Y167D	0.062 ± 0.001
F164A	0.004 ± 0.001
F164D Y167D	0.004 ± 0.001
Truncated IsdB Constructs	**k s^{-1}**
N1-L + N2	0.013 ± 0.001
L-N2	0.004 ± 0.001
35L-N2	0.004 ± 0.001
20L-N2	0.004 ± 0.001
N2	0.006 ± 0.001

This work complements the recently published article by Pishchany and colleagues, in which the authors also examined the amino acids responsible for $IsdB^{N1}$ binding to metHb (100). Using the sequence homology of $IsdB^{N1}$ with $IsdH^{N1}$, the same aromatic residues, F164 and Y167, in L2 were changed to alanine, which abolished Hb binding to the staphylococcal cell surface. Interestingly, engineering a chimera of $IsdB^{N1}$ that incorporated the YYHFF amino acid sequence of L2 from $IsdH^{N1}$ into L2 of $IsdB^{N1}$ showed that $IsdB^{N1}$ was not able to bind Hb. This suggests that specific contacts in the structure of $IsdB^{N1}$ are unique for interaction with metHb or for the stabilization of the α-helix upon binding. Given that the aromatic amino acid sequence of L2 in $IsdH^{N1}$ does bind to Hb in the published crystal structure (96). Although the F164 and Y167 changes knocked out Hb binding to the surface, they did not include these two variants in their heme transfer studies. Instead, they changed Y165 to an alanine and showed a similar effect on heme transfer as our double substitution, $IsdB^{N1LN2}$-F164D/Y167D. The rate

was closer to that of heme dissociation from Hb. In addition, the authors engineered the Y165A variant of IsdB into a strain of *S. aureus*. With hemoglobin as the sole iron source, the $isdB_{Y165A}$ mutant strain did not grow. Also in an established mouse model expressing human hemoglobin, a 100-fold decrease in colony forming units (CFUs) in the hearts and kidneys of infected mice was shown for both an IsdB deletion variant, *ΔisdB*, and for the Y165A variant, $isdB_{Y165A}$.

In comparing hemoglobin receptors, it is well documented that bacteria have evolved different structural architectures to achieve the goal of recognizing hemoglobin and abstracting heme. Many of the Gram-negative receptors adopt a β-barrel structure with surface loops that facilitate binding of Hb (16). In Gram-positive bacteria, conversely, the most common fold in hemoglobin recognition involves the NEAT domain. In Streptococcus pyogenes, the hemoglobin receptor, Shr, contains two NEAT domains that bind heme. In contrast to IsdB in *S. aureus*, an additional N-terminal domain of Shr binds to hemoglobin and thus allows the NEAT domains to abstract heme (33). In *S. aureus*, different NEAT domains are required for hemoglobin binding and heme abstraction, such as in IsdB where $IsdB^{N1}$ interacts with hemoglobin and $IsdB^{N2}$ abstracts the heme. Interestingly, in the spore forming *Bacillus anthracis*, the one NEAT domain of IsdX1 performs both of these functions (101). Unlike other gram-positive systems that attach the Isd proteins to the cell wall, *B. anthracis* excretes Isd proteins as hemophores into the extracellular milieu to bind hemoglobin and abstract heme. As a one NEAT domain protein of ~15 kDa, the 3_{10}-helix (equivalent position to L2 in $IsdB^{N1}$) (Figure 3-12A) appears to mitigate both the hemoglobin binding and heme abstraction

functions of IsdX1. Two key residues, Ser53 and Arg54, function as both Hb-binding residues and heme binding stabilizers. As shown in the sequence alignment in Figure 3-12B, there is no amino acid similarity between L2 of IsdB^{N1} and the 3_{10}-helix of IsdX1. This is not surprising due to the fact that IsdX1 probably interacts with hemoglobin at the heme-binding pocket as it only has the one domain to perform both functions.

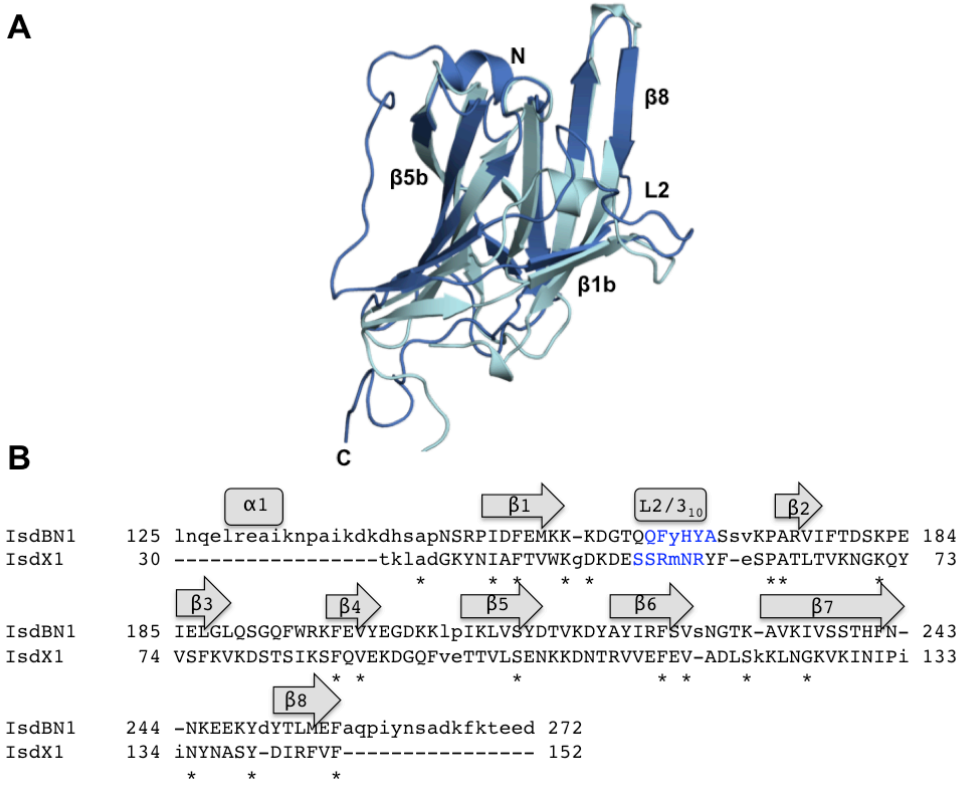


Figure 3-15. Comparison of IsdB^{N1} to IsdX1 from *Bacillus anthracis*. (A) Structural overlay of IsdB^{N1} in marine blue and IsdX1 (PDB: 3SZ6) in cyan. (B) Sequence alignment of IsdB^{N1} with IsdX1. The uppercase letters are the amino acids that have structural equivalence, while the lowercase amino acids are structurally nonequivalent. Secondary structure elements are denoted with arrows for β-strands and rectangles for helices. The amino acids in blue correspond to L2 of IsdB^{N1} and the 3_{10}-helix of IsdX1. The "*" below denotes the same amino acid. The sequence alignment was generated using DaliLite version 3.1 (87).

IsdB^{N1} adopts the characteristic immunoglobulin-like fold of other NEAT domains. The overall fold is unique from heme-binding NEAT domains with the addition of an N-terminal helix and a loop region that surrounds the first β-sheet. This feature has only been seen in the Hb-binding domains of IsdH, IsdH^{N1} and IsdH^{N2}. In addition, IsdB^{N1} shares a homologous aromatic motif (FYHY) located in loop 2, which was disordered in the structure. Amino acid substitution of Phe164 to an aspartic acid showed the inability of IsdB^{N1} to bind to metHb in the NMR. This substitution, however, did not fully inhibit heme transfer from metHb to a protein variant of IsdB^{N1LN2} F164D. Substituting and alanine for Phe164 did significantly decrease heme transfer from metHb to IsdB^{N1LN2} F164A confirming that this residue is important for metHb binding and subsequently heme transfer to the N2 domain.

CHAPTER FOUR

INVESTIGATION OF THE LINKER REGION OF ISDB

Introduction

This chapter discusses the first goal of this research project, which was to solve the 3D structure of the small region connecting $IsdB^{N1}$ and $IsdB^{N2}$, termed the linker region, $IsdB^{Linker}$. The linker domain comprises IsdB residues Glu271 through Thr338. Preliminary studies of $IsdB^{Linker}$ in Dr. Lei's lab showed that the linker region was important for heme acquisition of the N2 domain from methemoglobin *in vitro* (Zhu, et. al., manuscript under review). Elucidating the 3D structure of this small domain would help to establish a model of heme acquisition from metHb by IsdB. $IsdB^{Linker}$, unfortunately, did not appear to be fully folded in solution and formed an oligomer at millimolar concentrations so a NMR structure could not be obtained for this construct. Interactions of the linker region with the domains of IsdB, nevertheless, were explored, and the results of the NMR studies on $IsdB^{Linker}$ and their implication to the function of IsdB are examined in this chapter.

NMR Spectroscopy

Initial ^{1}H-^{15}N correlation HSQC spectra of $IsdB^{Linker}$ did not display a wide dispersion of HN resonances, which is indicative of proteins that are unfolded. Further optimization, therefore, was conducted on $IsdB^{Linker}$. From an initial micro drop screening (102) of common NMR buffers, potassium phosphate was chosen at 50mM pH 6.8 and

with the addition of 100mM potassium chloride. ^{1}H-^{15}N HSQC spectra were recorded for $IsdB^{Linker}$ from 285 K to 306 K to determine the optimal temperature for data collection. A lower temperature of 292 K displayed the best dispersion of the NMR signals. Although the initial 2D ^{1}H-^{15}N correlation spectrum displayed approximately 60 NMR signals out of the possible 65 backbone N/H protons, the extent of the NMR assignments based on the heteronuclear (^{1}H, ^{13}C, ^{15}N) 3D NMR experiments was only 78% of assignable $^{1}H_N$ and ^{15}N resonances (51 out 65), and excluding proline residues) Fig 4-1. The majority of the unassigned resonances were from the N-terminus of the protein. The completeness of the remaining backbone resonance assignments was 84% of $^{13}C_\alpha$ (57/68), 65% of $^{1}H_\alpha$ (44/68), and 84% of $^{13}C_\beta$ (56/68). Sidechain resonances were assigned to greater than 65% completeness. Despite the lower than expected assignment coverage, an attempt at assigning protons in the 3D NOESY spectra was made. Although the initial analysis on the protein indicated that the secondary structure was predominantly α-helical, inspection of the ^{15}N-edited NOESY spectrum did not show the characteristic NOE patterns for a well folded α-helix, i.e. $d_{\alpha N}$ (i, i+3), $d_{\alpha\beta}$ (i, i+3), $d_{\alpha N}$ (i, i+4), d_{NN} (i, i+1), and d_{NN} (i, i+2) (103). The lack of assignable resonance peaks in the NOESY data along with the oligomeric state of the protein at high concentrations meant a 3D structure could not be completed for this construct.

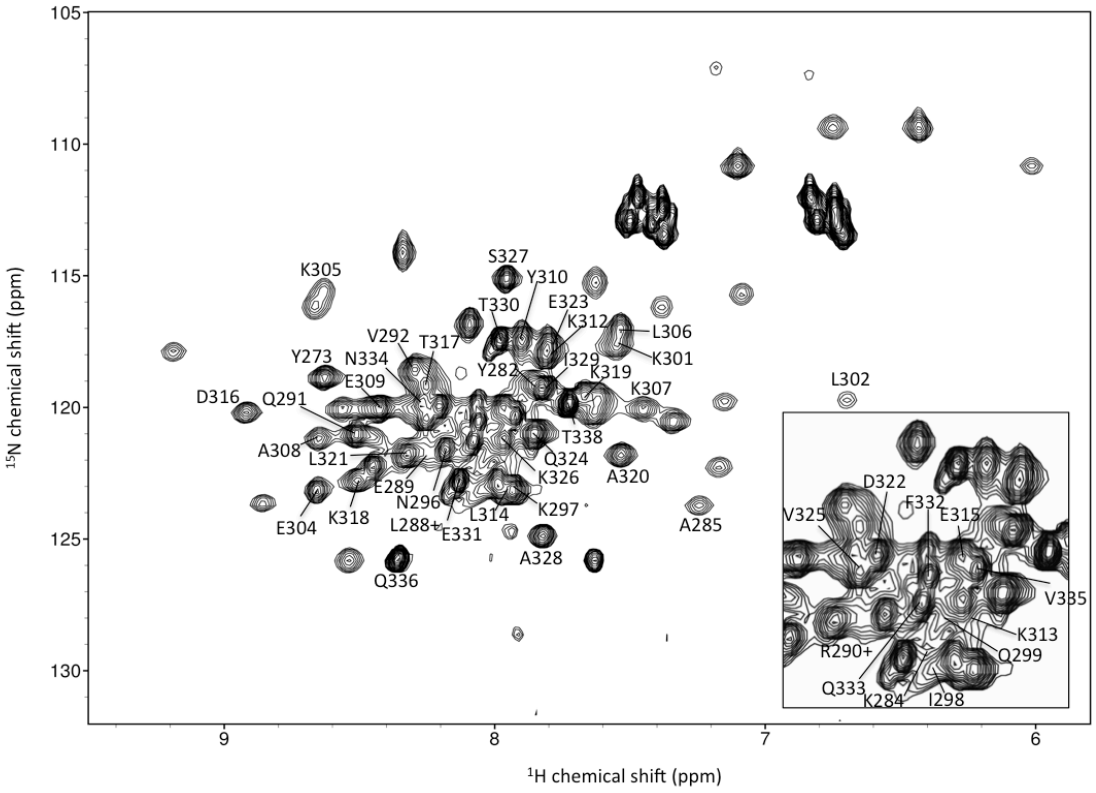


Figure 4-1. Labeled ^{1}H-^{15}N HSQC spectrum of IsdB^{Linker}. (a)2D ^{1}H-^{15}N HSQC spectrum of 2.4 mM IsdB^{Linker} in 50 mM potassium phosphate at pH 6.8, 100 mM potassium chloride, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectrum was recorded at 292 K on a Bruker DRX 600 MHz spectrometer. Backbone resonance assignments are indicated with the *one-letter* amino acid code and residue number. A ‘‘+’’ in front of residue numbers marks ^{1}H/^{15}N resonances that are too weak to be observed at the current intensity level of the displayed spectrum. Resonances from side chain NH_2 are not labeled. (b) *Insert* a close-up view of the congested middle region of the HSQC spectrum with assigned NH resonances labeled.

Characterization of the Linker Region

The overall secondary structural elements of a NEAT domain are well known; the secondary structure of the linker region, however, was not known. From the deviation of chemical shifts from random coil values for all of the assigned H_α, C_α, and C_β, the consensus chemical shift index was generated for IsdB^{Linker}. As shown in Figure 4-2, the

lack of assignment coverage for chemical shifts allowed only the prediction of one alpha helix from residues 307 to 324 based on the consensus chemical shift index. However, analysis by TALOS+ generated predicted phi and psi angles for two alpha helices from residues 290-297 and residues 304-327. To confirm the secondary structure predicted by the chemical shift data, UV circular dichroism (CD) spectroscopy revealed that the $IsdB^{Linker}$ was predominantly α-helical based on the characteristic minima at 208 and 222 nm as shown in Figure 4-3. Further characterization by size exclusion chromatography revealed that $IsdB^{Linker}$ formed a homodimer in the NMR conditions. This finding did not appear to be biologically relevant given that $IsdB^{N1\text{-}L\text{-}N2}$ does not form an oligomeric complex based on size exclusion chromatography.

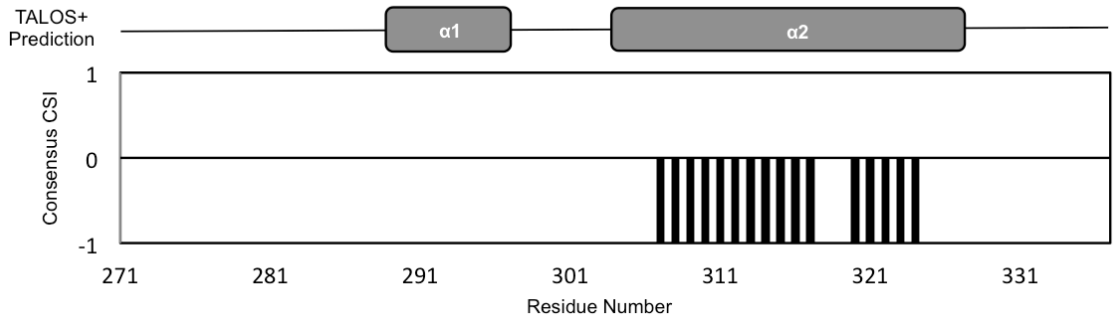


Figure 4-2. Secondary structural topology of $IsdB^{Linker}$ derived from chemical shift indexing and dihedral angle predictions from TALOS+ analysis of chemical shift data. The α-helices of $IsdB^{Linker}$ predicted by TALOS+ are shown as rectangles. The consensus chemical shift index (CSI) plot shown below $IsdB^{Linker}$ secondary structural elements is based on the CSI analysis of all H_α, C_α, and C_β chemical shifts.

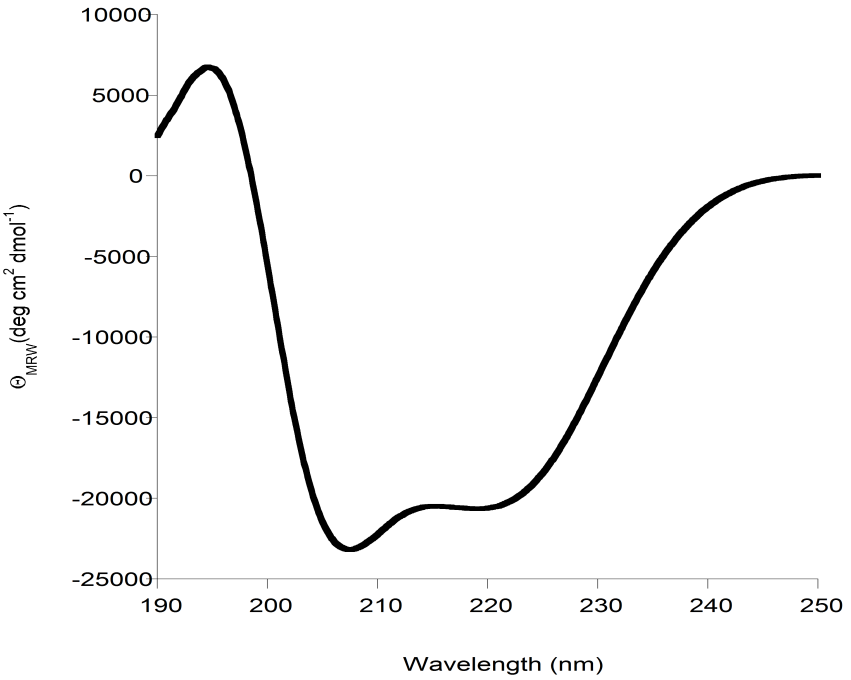


Figure 4-3. CD scan of IsdB^{Linker}. The minimums at 222 and 208 nm along with the maximum close to 190nm suggest that the secondary structure of the protein is predominantly α-helical.

Investigating Different IsdB^{Linker} Constructs

From the preliminary data collected by Dr. Lei and colleagues, the linker region is important for the acquisition of heme from metHb by N2. One question that arose during this time: are N2 and the linker region interacting? To test this, a ^{15}N-labled sample of IsdB^{Linker} was titrated with a 1:1 ratio of unlableled sample of IsdB^{N2}. As seen in Figure 4-4A, there are no discernable differences in the ^{1}H/^{15}N chemical shifts after the addition of unlabeled IsdB^{N2}. The same experiment was repeated with a ^{15}N-labled sample of IsdB^{N2} and an unlabeled sample of IsdB^{Linker}. As shown in Figure 4-4B, there are also no apparent differences in the ^{1}H/^{15}N chemical shifts of the N2 domain. If the two domains were interacting, the chemical environment of the interacting residues would be altered

and their HN resonances would possibly display a different chemical shift on the HSQC spectrum. From this data it did not appear that $IsdB^{Linker}$ and $IsdB^{N2}$ were interacting.

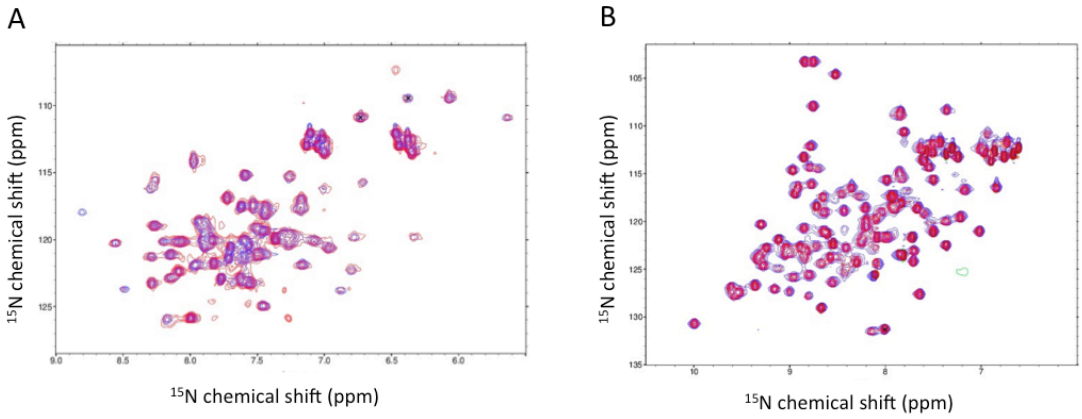


Figure 4-4. Titration experiment of $IsdB^{Linker}$ and $IsdB^{N2}$. (A)In red, a 2D ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled $IsdB^{Linker}$ after the addition of a 1:1 ratio of unlabeled $IsdB^{N2}$. The control spectrum of $IsdB^{Linker}$ is in blue. (B) In red, a 2D ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled $IsdB^{N2}$ after the addition of a 1:1 ratio of unlabeled $IsdB^{Linker}$. The control spectrum of $IsdB^{N2}$ is in blue. All spectra were recorded at 298 K and samples were prepared in 50 mM potassium phosphate at pH 6.8 with 100 mM potassium chloride and 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O.

This result, however, did not explain the preliminary data, whereby the presence of the linker region increased the ability of N2 to obtain heme from metHb. It is likely that the oligomeric state of the linker region in the NMR conditions would affect the ability of the linker to interact with N2. As opposed to the kinetic and thermodynamic experiments performed in Dr. Lei's lab where the protein concentrations are very small, in the micomolar range. In the NMR conditions, the linker region needed to be covalently bound to N2 to show significant chemical shift differences. As shown in Figure 4-5, $IsdB^{Linker\text{-}N2}$, a construct containing the IsdB residues Glu271 through Asn458, there were

significant changes in the $^{1}H/^{15}N$ chemical shifts of the HSQC spectrum (in red) when compared to the spectra of $IsdB^{N2}$ (in blue) and $IsdB^{Linker}$ (in black).

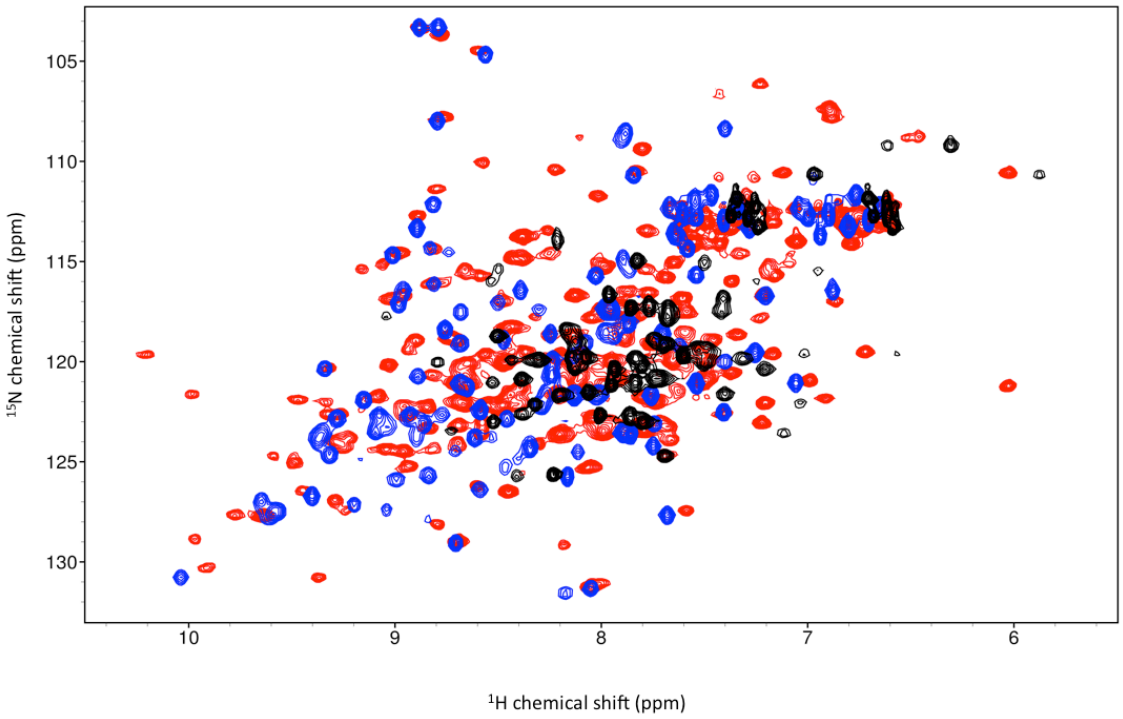


Figure 4-5. Overlay of ^{1}H-^{15}N HSQC spectra of $IsdB^{Linker}$, $IsdB^{N2}$, and $IsdB^{Linker-N2}$. The 2D ^{1}H-^{15}N HSQC of $IsdB^{L-N2}$ is shown in red, $IsdB^{N2}$ in blue, and $IsdB^{Linker}$ in black. All protein samples were 0.28 mM in 50 mM $KHPO_4/KH_2PO_4$, pH 6.8, 100 mM KCl, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectra were recorded at 298 K.

In order to determine how much of the linker region might interact with $IsdB^{N2}$, two constructs of varying length of $IsdB^{Linker-N2}$ were engineered. $IsdB^{20L-N2}$ contained the 20 C-terminal amino acids of the linker region covalently attached to N2 and $IsdB^{35L-N2}$ contained the C-terminal half of the linker residues attached to N2. As shown in Figure 4-6, an overlay of the ^{1}H-^{15}N HSQC spectra of these two constructs did not display significant differences despite the difference of 15 residues. If the additional 15 residues

in IsdB^{35L-N2} were well folded there would be an additional 15 NMR signals in the HSQC. Instead, only three additional signals appear for IsdB^{35L-N2}. Furthermore, the addition of the 20 residues of the linker did cause some shifts in the resonances when compared to the N2 spectra. This could point to the C-terminal end of the linker region interacting with the N2 domain as this region provided the greatest resonance assignment coverage in the initial NMR studies of IsdB^{Linker}.

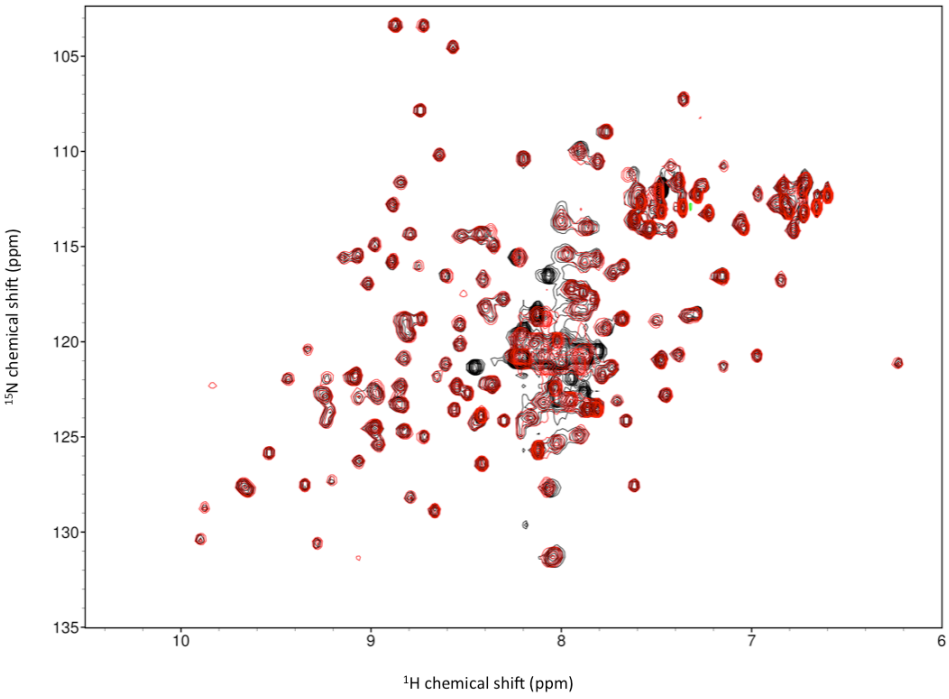


Figure 4-6. Overlay of ^{1}H-^{15}N HSQC spectra of IsdB^{35L-N2} and IsdB^{20L-N2}. The ^{1}H-^{15}N HSQC spectrum of IsdB^{35L-N2} is in black and the spectrum of IsdB^{20L-N2} is in red. Protein samples were at 0.5 mM in 50mM sodium phosphate at pH 6.8 and 100 mM sodium chloride, and 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectrum was recorded at 298 K on a Bruker DRX 600 MHz spectrometer.

To see if there were interactions of the linker region with the N-terminal NEAT domain of IsdB, a construct of IsdB$^{N1-Linker}$ was engineered. An overlay of the ^{1}H-^{15}N

HSQC of IsdB^{N1} onto the HSQC spectra of IsdB$^{N1\text{-}Linker}$ (Figure 4-7) revealed that there were no $^{1}H/^{15}N$ chemical shift differences for IsdB^{N1}. The same is true if the HSQC of the Linker region is overlaid with the HSQC of IsdB$^{N1\text{-}Linker}$. This result was unexpected due to the synergistic relationship between the domains of IsdB and their ability to abstract heme.

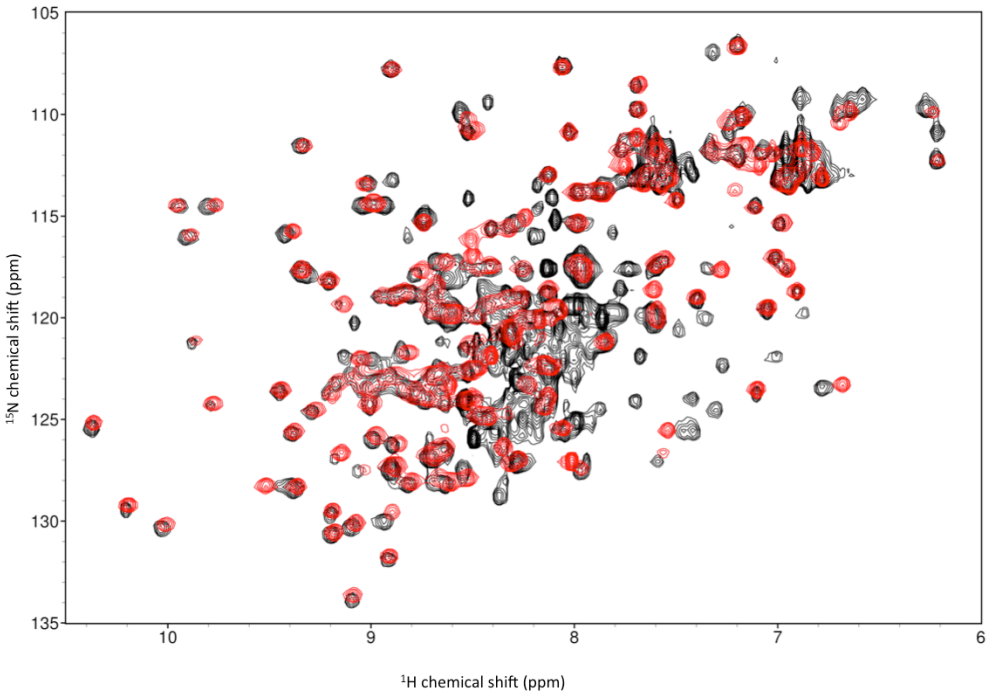


Figure 4-7. Overlay of ^{1}H-^{15}N HSQC spectra of IsdB^{N1} and IsdB$^{N1\text{-}Linker}$. The 2D ^{1}H-^{15}N HSQC spectrum of a 1 mM sample of IsdB^{N1} in red was overlaid with the HSQC spectrum of a 1 mM sample of IsdB$^{N1\text{-}Linker}$ in black. Both samples were 50 mM $NaHPO_4/NaH_2PO_4$, pH 6.8, 100 mM NaCl, 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectrum was recorded at 298 K on a Bruker DRX 600 MHz spectrometer.

Comparison to the Linker
Region of the IsdH Homolog

The IsdH protein in *S. aureus* contains three NEAT domains. The linker region that connect the second and third NEAT domains shares 72% sequence identity with $IsdB^{Linker}$ (Figure 4-8). Spirig and colleagues (98) solved the structure of $IsdH^{Linker}$ using NMR spectroscopy. The structure revealed a three-helix bundle held together by hydrophobic residues between the helices. An almost complete backbone resonance assignment for $IsdH^{Linker}$ indicated that it was well folded in solution. This was in contrast to $IsdB^{Linker}$, which did not appear to be completely folded leading to an incomplete backbone resonance assignment. Also, $IsdB^{Linker}$ formed a homodimer in at the high concentrations needed for NMR. Although the two linker regions share significant sequence identity, upon further inspection, $IsdB^{Linker}$ contains more hydrophilic lysine residues than $IsdH^{Linker}$ and their theoretical isoelectric points are drastically different. Calculation of the pI with ProtParam estimates an isoelectric point of ~5 for $IsdH^{Linker}$ where as the pI for $IsdB^{Linker}$ is estimated to be 9. Attempts to optimize the NMR conditions for $IsdB^{Linker}$ by changing salt concentration and/or the pH did not improve the spectra. Also, the extra lysine residues may not allow for the same helical bundle to form for $IsdB^{Linker}$. Although IsdB and IsdH appear to be similar, evidence suggests that the linker region of IsdB may be different than that of IsdH. As shown in the heme transfer studies presented in Chapter 3, $IsdB^{N1-L}$ mixed with $IsdB^{N2}$ is capable of abstracting heme from metHb at an increased rate. This was not the same result for IsdH, where $IsdH^{N2-L}$ with $IsdH^{N3}$ did not abstract heme from metHb (98). Also, from the published crystal structure of $IsdH^{N2LN3}$ bound to metHb, the linker region does not interact with metHb

(93). This is in contrast to size exclusion chromatography studies done by Dr. Tripet in the Copié laboratory, which showed that $IsdB^{N1}$ did not bind to metHb in the low concentrations used for size exclusion chromatography, however $IsdB^{N1-L}$ did bind to metHb, indicating that there is interaction taking place between the linker region and metHb. This might be an explanation for the difference in pI's of the two linker regions, where there is a unique segment if $IsdB^{Linker}$ that interacts with metHb, and that $IsdH^{Linker}$ lacks.

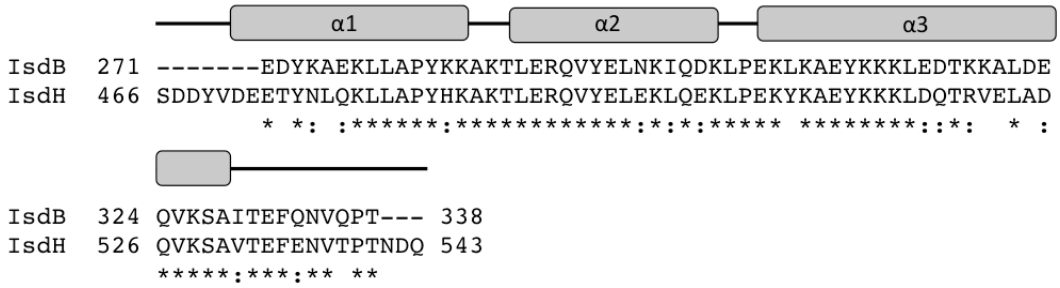


Figure 4-8. Sequence alignment of $IsdB^{Linker}$ and $IsdH^{Linker}$. The secondary structural elements of $IsdH^{Linker}$ are shown as rectangles above the alignment. An '*' represents the same amino acid and the ':' represents a similar amino acid. Alignment was performed using ClustalW2 (104) with the sequence of $IsdH^{Linker}$ acquired from the PDB entry 2LHR.

CHAPTER FIVE

THE SOLUTION STRUCTURE OF APO-ISDB^{N2}

Characterization of NEAT 2

There have been multiple attempts to solve the NMR solution structure of N2. The first construct attempted was the wild type N2 sequence N338-N458, cloned in a pET21-d vector obtained from Dr. Lei's group. The initial NMR ^{1}H-^{15}N spectrum revealed a wide dispersion of proton resonances from 6.5 ppm to 10 ppm and sharp rounded signals typical of a well-folded protein. The standard heteronuclear 3D NMR experiments were collected for this construct. There were a lack of resonances, unfortunately, and it appeared that the original $^{13}C/^{15}N$ labeled sample had degraded so another sample was produced. The new sample, although clear to the eye, appeared to be aggregated based on the initial NMR data collection. In order to optimize the protein expression, a new construct was engineered into the pET46 vector with an N-terminal His_6 tag. The resulting protein still possessed the same aggregation properties of its predecessor; so possible amino acid substitutions were explored. One of the reasons the protein might form a soluble aggregate would be a solvent exposed hydrophobic residue on the surface of the protein. From investigation into a predicted structure of IsdB, a likely candidate was Ile438. This residue was predicted to be in the loop region between β-strands 7 and 8, which could leave it exposed to solvent. The Ile was changed to an alanine and resulting protein was less aggregated with the addition of the detergents, 0.025% Tween 20 and 1mM CHAPS. A full standard NMR data set was collected on this

protein sample and assignment of the protein backbone was attempted. Many of the original backbone heteronuclear (^{1}H, ^{15}N, ^{13}C) NMR assignments of apo-IsdB^{N2} could be transferred to the data analysis of the new construct of apo-IsdB^{N2} I438A. A few more resonance assignments were added, nevertheless, this was not the optimal sample for structure determination and another construct was cloned. The additional change of Met362 to an alanine appeared to make a difference in the protein sample. As shown in Figure 5-1, the 2D ^{1}H-^{15}N correlation NMR (HSQC) spectrum of the new apo-IsdB^{N2} M362A/I438A construct (in blue) was almost identical to that of the wild type IsdB^{N2} spectrum (in black) and even displayed a few more signals. The NMR solution structure of apo-IsdB^{N2} M362A/I438A is presented in this chapter.

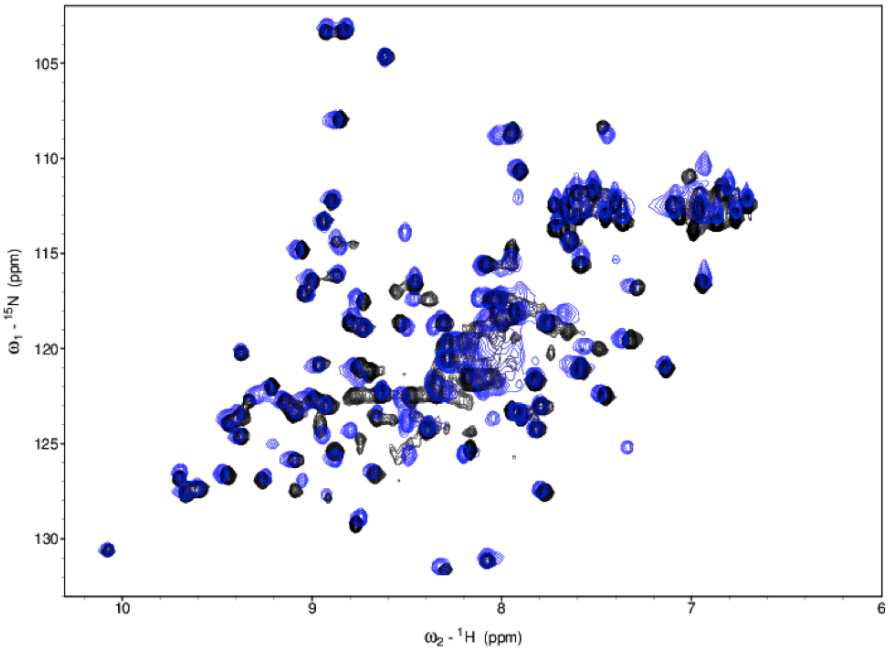


Figure 5-1. Overlay of the ^{1}H-^{15}N HSQC spectra of wild type apo-IsdB^{N2} and mutated apo-IsdB^{N2} M362A/I438A. The spectrum of the wild type apo-IsdB^{N2} is in black and the spectrum of apo-IsdB^{N2} M362A/I438A is in blue. The data was collected at 306 K in 50 mM sodium phosphate at pH 6.8, with 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O.

Extent of NMR Assignments

From standard two and three dimensional heteronuclear NMR experiments, 93 of the possible 118 $^{1}H_{N}$ resonances were assigned. Ten of the missing HN assignments corresponded to residues Thr387 to Met396. From the secondary structure prediction, this segment corresponded to part of β4 and the loop that connects β3 and β4. The structure of a NEAT domain appears to be well folded so the reason a large segment of backbone assignments are missing is not clear. The NEAT domains of *S. aureus* share a conserved tryptophan residue that provides hydrophobic packing near the heme-binding pocket. In $IsdB^{N2}$, Trp392 resides in this unassigned stretch of amino acids. As shown in Figure 5-2, there is no characteristic NMR signal for the tryptophan indol ring around 10.5 ppm in the proton chemical shift dimension. Despite these missing residues, the rest of the backbone assignment was completed with 88% of the $^{13}C_{\alpha}$ and $^{1}H_{\alpha}$, and 85% of the $^{13}C_{\beta}$. The side chain assignment was greater than 72% of the identified residues. Of the two changed residues, M362A was identified, while the I438A residue could not be assigned. This same residue position, Ile438, could not be assigned in the original wild type apo-$IsdB^{N2}$ NMR assignment either.

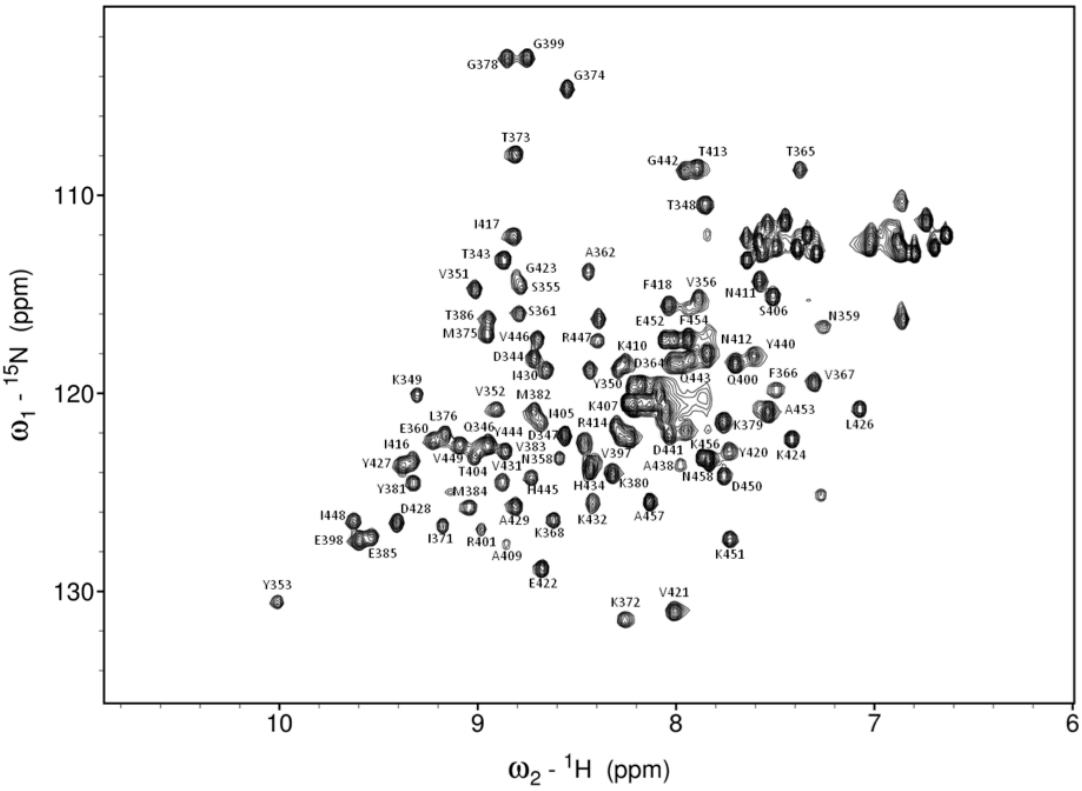


Figure 5-2. Labeled ^{1}H-^{15}N HSQC spectrum of apo-IsdB^{N2} M362A/I438A. The 2D ^{1}H-^{15}N HSQC spectrum of 1 mM IsdB^{N2} in 50 mM sodium phosphate at pH 6.8, with 1 mM EDTA, 0.1 mM PMSF, 0.01 % sodium azide, in 95 % H_2O and 5 % D_2O. The spectrum was recorded at 306 K on a Bruker DRX 600 MHz spectrometer. Backbone resonance assignments are indicated with the *one-letter* amino acid code and residue number. Resonances from side chain NH_2 are not labeled.

Structure Determination

A total of 1194 NOE assignments were extracted from the ^{15}N- and ^{13}C-edited NOESY spectra and used as distance restraints to calculate a preliminary ensemble of apo-IsdB^{N2} M362A/I438A structures using the standard annealing protocol in CYANA 2.1. The structure calculations additionally included 188 dihedral angles derived from chemical shift data using the program TALOS+ and 120 hydrogen-bond restraints (2

restraints per NH pair) derived from the $^{1}H/^{2}H$ solvent exchange experiments. The chosen 20 conformers of the preliminary ensemble for apo-IsdB^{N2} M362A/I438A have no NOE or dihedral angle restraint violations greater than 0.5 Å or 5°, respectively. The calculated ensemble needs further structural refinement, including water refinement with AMBER9. The root mean square deviation (RMSD) of backbone atoms of the secondary structure elements from the mean structure is 1.0 Å respectively. Analysis of the dihedral angles of the ordered residues in secondary structure elements (Figure 5-3) revealed that 81% of residues are in the most favored conformations on the Ramachadran plot and 19% are in additionally allowed regions. As shown in Figure 5-3 there are a few residues that sporadically appear in disallowed angle regions and these areas will need further inspection before the final water refinement protocol. The final restraint and structural statistics for the latest calculation of apo-IsdB^{N2} M362A/I438A are listed in Table 5-1.

As previously determined for other NEAT domains and for the crystal structure of holo-IsdB^{N2} (59), apo-IsdB^{N2} M362A/I438A adopts the characteristic immunoglobulin-like fold with two antiparallel sheets comprised of four β strands each. The first sheet includes the strands β1 (Y350-E354), β8 (D439-V449), β7 (L426-V435), and β4 (W392-V397) while the opposite face incorporates the strands β2 (K372-L376), β3 (K379-T386), β6 (N412-P418), and β5 R403-D408. There are two small α helices, α1 (M363-F366) that lies between β1 and β2, and α2 (K451-F454) that resides at the C-terminus (Figure 5-4A). The two amino acid substitutions that were employed to improve sample quality, M362A and I438A, are highlighted in Figure 5-4B. Both of these bulky hydrophobic residues reside near the surface of the protein creating a possible point of aggregation between the

domains in solution. Similarly, I438 created a crystal contact with the neighboring IsdB^{N2} domain's I438 in the published holo-IsdB^{N2} crystal structure (59).

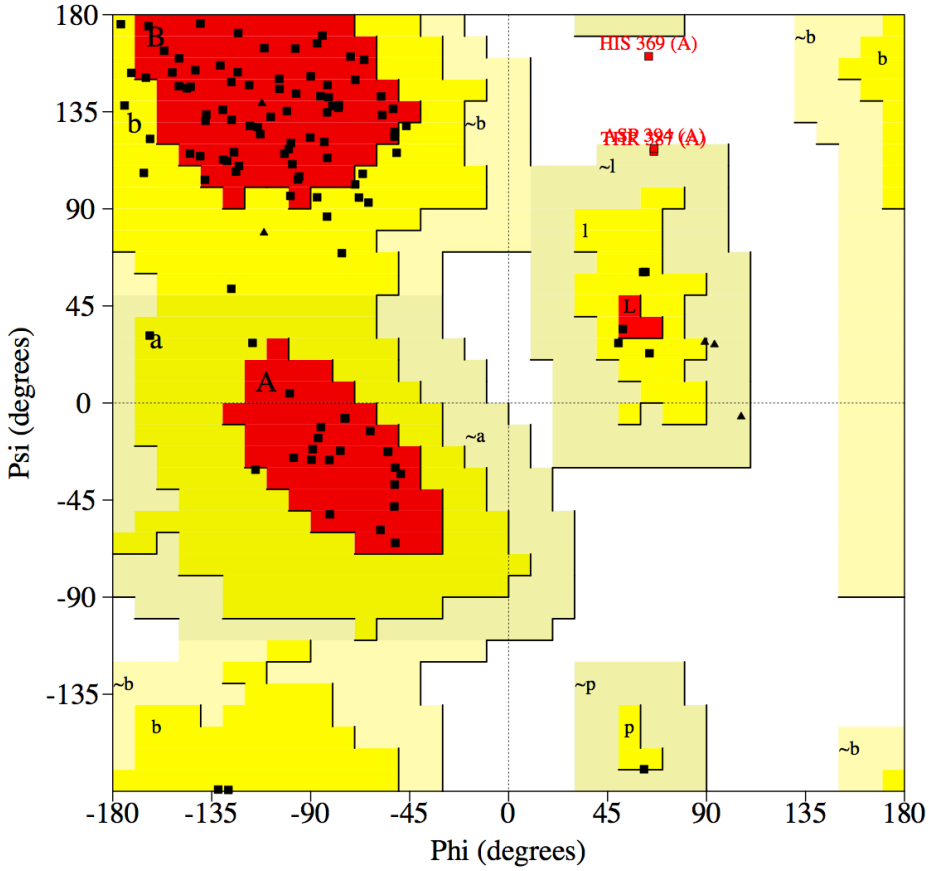


Figure 5-3. Ramachandran plot of apo-IsdB^{N2} M362A/I438A. The core areas in red correspond to the most favored regions (A, B, L), yellow areas to additionally allowed regions (a, b, l, p), tan areas to generously allowed regions (~a, ~b, ~l, ~p), and white areas to disallowed regions. Residues are represented by black squares in favored and additionally allowed regions, and residues in generously allowed or disallowed regions in red. His369, Thr387, and Asp394 are in generously allowed or disallowed regions. These three residues correspond to unassigned amino acids. The Ramachandran plot was generated by PROCHECK-NMR for the lowest energy structure of the NMR ensemble.

Table 5-1. NMR structure statistics for apo-IsdB^{N2} M362A/I438A

Constraints for final structure calculations	
Total NOE distance restraints	1194
Intraresidue	388
Sequential (| *i -j* |= 1)	367
Medium range (1 < | *i -j* | < 5)	105
Long range (| *i -j* | ≥ 5)	334
Dihedral angle restraints[a]	
ϕ angles	94
ψ angles	94
Hydrogen bonds restraints[b]	60
Structure statistics (20 conformers)	
CYANA target function (Å^2)	1.43
Residual distance violations	
Number > 0.5 Å	0
Residual dihedral angle violations	
Number > 5^o	0
Ramachandran plot statistics (%)[c,d]	
Residues in most favored regions	81.3
Residues in additionally allowed regions	18.7
Residues in generously allowed regions	0.1
Residues in disallowed regions	0
Average rmsd to mean structure (Å)[d]	
Protein backbone	0.94 ± 0.15
Protein heavy atoms	1.62 ± 0.19

[a] ϕ and ψ angles were derived from the program TALOS+, based on the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$ and ^{15}N chemical shifts. [b] Two distance restraints between the amide and carbonyl group atoms (NH-O = 1.8-2.8 Å, N-O =2.5-3.8 Å) were used for the hydrogen bond restraints. [c] Ramachandran plot analysis performed with PROCHECK. [d] Residues 350-354, 363-366, 372-376, 379-386, 403-408, 412-418, 426-435, 439-449, 451-454 were used. [e]

The preliminary NMR ensemble of the 20 lowest energy structures is shown in Figure 5-4C. The β-strands near the core of the protein are well defined as displayed in the narrow overlay of the ensemble in those regions. The loops, however, are not as well defined and it is in those regions that a few more NOEs would help improve the overall structure quality.

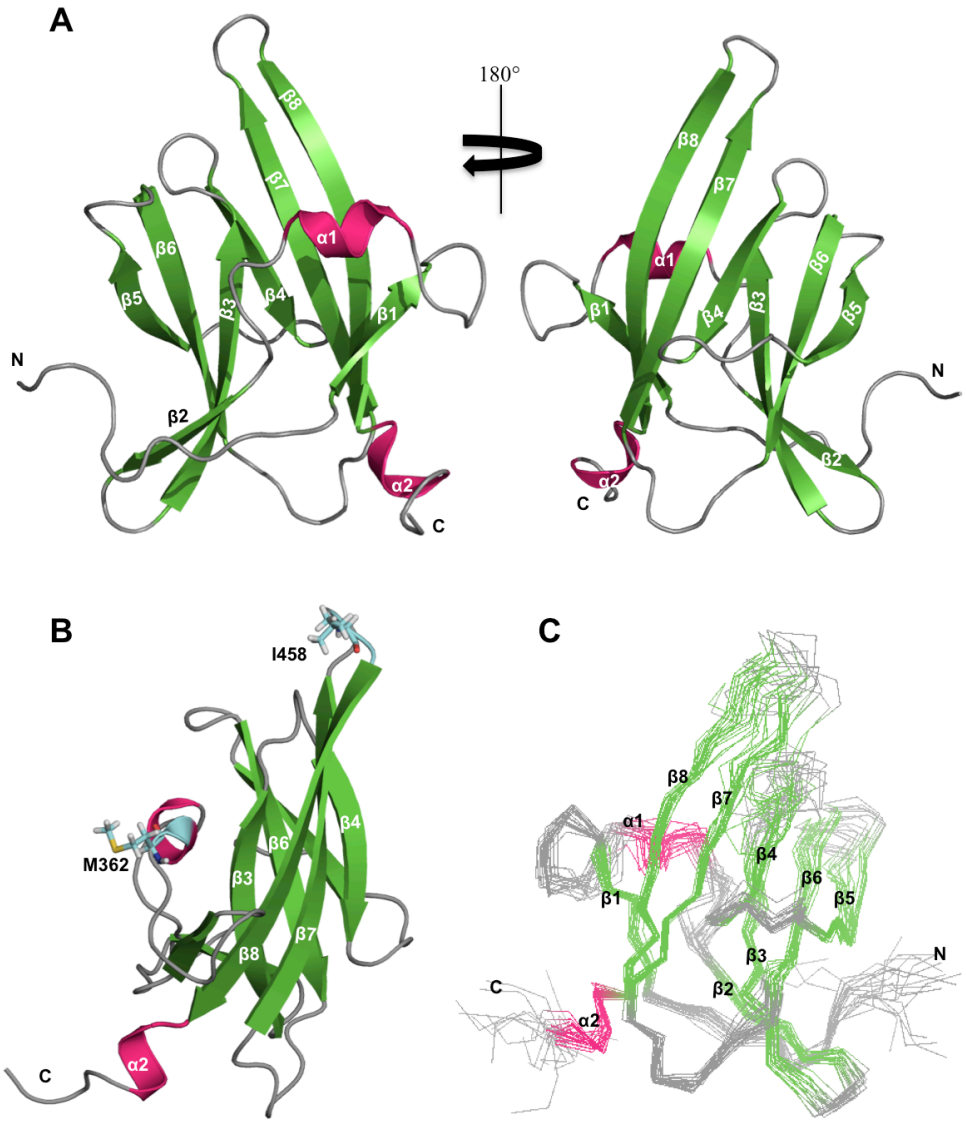


Figure 5-4. The NMR solution structure of apo-IsdB^{N2} M362A/I438A. (A) Ribbon representation of the lowest energy structure of IsdB^{N2} selected from the NMR ensemble. The secondary structure elements are labeled and color-coded with the β-strands in marine green, the α-helix in magenta, and the loops in gray. (B) Met362 and Ile438 are highlighted by stick representation. Both of these amino acids were mutated to alanine to improve sample aggregation. (C) The NMR ensemble of the 20 lowest energy structures of apo-IsdB^{N2} M362A/I438A. The backbone heavy atoms (N, C_α, and C') of the secondary structural elements 350-354, 363-366, 372-376, 379-386, 403-408, 412-418, 426-435, 439-449, 451-454 were superimposed using the program Theseus version 2.0.6. The β-strands are in green, the α-helices in pink, and the loops in gray.

Comparison to the Published Holo-IsdB^{N2} Crystal Structure

During the multiple attempts at the NMR structure determination of apo-IsdB^{N2}, the crystal structure of holo-IsdB^{N2} was published by Gaudin, et. al. (59). There were two unique findings from the crystal structure. The first was a unique Tyr/Met heme coordination. The conserved YXXXY motif of the NEAT domain provides the tyrosine axial ligand for the heme iron (Y440) and another conserved tyrosine hydrogen bonds to a preceeding tyrosine (Y444). In the solved crystal structure, however, M362 provided the other axial ligand for the heme iron. The second unique feature was the small C-terminal α-helix. The functional significance of this helix has yet to be shown, but it could be related to the longer C-terminal region of IsdB that anchors the protein to the cell wall. As shown in Figure 5-5, apo-IsdB^{N2} M362A/I438A aligns with holo-IsdB^{N2} (PDB 3RTL) with an RMSD of 2.2Å over 113 aligned residues. The apo- and holo-forms appear very similar; nevertheless, the heme-binding pocket does appear to have slight differences, especially in β7 and β8 (Figure 5-5B). This result corresponds to the 2D ^{1}H-^{15}N correlation HSQC spectral differences between apo- and holo-IsdB^{N2} as displayed in Figure 5-6.

Structural Homology

Using the DaliLite web server (87), the closest homologs to the structure of apo-IsdB^{N2} M362A/I438A were other *S. aureus* NEAT domains, including the crystal structures of apo- (PDB 3VUA) and holo-IsdH^{N3} (PDB 4IJ2) (93) with an RMSD of 2.0Å and 58% sequence identity over 113 aligned residues and the NEAT domain of

IsdA (PDB 2ITE) (58) with an RMSD of 2.4Å aligned with 111 residues and 22% identity (Figure 5-7).

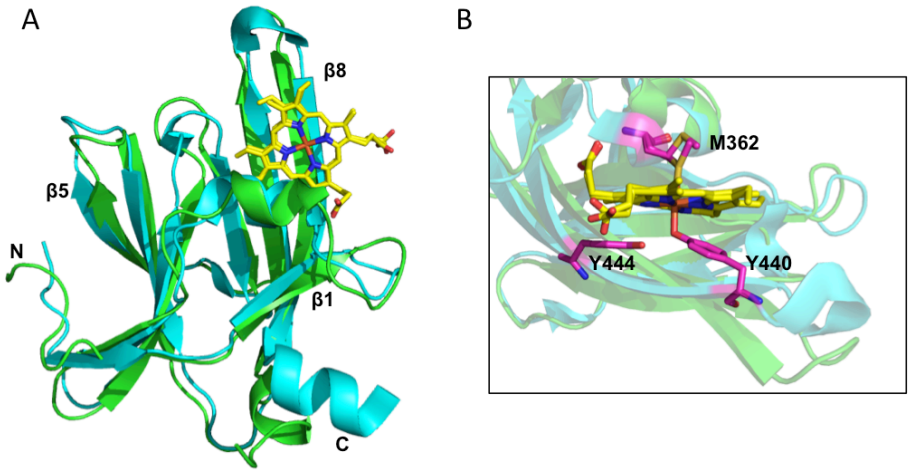


Figure 5-5. Structure overlay of apo-IsdB^{N2} M362A/I458A and holo-IsdB^{N2}. (A) An overlay of the first model of apo-IsdB^{N2} M362A/I458A in green with the crystal structure of holo-IsdB^{N2} in cyan. The heme group is in stick representation in yellow, with the nitrogen in blue, oxygen in red, and iron in orange (PDB 3RTL). (B) A close-up of the heme-binding pocket with the axial ligands, Y440 and M362, and the conserved Y444 in magenta.

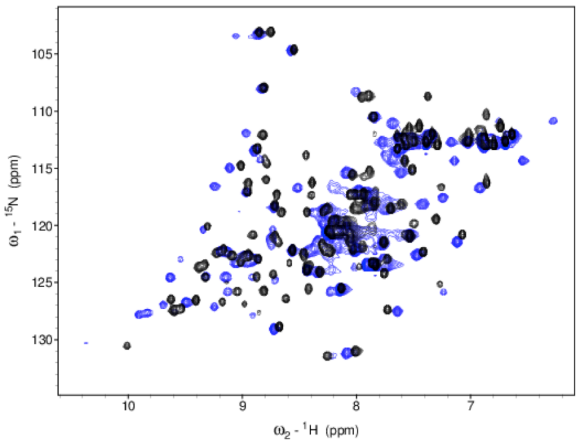


Figure 5-6. Overlay of the ^{1}H-^{15}N HSQC spectra of apo-IsdB^{N2} M362A/I438A and holo-IsdB^{N2} M362A/I438. The apo- spectrum is in black and the holo- spectrum is in blue.

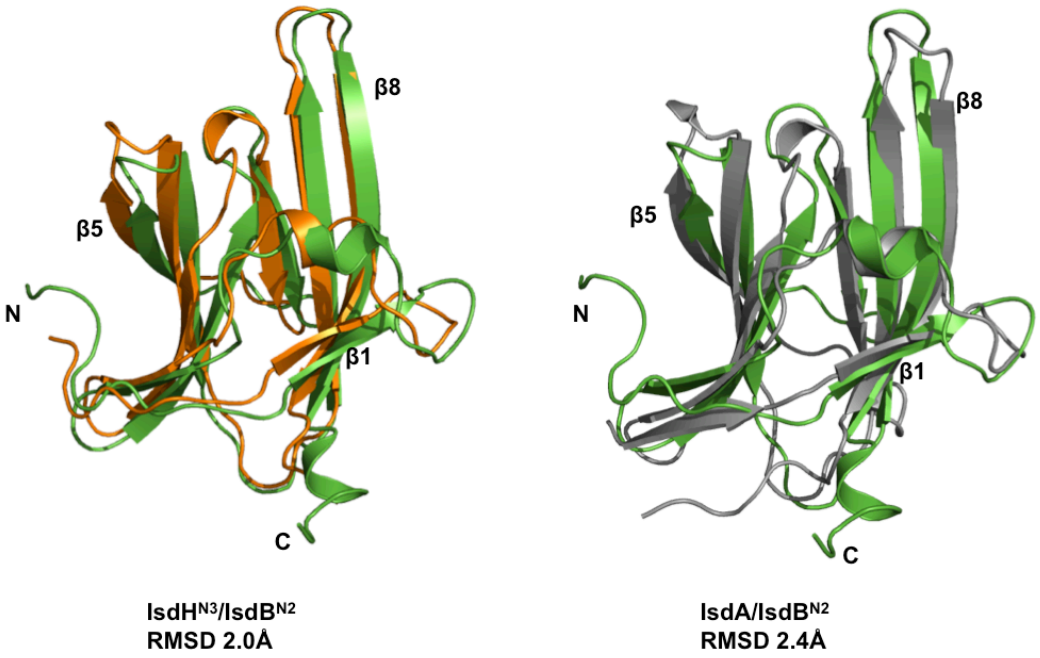


Figure 5-7. Comparison of the structure of apo-IsdB^{N2} M362A/I438A with (Left) the crystal structure of apo-IsdH^{N3} (PDB 3VUA) (Right) the crystal structure of IsdA (PDB 2ITE). IsdB^{N2} is colored green.

In addition to NEAT domain of *S. aureus*, other structural homologs were NEAT domains from *Bacillus anthracis* and *Streptococcus pyogenes*. These included the NEAT domain from IsdX1 (PDB 3SIK) (101) and the fifth NEAT domain of IsdX2 (4H8P) (105), the hemophores of *B. anthracis*. The structures aligned with 3.3Å and 3.0Å RMSD, respectively, and share 22% and 25% identity. Also, the crystal structure of Shp (2Q7A) (106) from *S. pyogenes* aligns with 3.5Å with only 10% identity. Shp uses two methionine axial ligands to coordinat the heme iron, which was a unique feature for a NEAT domain until the structure of holo-IsdB^{N2}.

CHAPTER SIX

CONCLUDING REMARKS

Staphylococcus aureus is part of the natural flora of ~30% of the world's population. It is opportunistic, so it will invade the body if given an open wound. Once inside, *S. aureus* is responsible for skin afflictions such as cellulitis and impetigo, but can infect the bloodstream and deep tissues causing bacteremia, endocarditis, and osteomyelitis. An increase in antibiotic resistant strains has affected the ability to treat these infections. The pathogenicity of *S. aureus* can be attributed to its plethora of strategies it employs to evade the innate immune system. Surface proteins, such as Protein A and Clumping factor B disguise the bacterium from neutrophils thereby avoiding phagocytosis, secreted staphylokinase proteins inhibit blood clotting of the host and thus increase the ability of *S. aureus* to travel through the body, and various adhesion proteins coating the surface allow *S. aureus* to adhere to many tissue types. The limiting factor for growth within the host is procuring nutrients, in particular the highly regulated nutrient, iron.

Iron is an essential nutrient for most of life and is required for a variety of biological processes including DNA biosynthesis, respiration, oxygen transport, and gene regulation. The ability to switch oxidation states from Fe^{2+} to Fe^{3+} allows it to serve as a catalytic center for many enzymes. Also, its ligand environment in the protein can control the iron's redox potential. Iron can be toxic to cells as Fe^{3+} is insoluble at physiological

conditions and Fe^{2+} can be used to generate hydroxyl radicals. Therefore, iron is tightly regulated in the human body and is usually found bound to proteins.

Both Gram-negative and Gram-positive bacteria have developed elaborate systems to acquire iron from the host. In order to acquire heme, Gram-negative bacteria either release hemophores or have protein receptors to heme binding proteins and abstract the heme. The heme is transported through the receptor by energy acquired by the TonB complex. The heme is transported through the periplasm, by periplasm binding proteins, and then transported through the inner membrane by an ATP-binding cassette (ABC) transporter. The Hb-receptor in Gram-negative adopts a β-barrel structure that translocates through the outer membrane. It binds to hemoglobin with extended loop structures such as ShuA from *Shigella dysenteriae*.

In Gram-positive bacteria, iron must be transported through the thick peptidoglycan cell wall, comprised of crosslinked sugars and peptides. To transport heme, Gram-positive bacteria use a protein relay system. In S. aureus, the iron-regulated surface determinant (Isd) pathway comprises nine proteins. The ferric uptake repressor protein (Fur) regulates transcription for the *isd* operon, by repressing transcription when iron levels in the cell are sufficient and allowing transcription when iron levels are low.

IsdB, IsdH, IsdA and IsdC are surface proteins that are attached to the cell wall by sortase enzymes. IsdB and IsdH are fully exposed on the staphylococcal cell surface, whereas IsdA is partially buried and IsdC is fully buried into the cell wall. In addition to their iron transport functions, the cell surface Isd proteins are protein receptors and adhesion molecules. IsdE is a lipoprotein that interacts the ATP-binding cassette (ABC)

transporter comprised of IsdF (permease) and possible FhuC (ATPase). The ABC transporter transport the heme through the membrane and into the cytoplasm where it can be degraded by the heme oxygenases, IsdG and IsdI to release iron for use by the cell.

The crucial first step in the Isd pathway is the binding of Hb by the Hb-receptor IsdB. IsdB, IsdH, IsdA, and IsdC all contain one or more copies of a near iron transporter (NEAT) domain. In IsdB, there aere two NEAT domains connected by a small ~70 amino acid linker region. The N-terminal NEAT domain was believed to bind Hb, while the C-terminal NEAT domain binds heme. The work presented in this thesis focused on understanding molecular mechanisms involved in the interactions of the first protein of the heme acquisition pathway of *S. aureus*, IsdB with its target protein, hemoglobin. When this project began, there was little structural knowledge of IsdB, and even less known about the mechanism by which IsdB interacted and abstracted heme from methemoglobin. The aims of this project were to determine the structures of the domains of IsdB and investigate their interactions with Hb and each other.

In Chapter 3, the solution structure of the first NEAT domain of IsdB was presented. Initial purification attempts of the predicted first NEAT domain of IsdB, residues 140-270, showed signs of aggregation. Inspection of the amino acid sequence and secondary structure prediction suggested an additional element of secondary structure 15 residues further toward the N-terminus. The expression of the new construct of $IsdB^{N1}$, residues 125-272 showed an increase in the solubility of the protein in the NMR conditions. This longer construct was used to determine the 3D solution structure of $IsdB^{N1}$. The 3D heteronuclear NMR data collected on the recombinant $IsdB^{N1}$ protein

enabled resonance assignments of 93 % all assignable $^{1}H_{N}$ and ^{15}N resonances. There were ten backbone amide resonances that could not be assigned due to significant resonance broadening and spectral overlap. Four of the missing assignments corresponded to aromatic residues F164, Y167, H166, and Y167.

$IsdB^{N1}$ adopts the characteristic immunoglobulin-like fold of other NEAT domains. The overall fold is unique from heme-binding NEAT domains with the addition of an N-terminal helix and a loop region that surrounds the first β-sheet. This feature has only been seen in the Hb-binding domains of IsdH, $IsdH^{N1}$ and $IsdH^{N2}$. In addition, $IsdB^{N1}$ shares a homologous aromatic motif (FYHY) located in loop 2, which was disordered in the structure. Amino acid substitution of Phe164 to an aspartic acid showed the inability of $IsdB^{N1}$ to bind to metHb in the NMR. This substitution, however, did not fully inhibit heme transfer from metHb to a protein variant of IsdBN1LN2 F164D. Substituting and alanine for Phe164 did significantly decrease heme transfer from metHb to $IsdB^{N1LN2}$ F164A confirming that this residue is important for metHb binding and subsequently heme transfer to the N2 domain.

The first goal of this research project involved studying the small linker region that connects the two NEAT domains of IsdB. Chapter 4 discusses the extent of structure determination of $IsdB^{Linker}$ and its interactions with N1 and N2. The initial structure determination by NMR for linker region proved to be difficult. Although the initial 2D ^{1}H-^{15}N HSQC spectrum displayed approximately 60 NMR signals out of the possible 65 backbone N/H protons, the extent of the NMR assignments based on the heteronuclear (^{1}H, ^{13}C, ^{15}N) 3D NMR experiments was only 78% of assignable $^{1}H_{N}$ and ^{15}N resonances

and the completeness of the remaining backbone resonance assignments was only 84% of the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ and 65% of side chains. Although an attempt was made to assign protons in the 3D NOESY experiments, there were not enough assignable resonance peaks and a structure was not determined. Some information about the secondary structure of the linker region could be gleaned from the backbone chemical shift assignments and circular dichroism (CD) spectroscopy. Both of these methods indicated that the linker region was predominantly α-helical. Unfortunately, in the high concentrations of protein needed for the NMR studies the linker region formed a homodimer based on size exclusion chromatography.

Preliminary studies of the domains of IsdB revealed that the linker region was important for heme acquisition of N2 from metHb. NMR titrations were performed where varying amounts of ^{15}N-labeled IsdB^{Linker} were added to unlabeled IsdB^{N2} and the H-N resonances were monitored via 2D ^{1}H-^{15}N HSQC experiments. The same experiment was duplicated with unlabeled IsdB^{Linker} protein with ^{15}N-labeled IsdB^{N2}. The resulting HSQC spectra did not indicate interaction between the two domains. This was in contrast to the preliminary data collected in Dr. Lei's laboratory and the most reasonable answer to this was that if the linker was forming a dimer in NMR conditions, then it might not be able to interact with IsdB^{N2} as it did in conditions used for the heme transfer studies. A construct of the linker region covalently bound to N2 was engineered and expressed and the resulting HSQC spectrum did show significant resonance changes as compared to IsdB^{N2}. This was indicative of the two domains interacting when they were covalently bound. To see if the linker region would interact with IsdB^{N1}, a construct of the two

covalently bound was also engineered and expressed. The resulting HSQC spectrum did not show significant resonance shifts as compared to $IsdB^{N1}$. This is also in contrast to the kinetic data that showed N1 covalently bound to the linker region was capable of enhancing the rate of which N2 extracted heme from metHb. One explanation to this would be that the linker region must also interact with metHb. From the kinetic heme transfer experiments; it would appear that the interaction with metHb takes place in the N-terminal region of the linker, as the two truncated $IsdB^{L-N2}$ constructs did not enhance heme transfer from metHb.

In Chapter 5, the solution structure of apo-$IsdB^{N2}$ was presented. Due to soluble aggregation, initial attempts to solve the structure of wild type $IsdB^{N2}$ were unsuccessful. In order to prevent aggregation, two amino acids were changed Met362 and Ile438. Out of the possible 118 $^{1}H_{N}$ resonances, 93 of the $^{1}H_{N}$ resonances were assigned. Ten of the missing HN assignments corresponded to residues Thr387 to Met396, which correspond to β4 and the loop that connects strands β3 and β4. In addition, there was no resonance assignment for the conserved tryptophan residue found in all of the NEAT domains of *S. aureus.* The rest of the backbone and side chain assignment was between 72 and 88%. The structure of apo-$IsdB^{N2}$ M362A/I438A adopts an Immunoglobulin-like fold with two antiparallel β sheets comprised of four strands each. The apo structure of $IsdB^{N2}$ revealed a similar overall fold as that of published crystal structure of holo-$IsdB^{N2}$. The more significant differences in the two structures appeared in the heme-binding pocket where strands β7 and β8 do not form the same curve as the holo structure where β7 and β8 are hugging the heme moiety. These differences between apo- and holo-$IsdB^{N2}$ could be seen

in the overlaid ^{1}H-^{15}N HSQC spectra of the two proteins. The closest structural homolog to $IsdB^{N2}$ was the third NEAT domain of the cell surface protein, IsdH, which shares 58% sequence identity.

As a means to fight *S. aureus* infection, understanding its essential nutrient pathways and creating new therapeutics that could block the import of these nutrients is of great importance. In war, blocking the enemy's supply lines effectively causes starvation. It was the goal of this research project to further the understanding of the heme supply line, the Isd pathway, for *S. aureus*, and in particular the interactions of the first protein in the pathway, IsdB. Although the two NEAT domains of IsdB adopt a similar immunoglobulin-like fold, there are essential differences in their structures that allow $IsdB^{N1}$ to interact with metHb and $IsdB^{N2}$ to abstract heme. The loop region (L2) of $IsdB^{N1}$ contains an important aromatic motif, conserved only in Hb-binding NEAT domains, that imparts the ability of $IsdB^{N1}$ to interact with metHb. This interaction with metHb was disrupted by amino acid substitution of this aromatic motif, and in particular Phe164. $IsdB^{N1}$ is not the only part of IsdB that makes contact with metHb. The covalent linkage of $IsdB^{N1}$ with the linker region also interacts and enhances N2's ability to obtain heme from metHb, indicating the importance of this small linker region to IsdB's overall function. The structure of $IsdB^{N2}$ shows a clear heme-binding pocket that allows the N2 domain to bind heme. This pocket is not present in $IsdB^{N1}$ as L2 does not form a fully folded helix in solution, but instead is disordered. The present, yet subtle differences in the apo- form of $IsdB^{N2}$ as compared to the crystal structure of holo-$IsdB^{N2}$ are found in the β strands responsible for embracing the heme moiety when it is in the binding pocket.

The data presented here shows a similar function to IsdB's closest homolog, IsdH, however significant differences exist in the mechanism of heme abstraction. Although their linker regions share ~70% sequence homology, the linker region of IsdH does not appear to interact with metHb and does not enhance the heme transfer reaction to the N3 domain when attached to IsdH's N2 domain. Future studies into the structure of the full length $IsdB^{N1LN2}$ construct along with metHb may elucidate the important interactions occurring between the linker region and metHb that enhances the ability of IsdB to abstract heme form metHb. In addition, further work into therapeutics that target the metHb-binding interface of $IsdB^{N1}$ could help to inhibit *S. aureus* ability to obtain heme.

REFERENCES CITED

1. Todd, A. E., Orengo, C. A., and Thornton, J. M. (2001) Evolution of function in protein superfamilies, from a structural perspective. *J Mol Biol* **307**, 1113-1143.

2. Ogsten, A. (1882) Micrococcus poisoning. *Journal of Anatomy and Physiology* **16**, 526-567.

3. Klevens, R. M., Morrison, M. A., Nadle, J., Petit, S., Gershman, K., Ray, S., Harrison, L. H., Lynfield, R., Dumyati, G., and Townes, J. M. (2007) Invasive methicillin-resistant Staphylococcus aureus infections in the United States. *JAMA* **298**, 1763-1771.

4. Casey, A. L., Lambert, P. A., and Elliott, T. S. J. (2007) Staphylococci. *Int J Antimicrob Ag* **29, Supplement 3**, S23-S32.

5. Enright, M. C., Robinson, D. A., Randle, G., Feil, E. J., Grundmann, H., and Spratt, B. G. (2002) The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *PNAS*, 7687-7692.

6. Foster, T. J. (2005) Immune evasion by staphylococci. *Nat Rev Microbiol* **3**, 948-958.

7. O'Riordan, K., and Lee, J. C. (2004) Staphylococcus aureus Capsular Polysaccharides. *Clin Microbiol Rev* **17**, 218-234.

8. Deivanayagam, C. C., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Höök, M., and Narayana, S. V. (2002) A novel variant of the immunoglobulin fold in surface adhesins of Staphylococcus aureus: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *The EMBO Journal* **21**, 6660-6672.

9. Zapotoczna, M., Jevnikar, Z., Miajlovic, H., Kos, J., and Foster, T. J. (2013) Iron-regulated surface determinant B (IsdB) promotes Staphylococcus aureusadherence to and internalization by non-phagocytic human cells. *Cell Microbiol* **15**, 1026-1041.

10. Andrews, S. C., Robinson, A. K., and Rodríguez Quiñones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215-237.

11. Weinberg, E. D. (1990) CELLULAR IRON-METABOLISM IN HEALTH AND DISEASE. *Drug Metab Rev* **22**, 531-579.

12. Wandersman, C., and Delepelaire, P. (2004) BACTERIAL IRON SOURCES: From Siderophores to Hemophores. *Annu Rev Microbiol* **58**, 611-647.

13. Pantopoulos, K., Porwal, S. K., Tartakoff, A., and Devireddy, L. (2012) Mechanisms of Mammalian Iron Homeostasis. *Biochem* **51**, 5705-5724.

14. Ammons, M. C., and Copié, V. (2013) Mini-review: Lactoferrin: a bioinspired, anti-biofilm therapeutic. *Biofouling* **29**, 443-455.

15. Tolosano, E., and Altruda, F. (2002) Hemopexin: structure, function, and regulation. *DNA Cell Biol* **21**, 297-306.

16. Wilks, A., and Burkhard, K. A. (2007) Heme and virulence: how bacterial pathogens regulate, transport and utilize heme. *Nat Prod Rep* **24**, 511-522.

17. Hantke, K. (1987) Ferrous Iron Transport Mutants in Escherichia-Coli-K12. *FEMS Microbiology Letters* **44**, 53-57.

18. Cartron, M. L., Maddocks, S., Gillingham, P., Craven, C. J., and Andrews, S. C. (2006) Feo – Transport of Ferrous Iron into Bacteria. *Biometals* **19**, 143-157.

19. Faraldo-Gómez, J. D., and Sansom, M. S. P. (2003) Acquisition of siderophores in Gram-negative bacteria. *Nat Rev Mol Cell Biol* **4**, 105-116.

20. Krewulak, K. D., and Vogel, H. J. (2008) Structural biology of bacterial iron uptake. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778**, 1781-1804.

21. Bracken, C. S., Baer, M. T., Abdur-Rashid, A., Helms, W., and Stojiljkovic, I. (1999) Use of heme-protein complexes by the Yersinia enterocolitica HemR receptor: histidine residues are essential for receptor function. *J Bacteriol* **181**, 6063-6072.

22. Burkhard, K. A., and Wilks, A. (2007) Characterization of the Outer Membrane Receptor ShuA from the Heme Uptake System of Shigella dysenteriae: SUBSTRATE SPECIFICITY AND IDENTIFICATION OF THE HEME PROTEIN LIGANDS. *J Biol Chem* **282**, 15126-15136.

23. Wolff, N., Izadi-Pruneyre, N., Couprie, J., Habeck, M., Linge, J., Rieping, W., Wandersman, C., Nilges, M., Delepierre, M., and Lecroisey, A. (2008) Comparative Analysis of Structural and Dynamic Properties of the Loaded and Unloaded Hemophore HasA: Functional Implications. *J Mol Biol* **376**, 517-525.

24. Braun, V., and Hantke, K. (2011) Recent insights into iron import by bacteria. *Curr Opin Chem Biol* **15**, 328-334.

25. Baelen, S., Dewitte, F., Clantin, B., and Villeret, V. (2013) Structure of the secretion domain of HxuA from Haemophilus influenzae. *Acta Cryst* **F69***,* 1322-1327.

26. Brown, J. S., and Holden, D. W. (2002) Iron acquisition by Gram-positive bacterial pathogens. *Microbes Infect* **4**, 1149-1156.

27. Hammer, N. D., and Skaar, E. P. (2011) Molecular Mechanisms of Staphylococcus aureusIron Acquisition. *Annu Rev Microbiol* **65**, 129-147.

28. Maresso, A. W., and Schneewind, O. (2006) Iron acquisition and transport in Staphylococcus aureus. *Biometals* **19**, 193-203.

29. Grigg, J. C., Cheung, J., Heinrichs, D. E., and Murphy, M. E. P. (2010) Specificity of Staphyloferrin B Recognition by the SirA Receptor from Staphylococcus aureus. *J Biol Chem* **285**, 34579-34588.

30. Maresso, A. W., Garufi, G., and Schneewind, O. (2008) Bacillus anthracis Secretes Proteins That Mediate Heme Acquisition from Hemoglobin. *PLoS Pathogens* **4**, e1000132.

31. Fabian, M., Solomaha, E., Olson, J. S., and Maresso, A. W. (2009) Heme transfer to the bacterial cell envelope occurs via a secreted hemophore in the Gram-positive pathogen Bacillus anthracis. *J Biol Chem* **284**, 32138-32146.

32. Lei, B., Smoot, L. M., Menning, H. M., Voyich, J. M., Kala, S. V., Deleo, F. R., Reid, S. D., and Musser, J. M. (2002) Identification and Characterization of a Novel Heme-Associated Cell Surface Protein Made by Streptococcus pyogenes. *Infect Immun* **70**, 4494-4500.

33. Ouattara, M., Bentley Cunha, E., Li, X., Huang, Y.-S., Dixon, D., and Eichenbaum, Z. (2010) Shr of group A streptococcus is a new type of composite NEAT protein involved in sequestering haem from methaemoglobin. *Mol Microbiol* **78**, 739-756.

34. Zhu, H., Liu, M., and Lei, B. (2008) The surface protein Shr of Streptococcus pyogenes binds heme and transfers it to the streptococcal heme-binding protein Shp. *BMC Microbiology* **8**, 15.

35. Torres, V. J., Pishchany, G., Humayun, M., Schneewind, O., and Skaar, E. P. (2006) Staphylococcus aureus IsdB is a hemoglobin receptor required for heme iron utilization. *J Bacteriol* **188**, 8421-8429.

36. Mazmanian, S. K., Ton That, H., Su, K., and Schneewind, O. (2002) An iron-regulated sortase anchors a class of surface protein during Staphylococcus aureus pathogenesis. *PNAS* **99**, 2293-2298.

37. Escolar, L., Pérez-Martín, J., and de Lorenzo, V. (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* **181**, 6223-6229.

38. Mazmanian, S. K., Skaar, E. P., Gaspar, A. H., Humayun, M., Gornicki, P., Jelenska, J., Joachmiak, A., Missiakas, D. M., and Schneewind, O. (2003) Passage of Heme-Iron Across the Envelope of Staphylococcus aureus. *Science* **299**, 906-909.

39. Skaar, E. P., and Schneewind, O. (2004) Iron-regulated surface determinants (Isd) of Staphylococcus aureus: stealing iron from heme. *Microbes Infect* **6**, 390-397.

40. Mazmanian, S. K., Ton That, H., and Schneewind, O. (2001) Sortase−catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus. *Mol Microbiol* **40**, 1049-1057.

41. Kim, H. K., DeDent, A., Cheng, A. G., McAdow, M., Bagnoli, F., Missiakas, D. M., and Schneewind, O. (2010) IsdA and IsdB antibodies protect mice against Staphylococcus aureus abscess formation and lethal challenge. *Vaccine* **28**, 6382-6392.

42. Pishchany, G., Dickey, S. E., and Skaar, E. P. (2009) Subcellular Localization of the Staphylococcus aureus Heme Iron Transport Components IsdA and IsdB. *Infect Immun* **77**, 2624-2634.

43. Dryla, A., Gelbmann, D., Von Gabain, A., and Nagy, E. (2003) Identification of a novel iron regulated staphylococcal surface protein with haptoglobin-haemoglobin binding activity. *Mol Microbiol* **49**, 37-53.

44. Clarke, S. R., Wiltshire, M. D., and Foster, S. J. (2004) IsdA of Staphylococcus aureus is a broad spectrum, iron-regulated adhesin. *Mol Microbiol* **51**, 1509-1519.

45. Heinrichs, D., and Taylor, J. (2002) Transferrin binding in Staphylococcus aureus: involvement of a cell wall-anchored protein: involvement of a cell wall-anchored protein. *Mol Microbiol* **43**, 1603-1614.

46. Miajlovic, H., Zapotoczna, M., Geoghegan, J. A., Kerrigan, S. W., Speziale, P., and Foster, T. J. (2010) Direct interaction of iron-regulated surface determinant IsdB of Staphylococcus aureus with the GPIIb/IIIa receptor on platelets. *Microbiology* **156**, 920-928.

47. Sharp, K. H., Schneider, S., Cockayne, A., and Paoli, M. (2007) Crystal structure of the heme-IsdC complex, the central conduit of the isd iron/heme uptake system in Staphylococcus aureus. *J Biol Chem* **282**, 10625-10631.

48. Liu, M., Tanaka, W. N., Zhu, H., Xie, G., Dooley, D. M., and Lei, B. (2008) Direct hemin transfer from IsdA to IsdC in the iron-regulated surface determinant (Isd) heme acquisition system of Staphylococcus aureus. *J Biol Chem* **283**, 6668-6676.

49. Zhu, H., Xie, G., Liu, M. Y., Olson, J. S., Fabian, M., Dooley, D. M., and Lei, B. F. (2008) Pathway for heme uptake from human methemoglobin by the iron-regulated surface determinants system of Staphylococcus aureus. *J Biol Chem* **283**, 18450-18460.

50. Maresso, A. W., Chapa, T. J., and Schneewind, O. (2006) Surface Protein IsdC and Sortase B Are Required for Heme-Iron Scavenging of Bacillus anthracis. *J Bacteriol* **188**, 8145-8152.

51. Bierne, H., Garandeau, C., Pucciarelli, M. G., Sabet, C., Newton, S., Garcia-del Portillo, F., Cossart, P., and Charbit, A. (2004) Sortase B, a New Class of Sortase in Listeria monocytogenes. *J Bacteriol* **186**, 1972-1982.

52. Grigg, J. C., Vermeiren, C. L., Heinrichs, D. E., and Murphy, M. E. (2007) Heme coordination by Staphylococcus aureus IsdE. *J Biol Chem* **282**, 28815-28822.

53. Lee, W. C., Reniere, M. L., Skaar, E. P., and Murphy, M. E. (2008) Ruffling of metalloporphyrins bound to IsdG and IsdI, two heme-degrading enzymes in Staphylococcus aureus. *J Biol Chem* **283**, 30957-30963.

54. Wu, R. (2004) Staphylococcus aureus IsdG and IsdI, Heme-degrading Enzymes with Structural Similarity to Monooxygenases. *J Biol Chem* **280**, 2840-2846.

55. Andrade, M. A., Ciccarelli, F. D., Perez-Iratxeta, C., and Bork, P. (2002) NEAT: a domain duplicated in genes near the components of a putative Fe3+ siderophore transporter from Gram-positive pathogenic bacteria. *Genome Biol* **3**, 0047-1.

56. Grigg, J. C., Ukpabi, G., Gaudin, C. F. M., and Murphy, M. E. P. (2010) Structural biology of heme binding in the Staphylococcus aureus Isd system. *J Inorg Biochem* **104**, 341-348.

57. Pilpa, R. M., Fadeev, E. A., Villareal, V. A., Wong, M. L., Phillips, M., and Clubb, R. T. (2006) Solution structure of the NEAT (NEAr Transporter) domain from IsdH/HarA: The human hemoglobin receptor in Staphylococcus aureus. *J Mol Biol* **360**, 435-447.

58. Grigg, J. C., Vermeiren, C. L., Heinrichs, D. E., and Murphy, M. E. P. (2007) Haem recognition by a Staphylococcus aureus NEAT domain. *Mol Microbiol* **63**, 139-149.

59. Gaudin, C. F. M., Grigg, J. C., Arrieta, A. L., and Murphy, M. E. P. (2011) Unique Herne-Iron Coordination by the Hemoglobin Receptor IsdB of Staphylococcus aureus. *Biochem* **50**, 5443-5452.

60. Villareal, V. A., Pilpa, R. M., Robson, S. A., Fadeev, E. A., and Clubb, R. T. (2008) The IsdC Protein from Staphylococcus aureus Uses a Flexible Binding Pocket to Capture Heme. *J Biol Chem* **283**, 31591-31600.

61. Watanabe, M., Tanaka, Y., Suenaga, A., Kuroda, M., Yao, M., Watanabe, N., Arisaka, F., Ohta, T., Tanaka, I., and Tsumoto, K. (2008) Structural basis for multimeric heme 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.

62. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* **89**, 496-514.

63. Wittekind, M., and Mueller, L. (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha-Carbon and Beta-Carbon Resonances in Proteins. *J Magn Reson* **101**, 201-205.

64. Grzesiek, S., and Bax, A. (1992) An Efficient Experiment for Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins. *J Magn Reson* **99**, 201-207.

65. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. *J Magn Reson, Series B* **101**, 114-119.

66. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* **114**, 6291-6293.

67. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry* **28**, 6150-6156.

68. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J Magn Reson* **88**, 425-431.

69. Vuister, G. W., and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson* **98**, 428-435.

70. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J Am Chem Soc* **111**, 1515-1517.

71. Fonner, B. A., Tripet, B. P., Lui, M., Zhu, H., Lei, B., and Copie, V. (2013) H, C, N backbone and side chain NMR resonance assignments of the N-terminal NEAr iron transporter domain 1 (NEAT 1) of the hemoglobin receptor IsdB of Staphylococcus aureus. *Biomol Nmr Assign* **8**, 201-205.

72. Hwang, T. L., and Shaka, A. J. (1998) Multiple-pulse mixing sequences that selectively enhance chemical exchange or cross-relaxation peaks in high-resolution NMR spectra. *J Magn Reson* **135**, 280-287.

73. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J Biomol Nmr* **6**, 277-293.

74. Goddard, T. D., and Kneller, D. G. (2008) SPARKY 3. University of California, San Francisco

75. Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J Biomol Nmr* **3**, 185-204.

76. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol Nmr* **24**, 171-189.

77. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319**, 209-227.

78. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS plus : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol Nmr* **44**, 213-223.

79. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **273**, 283-298.

80. Case, D. A., Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) AMBER 9. University of California, San Francisco.

81. Bertini, I., Case, D. A., Ferella, L., Giachetti, A., and Rosato, A. (2011) A Grid-enabled web portal for NMR structure refinement with AMBER. *Bioinformatics* **27**, 2384-2390.

82. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol Nmr* **8**, 477-486.

83. Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics* **50**, 437-450.

84. Luthy, R., Bowie, J. U., and Eisenberg, D. (1992) Assessment of Protein Models with 3-Dimensional Profiles. *Nature* **356**, 83-85

85. Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins: Structure, Function, and Bioinformatics* **66**, 778-795.

86. DeLano, W. L. (2002) The PyMol Molecular Graphics System. DeLano Scientific, San Carlos, CA

87. Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-W549.

88. Fonner, B. A., Tripet, B. P., Lui, M., Zhu, H., Lei, B., and Copie, V. . (2014) ^{1}H, ^{13}C, ^{15}N backbone and side chain NMR resonance assignments of the N-terminal NEAr iron transporter domain 1 (NEAT 1) of the hemoglobin receptor IsdB of Staphylococcus aureus. *Biomol. NMR Assign.* **8,** 201-205**.**

89. Ascoli, F., Fanelli, M. R., and Antonini, E. (1981) Preparation and Properties of Apohemoglobin and Reconstituted Hemoglobins. *Methods Enzymol* **76**, 72-87.

90. Wishart, D. S., and Sykes, B. D. (1994) The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure Using C-13 Chemical-Shift Data. *J Biomol Nmr* **4**, 171-180.

91. Eisenberg, D., Luthy, R., and Bowie, J. U. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **277**, 396-404.

92. Zhang, C., and Kim, S. H. (2000) A comprehensive analysis of the Greek key motifs in protein β−barrels and β−sandwiches. *Proteins: Structure, Function, and Bioinformatics* **40**, 409-419.

93. Dickson, C. F., Krishna Kumar, K., Jacques, D. A., Malmirchegini, G. R., Spirig, T., Mackay, J. P., Clubb, R. T., Guss, J. M., and Gell, D. A. (2014) Structure of the Hemoglobin-IsdH Complex Reveals the Molecular Basis of Iron Capture by Staphylococcus aureus. *J Biol Chem* **289,** 6728-6738.

94. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6**, 197-208.

95. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501.

96. Kumar, K. K., Jacques, D. A., Pishchany, G., Caradoc-Davies, T., Spirig, T., Malmirchegini, G. R., Langley, D. B., Dickson, C. F., Mackay, J. P., Clubb, R. T., Skaar, E. P., Guss, J. M., and Gell, D. A. (2011) Structural Basis for Hemoglobin Capture by Staphylococcus aureus Cell-surface Protein, IsdH. *J Biol Chem* **286**, 38439-38447.

97. Schlenker, C., Goel, A., Tripet, B. P., Menon, S., Willi, T., Dlakic, M., Young, M. J., Lawrence, C. M., and Copie, V. (2012) Structural Studies of E73 from a Hyperthermophilic Archaeal Virus Identify the "RH3" Domain, an Elaborated Ribbon-Helix-Helix Motif Involved in DNA Recognition. *Biochem* **51**, 2899-2910.

98. Spirig, T., Malmirchegini, G. R., Zhang, J., Robson, S. A., Sjodt, M., Liu, M. Y., Kumar, K. K., Dickson, C. F., Gell, D. A., Lei, B. F., Loo, J. A., and Clubb, R. T. (2013) Staphylococcus aureus Uses a Novel Multidomain Receptor to Break Apart Human Hemoglobin and Steal Its Heme. *J Biol Chem* **288**, 1065-1078.

99. Hargrove, M. S., Whitaker, T., Olson, J. S., Vali, R. J., and Mathews, A. J. (1997) Quaternary structure regulates hemin dissociation from human hemoglobin. *J Biol Chem* **272**, 17385-17389.

100. Pishchany, G., Sheldon, J. R., Dickson, C. F., Alam, M. T., Read, T. D., Gell, D. A., Heinrichs, D. E., and Skaar, E. P. (2013) IsdB-dependent Hemoglobin Binding Is Required for Acquisition of Heme by Staphylococcus aureus. *Journal of Infectious Diseases*, jit817.

101. Ekworomadu, M. T., Poor, C. B., Owens, C. P., Balderas, M. A., Fabian, M., Olson, J. S., Murphy, F., Balkabasi, E., Honsa, E. S., He, C., Goulding, C. W., and Maresso, A. W. (2012) Differential Function of Lip Residues in the Mechanism and Biology of an Anthrax Hemophore. *PLoS Pathogens* **8**, e1002559.

102. Lepre, C. A., and Moore, J. M. (1998) Microdrop screening: a rapid method to optimize solvent conditions for NMR spectroscopy of proteins. *J Biomol Nmr* **12**, 493-499.

103. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley.

104. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947-2948.

105. Honsa, E. S., Owens, C. P., Goulding, C. W., and Maresso, A. W. (2013) The Near-iron Transporter (NEAT) Domains of the Anthrax Hemophore IsdX2 Require a Critical Glutamine to Extract Heme from Methemoglobin. *J Biol Chem* **288**, 8479-8490.

106. Aranda, I., Roman, Worley, C. E., Liu, M., Bitto, E., Cates, M. S., Olson, J. S., Lei, B., and Phillips, J., George N. (2007) Bis-methionyl Coordination in the Crystal Structure of the Heme-binding Domain of the Streptococcal Cell Surface Protein Shp. *J Mol Biol* **374**, 374-383.

APPENDICES

APPENDIX A

NMR EXPERIMENTS AND PARAMETERS USED FOR DATA COLLECTION FOR $IsdB^{N1}$

Experiment	Sweep Width/Carrier (ppm)			Acquired Data (Complex Points)			Mixing Time	Number of
	F1	F2	F3	F1	F2	F3	(ms)	Scans
^{1}H-^{15}N HSQC	36.0/118.0	14.03/4.71		128	2048			16
HNCA	32.1/54.0	36.0/119.5	16.02/4.71	128	40	2048		16
HNCACB	75.0/39.0	36.0/119.5	14.03/4.71	128	40	2048		16
CBCA(CO)NH	75.0/39.0	36.0/119.5	14.03/4.71	128	40	2048		16
C(CO)NH	75.0/39.0	36.0/119.5	16.02/4.71	128	40	2048		16
HBHA(CO)NH	10.0/4.70	32.0/120.0	13.35/4.70	128	64	2048		16
HC(CO)NH	16.02/4.71	36.0/119.5	16.02/4.71	128	40	2048		16
^{15}N-NOESY-HSQC	14.03/4.71	36.0/118.0	14.03/4.71	128	64	2048	100	16
3D HCCH-TOCSY	8.33/4.70	75.0/39.0	13.98/4.70	128	64	1024	50	8
3D HCCH-NOESY	14.03/4.71	75.0/39.0	14.03/4.71	128	64	2048	100	16
^{1}H-^{13}C CT-HSQC	75.0/46.0	11.97/4.70		256	2048			8
HN-TOCSY	12.6/4.71	36.0/120.0	12.6/4.71	128	40	2048	60	16
CLEANEX	26.0/121.0	13.35/4.69		128	2048			16

APPENDIX B

EXAMPLE INPUT FILES FOR STRUCTURE CALCULATION OF $IsdB^{N1}$ IN CYANA

```
# cyana.seq
# Amino Acid sequence file

LEU   125
ASN   126
GLN   127
GLU   128
LEU   129
ARG   130
GLU   131
ALA   132
ILE   133
LYS   134
ASN   135
PRO   136
ALA   137
ILE   138
LYS   139
ASP   140
LYS   141
ASP   142
HIS   143
SER   144
ALA   145
PRO   146
ASN   147
SER   148
ARG   149
PRO   150
ILE   151
ASP   152
PHE   153
GLU   154
MET   155
LYS   156
LYS   157
LYS   158
ASP   159
GLY   160
THR   161
GLN   162
GLN   163
PHE   164
TYR   165
HIS   166
TYR   167
ALA   168
SER   169
SER   170
VAL   171
LYS   172
PRO   173
ALA   174
ARG   175
```

shift_q.prot
Chemical shift file from Sparky where stereospecific protons have been changed to Q.

```
# Chemical shift file from UNIO
# UNIO version 2.0.2
# File format   : cyana
# File generated: 2013 March  7th, 12:46:13

     1  56.175   0.189 CA    125
     2  42.568   0.066 CB    125
     3  25.800   0.001 CD1   125
     4  27.445   0.000 CG    125
     5   4.373   0.008 HA    125
     6   1.912   0.008 HB2   125
     7   1.387   0.008 HB3   125
     8   0.953   0.007 QD1   125
     9   0.918   0.000 QD2   125
    10   1.701   0.000 HG    125
    11   8.750   0.004 H     125
    12 123.733   0.018 N     125
    13  53.319   0.035 CA    126
    14  39.717   0.111 CB    126
    15   4.903   0.005 HA    126
    16   2.826   0.023 HB2   126
    17   2.768   0.003 HB3   126
    18   8.606   0.009 H     126
    19 124.214   0.073 N     126
    20  55.020   0.052 CA    127
    21  32.810   0.057 CB    127
    22  34.268   0.000 CG    127
    23   4.589   0.007 HA    127
    24   2.226   0.017 HB2   127
    25   2.002   0.013 HB3   127
    26   2.368   0.010 QG    127
    27   7.565   0.010 H     127
    28 117.458   0.055 N     127
    29  57.019   0.043 CA    128
    30  31.626   0.080 CB    128
    31  37.398   0.000 CG    128
    32   4.261   0.007 HA    128
    33   2.058   0.001 QB    128
    34   2.555   0.011 QG    128
```

hcch.peaks
XEASY cross peak list of NOE assignments from 3D HCCH-NOESY experiment

#Number of dimensions 3
#FORMAT xeasy3D
#INAME 1 H
#INAME 2 C
#INAME 3 HC
NMR peak file from UNIO
UNIO version 2.0.2
File format : xeasy
File generated: 2013 March 7th, 13:04:53

1	0.956	59.837	4.979	1 U	3.750E+06	0.000E+00	e 0	657	650	655	#VC	1.00000
2	0.963	41.229	1.902	1 U	3.030E+07	0.000E+00	e 0	657	651	656	#VC	1.00000
3	4.922	41.229	1.903	1 U	4.460E+06	0.000E+00	e 0	655	651	656	#VC	1.00000
4	0.628	14.843	0.965	1 U	2.820E+07	0.000E+00	e 0	782	652	657	#VC	1.00000
5	4.963	14.843	0.963	1 U	4.940E+06	0.000E+00	e 0	655	652	657	#VC	1.00000
6	0.511	34.224	1.570	1 U	9.500E+06	0.000E+00	e 0	412	387	391	#VC	1.00000
7	0.802	42.160	1.182	1 U	1.470E+07	0.000E+00	e 0	891	885	889	#VC	1.00000
8	1.946	42.160	1.175	1 U	4.500E+06	0.000E+00	e 0	1021	885	889	#VC	1.00000
9	3.405	42.160	1.175	1 U	4.420E+06	0.000E+00	e 0	1002	885	889	#VC	1.00000
10	1.176	71.807	3.410	1 U	7.830E+06	0.000E+00	e 0	889	999	1002	#VC	1.00000
11	0.074	71.807	3.408	1 U	4.020E+06	0.000E+00	e 0	890	999	1002	#VC	1.00000
12	4.699	51.860	5.407	1 U	1.230E+07	0.000E+00	e 0	1029	856	858	#VC	1.00000
13	5.568	50.752	5.082	1 U	3.840E+06	0.000E+00	e 0	704	759	761	#VC	1.00000
14	4.884	58.549	5.356	1 U	1.730E+07	0.000E+00	e 0	461	802	804	#VC	1.00000
15	4.768	60.765	4.471	1 U	3.460E+06	0.000E+00	e 0	471	395	399	#VC	1.00000
16	2.508	45.200	4.231	1 U	5.100E+06	0.000E+00	e 0	498	496	497	#VC	1.00000
17	5.042	45.200	4.235	1 U	1.790E+06	0.000E+00	e 0	767	496	497	#VC	1.00000
18	4.905	60.470	3.651	1 U	5.700E+06	0.000E+00	e 0	204	405	410	#VC	1.00000
19	5.172	53.670	5.532	1 U	8.800E+06	0.000E+00	e 0	1020	255	258	#VC	1.00000
20	6.165	54.480	5.397	1 U	1.740E+06	0.000E+00	e 0	789	478	481	#VC	1.00000
21	5.391	54.422	6.159	1 U	2.880E+06	0.000E+00	e 0	481	785	789	#VC	1.00000
22	4.925	41.360	3.534	1 U	1.730E+06	0.000E+00	e 0	1037	1036	1038	#VC	1.00000
23	0.669	60.765	4.479	1 U	4.990E+06	0.000E+00	e 0	401	395	399	#VC	1.00000
24	4.480	33.471	1.488	1 U	6.820E+06	0.000E+00	e 0	399	396	400	#VC	1.00000
25	0.888	43.014	1.713	1 U	8.220E+06	0.000E+00	e 0	229	222	227	#VC	1.00000
26	0.559	43.014	1.716	1 U	7.290E+06	0.000E+00	e 0	402	222	227	#VC	1.00000
27	1.673	19.020	0.898	1 U	3.140E+07	0.000E+00	e 0	227	225	229	#VC	1.00000
28	2.159	19.020	0.898	1 U	1.010E+07	0.000E+00	e 0	219	225	229	#VC	1.00000
29	5.206	19.020	0.897	1 U	6.980E+06	0.000E+00	e 0	218	225	229	#VC	1.00000
30	4.836	19.020	0.897	1 U	8.080E+06	0.000E+00	e 0	226	225	229	#VC	1.00000

Dihedral angle file as generated by TALOS+
cyana.aco

```
 126   ASN  PHI    -121.0   -73.0
 126   ASN  PSI     -55.0   -15.0
 127   GLN  PHI    -227.0   -87.0
 127   GLN  PSI     127.0   183.0
 128   GLU  PHI    -119.0    21.0
 128   GLU  PSI      78.0   180.0
 129   LEU  PHI     -82.0   -34.0
 129   LEU  PSI     -60.0   -20.0
 130   ARG  PHI     -85.0   -45.0
 130   ARG  PSI     -46.0     2.0
 131   GLU  PHI    -100.0   -60.0
 131   GLU  PSI     -63.0    -7.0
 132   ALA  PHI     -80.0   -40.0
 132   ALA  PSI     -55.0   -15.0
 133   ILE  PHI     -83.0   -43.0
 133   ILE  PSI     -49.0    -9.0
 134   LYS  PHI    -108.0   -68.0
 134   LYS  PSI     -32.0    28.0
 135   ASN  PHI    -119.0   -39.0
 135   ASN  PSI      71.0   175.0
 136   PRO  PSI     -50.0   -10.0
 137   ALA  PHI     -95.0   -51.0
 137   ALA  PSI     -66.0     6.0
 138   ILE  PHI    -133.0   -85.0
 138   ILE  PSI     -26.0    30.0
 139   LYS  PHI    -162.5    12.5
 139   LYS  PSI      62.5   180.0
 140   ASP  PHI      35.2    75.2
 140   ASP  PSI      23.2    63.2
 141   LYS  PHI    -204.0   -68.0
 141   LYS  PSI     118.0   158.0
 142   ASP  PHI    -148.0   -32.0
 142   ASP  PSI      81.0   193.0
 143   HIS  PHI    -155.0   -63.0
 143   HIS  PSI      70.0   196.0
 144   SER  PHI    -144.0   -34.0
 144   SER  PSI      92.0   185.0
 145   ALA  PHI    -155.3   -28.5
 145   ALA  PSI     116.4   180.4
```

distance restraint file generated from the XEASY peaks list by CYANA
#all.upl

```
126 ASN  H     127 GLN  H       4.74            #peak   860 #SUP  1.00
126 ASN  HA    127 GLN  QB      5.10            #peak   208
126 ASN  HA    127 GLN  QG      5.11            #peak   213 #SUP  1.00
126 ASN  HB2   127 GLN  HA      5.50            #peak   206 #SUP  1.00
126 ASN  HB3   127 GLN  HA      5.50            #peak   207 #SUP  1.00
126 ASN  QB    127 GLN  HA      4.72            #peak   207
127 GLN  QG    132 ALA  QB      5.20            #peak   212 #SUP  1.00
128 GLU  H     129 LEU  H       5.50            #peak  1047 #SUP  1.00
128 GLU  HA    130 ARG  H       5.33            #peak   869 #SUP  1.00
128 GLU  QG    129 LEU  H       5.50            #peak   864 #SUP  1.00
129 LEU  H     129 LEU  HD12    4.77            #peak   866 #SUP  1.00
129 LEU  H     130 ARG  H       5.25            #peak   861 #SUP  1.00
129 LEU  H     131 GLU  H       5.50            #peak   877 #SUP  1.00
129 LEU  HA    129 LEU  HD11    5.07            #peak    77 #SUP  1.00
129 LEU  HA    132 ALA  QB      4.50            #peak   216 #SUP  1.00
129 LEU  HA    133 ILE  QG2     4.45            #peak   217 #SUP  1.00
129 LEU  HD11  195 TRP  HE3     5.50            #peak    76 #SUP  1.00
129 LEU  HD11  221 ALA  QB      4.63            #peak    73 #SUP  1.00
129 LEU  HD12  130 ARG  HA      4.78            #peak   221 #SUP  1.00
129 LEU  QB    130 ARG  H       3.97            #peak   873
129 LEU  QB    132 ALA  QB      4.46            #peak   233
129 LEU  QQD   130 ARG  H       5.44            #peak   874
129 LEU  QQD   132 ALA  H       5.44            #peak   890
129 LEU  QQD   132 ALA  QB      5.44            #peak   235
129 LEU  QQD   221 ALA  H       6.00            #peak   415
129 LEU  QQD   221 ALA  QB      4.42            #peak   123
130 ARG  H     130 ARG  QD      5.30            #peak   871 #SUP  1.00
130 ARG  H     131 GLU  H       5.03            #peak   867 #SUP  1.00
130 ARG  H     196 ARG  H       5.50            #peak   652 #SUP  1.00
130 ARG  H     196 ARG  HA      5.10            #peak   868 #SUP  1.00
130 ARG  QG    194 PHE  H       5.50            #peak   644 #SUP  1.00
130 ARG  QG    196 ARG  H       5.47            #peak   655 #SUP  1.00
131 GLU  H     131 GLU  QB      3.88            #peak   881 #SUP  1.00
131 GLU  H     132 ALA  H       4.07            #peak   884 #SUP  1.00
131 GLU  H     132 ALA  HA      4.81            #peak   879 #SUP  1.00
131 GLU  H     132 ALA  QB      4.35            #peak   882 #SUP  1.00
131 GLU  H     133 ILE  H       5.50            #peak   876 #SUP  1.00
131 GLU  H     196 ARG  HA      5.50            #peak   880 #SUP  1.00
131 GLU  HA    133 ILE  H       5.50            #peak   895 #SUP  1.00
```

APPENDIX C

SCRIPTS NEEDED FOR STRUCTURE CALCULATIONS IN CYANA

#Script to initialize CYANA
#init.cya

```
name:=cyana
nproc=2
rmsdrange:=129..133,151..156,175..178,184..189,198..202,209..215,221..2
27,233..242,251..256
cyanalib
read seq cyana.seq
```

Script to perform a structure calculation with CYANA by generating distance restraint files from NOESY peak lists and automatic calibration.
#CALC.cya

```
peaks      := nh,hcch,hcch_2,hcch_loop  # names of peak lists
prot       := shift_q                   # names of proton lists
tolerance  := 0.040,0.030,0.45   # chemical shift tolerances
                      # order: 1H(a), 1H(b), 13C/15N(b), 13C/15N(a)
calibration:=         # calibration constants (will be determined
                             #   automatically, if commented out)
dref       := 5.0           # average upper distance limit for
                           #   automatic calibration
#set upl_values:=2.4,6.0   # changes the upper bound from 5.5 to 6.0A

if (master) then

  # ---- check consistency of peak and chemical shift lists----

  peakcheck peaks=$peaks prot=$prot

  # ---- calibration ----

  calibration prot=$prot peaks=$peaks constant=$calibration dref=$dref
  peaks calibrate "**" simple
  write upl $name-in.upl
  distance modify info=full
  write upl $name.upl

end if
synchronize

# ---- structure calculation ----

read seq $name.seq                   # re-read sequence to initialize
read upl $name.upl                   # read upper distance limits
read aco $name.aco                   # read angle constraints
read upl hbond.upl append
seed=5671                             # random number generator seed
calc_all structures=200 command=anneal steps=1000
 # calculate 100 conformers
overview $name.ovw structures=40 pdb
 # write overview file and coordinates
```

Script to perform a structure calculation with CYANA from the upper distance restraint file
#CALC2.cya

```
# ------ read input files ------

read upl all.upl              # read upper distance bounds
read aco cyana.aco            # read dihedral angle restraints


# ------ structure calculation ------

seed=5671                       # random number generator seed
calc_all 200 command=anneal steps=10000               # calculate
conformers


# ------ structure calculation ------

overview cyana.ovw structures=40 pdb      # write overview file and 40
best conformers
```