Structural and Biochemical Investigations of the Heme Loading Mechanisms in Hemophores HasAp from *Pseudomonas aeruginosa* and $HasA_{yp}$ from *Yersinia pestis*

By

Ritesh Kumar

Submitted to the graduate degree program in Bioinformatics
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chairperson: Mario Rivera, Ph.D.

Ilya A. Vakser, Ph.D.

Eric J. Deeds, Ph.D.

J Christian J Ray, Ph.D.

David R. Benson, Ph.D.

Scott Lovell, Ph.D.

Date Defended: April 16, 2014

The Dissertation Committee for **Ritesh Kumar** certifies that this is

the approved version of the following dissertation:

Structural and Biochemical Investigations of the Heme Loading Mechanisms in Hemophores HasAp from *Pseudomonas aeruginosa* and $HasA_{yp}$ from *Yersinia pestis*

_____

Chairperson: Mario Rivera, Ph.D.

Date approved: April 16, 2014

ABSTRACT

Under iron-limiting conditions, several Gram-negative pathogens secrete hemophore (HasA) protein that binds heme and deliver it to the outer membrane receptors in bacteria. Hemophore proteins have been structurally and biochemically characterized in their heme-free (apo) and heme-bound (holo) forms in *Serratia marcescens* (HasAs) and *Pseudomonas aeruginosa* (HasAp). Hemophores have unique structural features which are distinguished from other known proteins. The structure of HasAp (and HasAs), are composed of an "α-helical wall" on one side and a "β-sheet wall" on another. Heme is bound between the two extended loops, the Y75 loop harboring the conserved Tyr75 ligand and the H32 loop harboring the His32 ligand. The major difference between the apo- and holo-HasAp is the large conformational rearrangements of the H32 loop upon heme binding which relocates His32 side chain ~ 30 Å from open to closed form. It has been established through kinetics experiments that heme first loads onto the Y75 loop where it is coordinated by Tyr75, followed by the closure of the H32 loop and coordination of heme-iron ion by His32.

Multiple sequence alignment of known hemophore sequences from Gram-negative bacteria reveals the conservation of Tyr75 but not the His32; instead in *Yersinia* species Gln32 is present at a similar position. To investigate the role played by the Gln32 in heme coordination, we structurally characterized hemophore from *Yersinia pestis* ($HasA_{yp}$) in apo- and holo-form. Surprisingly, the Q32 loop in apo-$HasA_{yp}$ is already in closed conformation and the heme is only coordinated by one protein provided ligand, Tyr75 from the Y75 loop. Further, heme loading is fast compared to the HasAp and occurs with minimal structural changes.

Analyzing the available structural information from hemophores (HasAp, $HasA_{yp}$ and HasAs), we looked for the interactions that stabilized the open H32 loop conformation in apo-HasAp. Our analysis revealed the presence of Arginine at position 33 (Arg33) anchoring the H32 loop to the body of the protein. In order to investigate the role played by the Arg33, we replaced it with Alanine (Ala) and carried out mutagenesis investigations. Results from X-ray crystallography, solution NMR and molecular dynamics simulations clearly indicate that single replacement of Arg33 to Ala shifts the open H32 loop conformation to a closed conformation which is very similar to the H32 loop in the wt holo-HasAp.

The structure of Y75 loop is conserved with Tyr75 and its hydrogen bond partner His83 (His81 in $HasA_{yp}$). To investigate the proposed requirement for the Tyr75 – His83 unit for heme coordination by Tyr75, we conducted structural and biochemical study of Y75A and His83A HasAp mutants. The high resolution crystal structure of Y75A and H83A mutants show the original α+β protein fold similar to the wt protein. In the absence of Tyr75 in Y75A mutant, His83 do not coordinate the heme-iron ion instead, a formate ion occupies the same position as that of the side chain of Tyr75, coordinating heme iron ion by one of its oxygen atoms. In H83A mutants, Tyr75 and His32 coordinate the heme-iron ion even at the pH 5.4. Despite the severe changes made in the Y75 loop residues, kinetic analysis suggests that the heme loads with biphasic kinetics similar to the wt HasAp. Taken together, heme loading onto to the Y75 loop is mainly driven by hydrophobic, π–π interactions onto the Y75 loop.

ACKNOWLEDGEMENTS

There are not enough words in English dictionary to thank my advisor Prof. Mario Rivera for his confidence and faith in me. I would like to thank him for giving me an opportunity to learn and work in his laboratory, support and encouragement throughout my graduate studies. I am also indebted for his wisdom and insightful discussions during the times when I was really down. I would simply put, Prof. Rivera, you are an amazing teacher, a great mentor and a good friend to me.

I would like to extend my sincere gratitude to Dr. Scott Lovell, with whom I worked closely and learned X-ray crystallography. He made learning fun and I was really inspired by his passion and immense knowledge of crystallography. Sometimes I would think of recording our discussions while solving the structures. I would also thank him for his valuable time, suggestions and sharing his experiences during my job search. I am also very thankful to Dr. David Benson for his insightful suggestions, opinions and ideas during my graduate work. I would also like to thank my committee members Prof. Ilya Vakser, Dr. Eric Deeds and Dr. J Christian J Ray for their valuable time reading my dissertation and providing me with suggestions. I would like to express my sincere regards to our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University, Portland, Oregon for performing all the kinetics experiments for my research work. I also thank Dr. Kevin P. Battaile for collecting diffraction data at the synchrotron facility at Argonne National Laboratory, Illinois. I would like to thank our collaborators Dr. Wonpil Im and Dr. Yifei Qi, Center for Bioinformatics and Department of Molecular Biosciences, University of Kansas, Lawrence, KS for the molecular dynamics simulation results which added additional insights into my research work.

I would like to thank my wonderful colleagues, and was fortunate to work with them in the lab. I thank past and present group members Dr. Grace Jepkorir for teaching me purification of HasA proteins, Dr. Saroja Kumari Weeratunga, Dr. Juan C. Rodríguez, Pavithra Vani Nama, Dr. Huili Yao for her help in performing solution NMR experiments for my research work, Yan Wang, Kate Eshelman, Mary Erickson, Kena Senegal, Pearl Ashitey, Rachel Lietz and Nurjahan Mehzabeen from Dr. Scott's Lab.

I also wish to thank my friends Dr. Ambrish Roy, Dr. Vishwendra Patel, Dr. Pallabi Sarkar, Dr. Veer Sandeep Bhatt and Dr. Nikunj Shukla for their continuous support and help.

Finally, I would like to thank my family and my in-laws for their never ending love, support and courage to pursue my dreams. I thank my father Mr. Arvind Kumar for teaching me the importance of hard work, dedication and responsibilities. My mother Mrs. Madhuri Devi for her unconditional love, affection and support and her dream to see me getting a Ph.D. You were a source of strength for me wherever I found myself in dilemma or confusion. My sweet little sister Surbhi Sonal and my brother Saurabh Kumar for their love and support. My wife Sweta Kumari Gupta for her unwavering love, support, and suggestions during my graduate studies. My in-laws (Mr. Arvind Ram and Mrs. Rekha Devi) for their love and support.

Finally, I would like to thank everyone who directly or indirectly helped me make my graduate journey successful.

DEDICATION

I dedicate this dissertation to my parents, **Mr. Arvind Kumar** and **Mrs. Madhuri Devi** and my wife **Sweta Kumari Gupta**, for their unwavering love, strong support and continuous encouragement.

Table of Contents

List of Figures

Chapter 1

Chapter 2

List of Tables

List of Schemes

Chapter 1

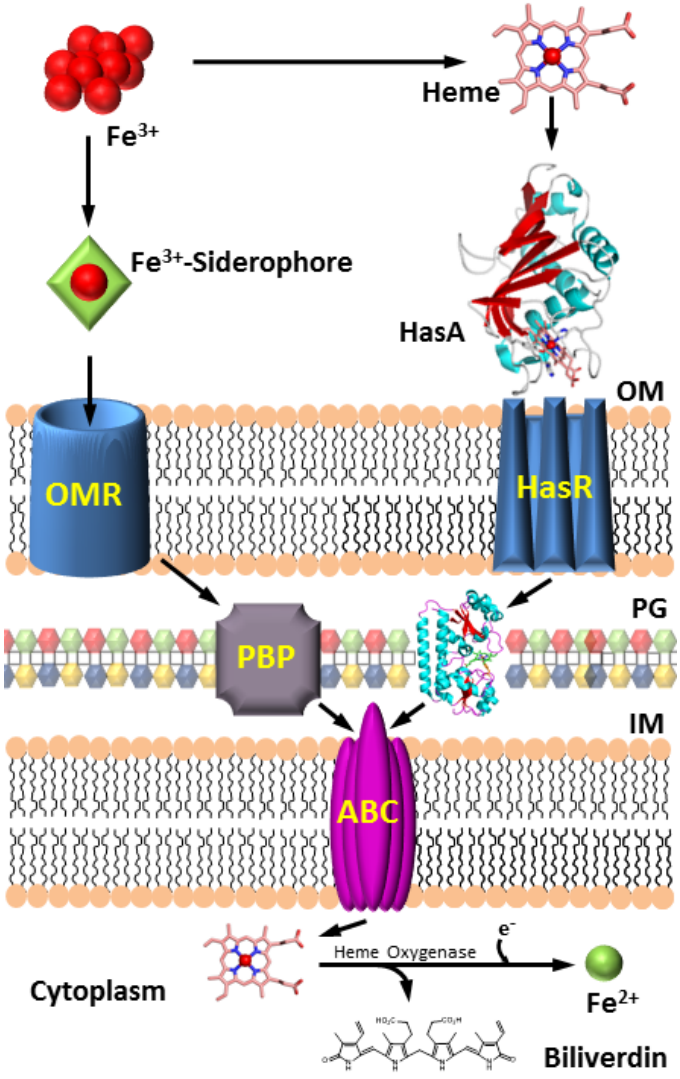
List of Appendices

Abbreviations

ABC	ATP-binding cassette
DMSO	Dimethyl sulfoxide
Has	Heme Acquisition System
HasR	Heme Acquisition System Receptor
Hb	Hemoglobin
HDL	High density lipoprotein
Heme	Fe^{2+} complex of protoporphyrin IX
Hemin	Fe^{3+} complex of protoporphyrin IX
Hp	Haptoglobin
HPX	Hemopexin
HSA	Human serum albumin
IPTG	Isopropyl-β-D-thiogalactopyranoside
LDL	Low density lipoprotein
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PCR	Polymerase chain reaction
PDB	Protein Data Bank
RMSD	Root mean square deviation
RR	Resonance Raman
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
UV-vis	Ultraviolet-visible
WT	Wild type

Chapter 1

Introduction



Preface

With the growing resistance of the opportunistic pathogens to available antibacterial medicines, there has been an accelerated effort in the scientific community to re-design strategies to develop novel therapeutics, which could be efficacious and at the same time provide long-term resistance from bacterial infections. Some opportunistic pathogens, such as *Pseudomonas aeruginosa, Yersinia pestis, Escherichia coli, Mycobacterium tuberculosis, Haemophilus influenzae, Salmonella typhimurium, Staphylococcus aureus, Shigella dysenteriae* etc., are commensal with the vertebrate host until the host immune system is compromised.

A very common example of Gram-negative opportunistic pathogen is *Pseudomonas aeruginosa*, which is ubiquitously present. In immunocompromised patients, *P. aeruginosa* infection could be potentially fatal [1]. In the United States alone, on an average 4 cases per 1000 patients (0.4 %) released from hospitals are infected by *P. aeruginosa* [2]. Moreover, it's been estimated that almost 50 % of the patients suffering from burns, cystic fibrosis and cancer do not survive *P. aeruginosa* infections [2].

Irrespective of the nature of the pathogen that invades the vertebrate host, one thing common between all of them is the requirement for iron which is a nutrient need to multiply and cause infection in the host. Research finds that iron is necessary to successfully colonize in the vertebrate host during the initial phase of infections [3].

This strongly supports the idea that pathogenicity of the bacteria could be targeted if they are deprived of the essential nutrient iron. Genes involved in the iron uptake and regulation have been characterized but there is a gap in the knowledge about the proteins they code, their structure, functions and various dynamic interactions during infection. This brings the attention

for a better understanding of the mechanism by which bacteria scavenge iron from the vertebrate host. Structural characterization of these proteins at atomic detail would help us to design inhibitors which would target specific protein-protein interactions. This in turn would help to design better therapeutics which would be different from the conventional therapies and would provide long term resistance from the bacterial infections.

1.1 Iron: An Important Metal Ion in Biological Systems

Iron is the fourth most abundant element in the earth's crust [4], and is an essential nutrient for nearly all organisms. It is a key element in oxidation/reduction reactions, due to its Fe^{2+}/ Fe^{3+} redox potentials which depend on the ligands that are coordinated with the iron ion. It acts as a catalyst for many oxidative enzymes and is important for many physiologically important reactions like oxygen transport (hemoglobin and myoglobin), cytochromes in respiration, biosynthesis of genetic material DNA, cell division and gene regulation [5, 6]. Despite its abundance and significance, iron acquisition and incorporation in biological systems is difficult, leading to cellular toxicity from the soluble ferrous (Fe^{2+}) form (0.1 M at pH 7) which generates highly reactive free hydroxyl radicals via Fenton-type reaction causing damage to DNA, proteins, lipid membranes and other cellular components, and the ferric (Fe^{3+}) form (10^{-18} M at pH 7) which is insoluble at the physiological pH [7].

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + OH^{\cdot} \qquad \text{(Reaction 1; Fenton reaction)}$$

In addition, one-electron reduction of O_2 by ferrous ions produces superoxide radicals which in turn reduce the ferric ions to ferrous ions:

$$O_2^{-} + Fe^{3+} \rightarrow O_2 + Fe^{2+} \qquad \text{(Reaction 2)}$$

Summation of the above two reactions (Reaction 1 and Reaction 2) is called the Haber-Weiss cycle which generates the hydroxyl anions, oxygen and reactive hydroxyl radicals as the reaction products.

$$H_2O_2 + O_2^{\cdot-} \rightarrow OH^- + OH^{\cdot} + O_2 \qquad \text{(Reaction 3; Haber-Weiss reaction)}$$

As a consequence of the potential toxicity of the ferrous form and the insolubility of the ferric form, all living organisms have developed efficient ways of acquiring, storing and utilizing iron.

1.2 Iron Uptake Mechanisms in Bacteria

Similar to vertebrates, bacterial pathogens require iron to meet their metabolic needs and function normally. Gram-positive bacteria have thick peptidoglycan lipid layer around the plasma membrane whereas in Gram-negative bacteria, the cell membrane is compartmentalized into outer membrane, periplasmic space and inner membrane. Gram-positive and Gram-negative bacteria have different mechanism to uptake iron/ heme from the vertebrate host.

1.2.1 Siderophores

Under iron limiting conditions, bacteria produces and secretes small molecular weight compounds called siderophores that strongly and specifically bind ferric (Fe^{3+}) iron ion [8, 9]. Siderophores solubilize the ferric ion from organic complexes, and actively transport against the concentration gradient via specific cell surface receptors through the inner membrane by the periplasmic protein-dependent TonB transporters [10-12]. Along with siderophores secretion, some bacteria secrete extracellular proteases that lyse host iron-binding proteins [13, 14], or produce exotoxins that target iron containing proteins [15, 16] or directly bind host iron-binding

proteins and reduce the insoluble Fe^{3+} to the usable soluble Fe^{2+} form [17]. Two different mechanisms operate in Gram-negative and Gram-positive bacteria to internalize iron bound siderophores.

1.2.1.1 Siderophore Uptake in Gram-negative Bacteria

In Gram-negative bacteria, the iron-siderophore complex is first bound by the specific outer membrane receptor proteins such as FepA, FecA and FhuA and are actively transported across the outer membrane into the periplasm, using the energy dependent TonB-ExbB-ExbD system [18]. Once in the periplasm, they are bound by the ATP-binding cassette (ABC) transporters in the inner membrane and are delivered to the cytoplasm.

In *E. coli*, low levels of intracellular Fe^{2+} induce the production of siderophore called enterobactin [19, 20]. The *fur* gene encodes a DNA binding protein called Ferric Uptake Regulator (Fur) which binds to the upstream of the specific DNA sequence and with Fe^{2+} co-represses siderophore synthesis (Figure 1.1). Fe^{2+} acts as a co-repressor of the transcription machinery. Since concentrations of Fe^{2+} transcriptionally control the expression of genes that promote iron acquisition, a mutation identified in Fur leads to constitutive expression of the iron uptake genes in *E. coli* [21].

In *Pseudomonas aeruginosa*, two siderophores, pyoverdine and pyochelin [22, 23] are produced under the regulation of Fur [24]. Under iron limiting conditions, *P. aeruginosa* first synthesizes pyochelin and then switches to pyoverdine production under extremely low iron concentrations [12, 25]. Pyochelin biosynthesis employs fewer genes and has lower iron binding affinity ($K_f = 10^5\ M^{-1}$) compared to pyoverdine ($K_f = 10^{24}\ M^{-1}$) [25]. The high affinity of pyoverdine can displace iron ions from storage proteins ferritins. It also acts as a regulator for

siderophore receptors on the cell membrane as well as signal molecule by triggering the production of virulence factors, PrpL and exotoxin A [12].

Yersinia pestis secretes a highly potent siderophore called yersiniabactin under iron limiting conditions [26]. Yersiniabactin has a formation constant of 4×10^{-36} for ferric iron ions and no affinity to bind ferrous ions [26].

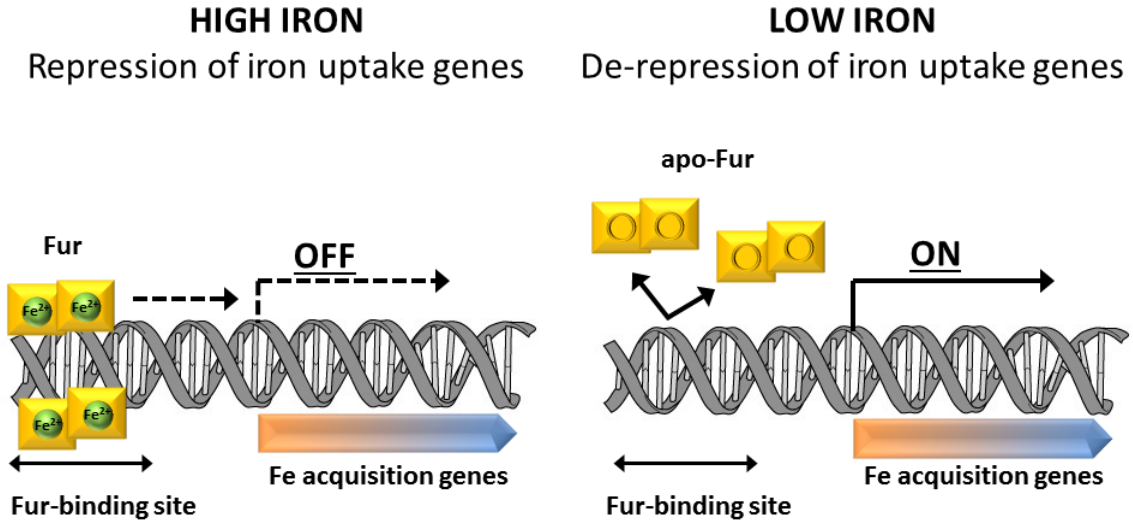


Figure 1.1. Schematic representation of Fur-mediated gene expression. Figure adapted and re-drawn from reference [4].

1.2.1.2 Siderophore Uptake in Gram-positive Bacteria

In Gram-positive bacteria, membrane anchored binding proteins (MABP) bind the ferri-siderophore complexes which are then transported across the inner membrane. Once these iron-siderophore complexes are inside the cytosol, they are either reduced to ferrous ions or in some organisms siderophores are broken down to release iron ions (Figure 1.2).

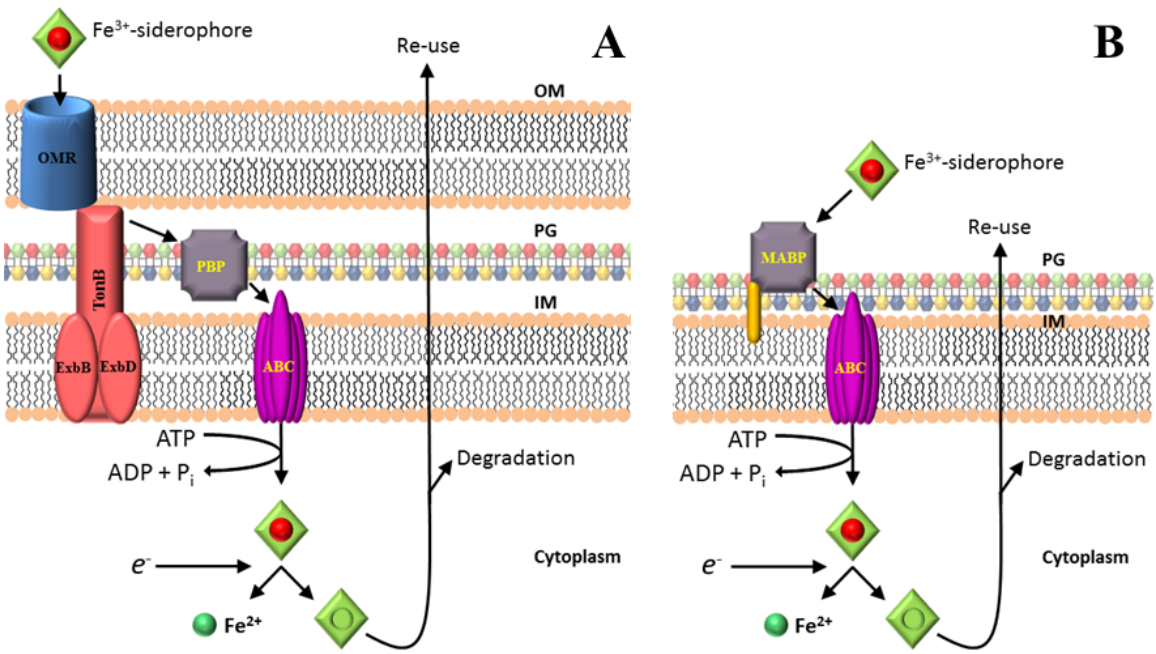


Figure 1.2. Schematic representation of siderophore-mediated iron uptake by (A) Gram-negative and (B) Gram-positive bacteria. OMR: Outer membrane receptor, PBP: Periplasmic binding protein, ABC: ATP-Binding Cassette proteins, MAPB: Membrane-anchored binding protein, OM: Outer membrane, PG: Peptidoglycan, IM: Inner membrane. Figure adapted and re-drawn from reference [27].

1.3 Heme: Most Abundant Iron-containing Prosthetic Group

In the vertebrates, the blood plasma concentration of free iron is kept below 10^{-18} M due to sequestration into high-affinity iron-binding proteins like heme in hemoglobin, iron-sulfur clusters, transferrin, lactoferrin and storage proteins like ferritins. The regulation of heme biosynthesis and its degradation is tightly controlled as heme is tightly bound, as a prosthetic group, to heme proteins or to the transport proteins like hemopexin (HPX) or haptaglobin (HPT) [28]. However, like iron, free heme is toxic and causes oxidative damage to the cells by catalyzing formation of the reactive oxygen intermediates. To avoid this deleterious effect, free

heme released into the blood plasma as a result of the breakdown of the erythrocytes are quickly scavenged for recycling by human serum albumin (HSA), HPX and hemoglobin-haptaglobin (Hb-HPT) complex [29, 30]. Under pathological conditions, free heme, HPX, HSA and Hb-Hp complex all serve as a source of heme for the invading pathogens. The low availability of free heme in the vertebrate host has led to the evolution of dedicated and sophisticated mechanisms in bacteria for the acquisition of heme for their nutrient requirements [31].

Majority of iron (~ 95 %) is sequestered in the form of heme, in hemoglobin (Hb) within the vertebrate cells [31, 32]. Hemoglobin (Hb) is the most abundant heme protein in the erythrocytes that transports oxygen, carbon dioxide and nitric oxide [30]. However, as the life span of erythrocytes end, they undergo lysis resulting in the release of their contents into the plasma. Free heme released from the lysis of Hb is sequestered rapidly by the high and low density lipoproteins (HDL and LDL), human serum albumin (HSA) and hemopexin (HPX) [30].

During the initial release of heme in the plasma, around 80% is quickly sequestered in LDL and HDL and the remaining 20% binds to HSA and HPX, which is slowly removed from the LDL and HDL [30]. Final fate of the bound heme is the delivery to specific cell receptors in liver, spleen and macrophages [33-35]. Pathogenic bacteria can scavenge free heme, Hb, Hb-Hp complex, and the heme-bound HSA and HPX for their metabolic needs.

1.4 Heme as an Iron Source for Bacteria

1.4.1 Bacterial Heme Uptake Systems

Heme uptake in Gram-negative bacteria can be classified into two groups: in one, there is direct uptake of the heme or heme associated with the host hemoproteins like hemoglobin, heme-serum albumin, heme-hemopexin, hemoglobin-haptoglobin and myoglobin. And in another,

heme uptake involves the secretion of small heme-binding proteins called hemophores that either bind/ extract heme from the lysed or degraded iron containing proteins and deliver it to the specific outer membrane receptors on the cell surface of the bacteria. Once the heme or is delivered to the receptor, the cargo (heme) is shuttled to the cytosol via a series of membrane bound proteins like TonB-dependent outer membrane receptor, a periplasmic binding protein (PBP), and an ABC transporter. Inside the cytosol, cargo is acted upon and degraded by the heme-degrading enzymes (heme oxygenases) to release iron ion (Figure 1.3-A) [31]. In Gram-positive bacteria, the cargo is bound by the lipoproteins anchored in the peptidoglycan and are then transported to the ABC transport proteins [27, 31, 32, 36-38] (Figure 1.3-B). Details of the heme uptake systems in Gram-negative bacteria will be discussed in the following sections.

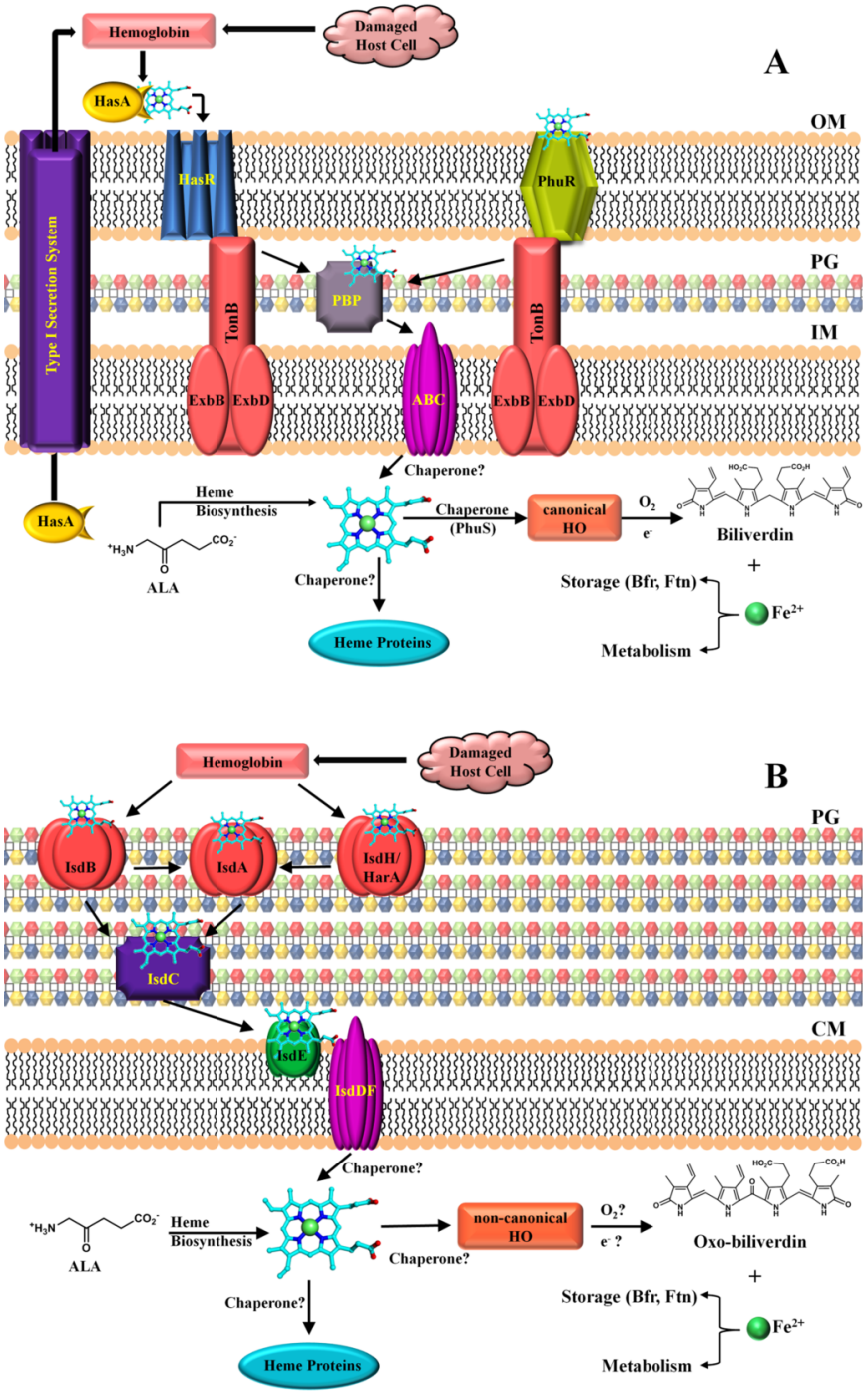


Figure 1.3. Schematic representation of heme uptake and degradation in (A) Gram-negative bacteria and (B) Gram-positive bacteria (*Staphylococcus aureus*). Figure adapted from reference [31].

1.4.2 Mechanisms of Heme Uptake in Gram-negative Bacteria

Extensive studies have been done to understand the heme uptake systems in Gram-negative bacteria. Numerous proteins and various mechanisms been suggested which are broadly classified by Wandersman and co-workers [39] in two groups: in the first group, heme uptake involves direct binding of the heme and/or heme-proteins (hemoglobin, human serum albumin, hemopexin and hemoglobin-haptoglobin complex) to the specific cell surface receptors on the outer membrane. This class includes *shuASTUV* from *Shigella dysenteriae* [40-42], *ChuA* from *Escherichia coli* O157:H7 [43], *phuRSRUVW* from *Pseudomonas aeruginosa* [44-46], *hmuRSTUV* from *Yersinia pestis* [47] and the *hemR-hemSTUV* system in *Yersinia enterocolitica* [48]. Genetic organization of *phu* (**P**seudomonas **h**eme **u**ptake) operon in *Pseudomonas aeruginosa* involves a single outer membrane receptor (phuR) for uptake of heme, a periplasmic transport protein (phuT), inner membrane proteins (phuUVW) and a cytoplasmic protein (phuS) [49] (Figure 1.4).

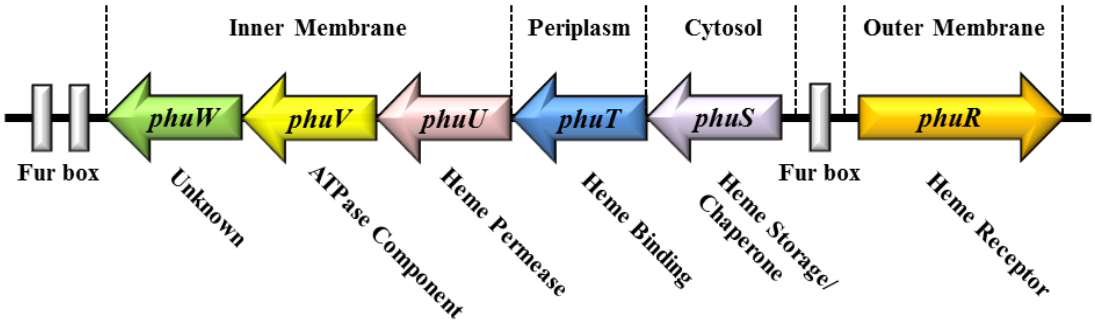


Figure 1.4. Genetic organization of the heme uptake locus in *Pseudomonas aeruginosa* containing the *phuR* gene and the *phuSTUVW* operon. The three Fur binding domains are shown as white rectangles. Figure adapted and re-drawn from reference [49].

The second group has the small secreted protein called hemophore which delivers the heme to its outer membrane receptor. This hemophore/receptor pair is found in *Serratia*

marcescens (*hasRADEBF*) [50, 51], *Pseudomonas aeruginosa* (*hasRADEF*) [52], *Pseudomonas fluorescens* [53], *Yersinia pestis* (*hasRADEBF*) [54]. A different type of hemophore have been identified in *Haemophilus influenzae* (HuxA) (*hxuCBA*) [55, 56] and *Porphyromonas gingivalis* (HusA) [57] [31]. The *has* operon in *Serratia marcescens*, which is similar to that of *Pseudomonas aeruginosa*, is regulated by intracellular iron concentrations via Fur repression. Under iron-replete conditions, Fur protein binds Fe^{2+} at the upstream in the promoter region of DNA and represses the transcription of iron uptake genes (Figure 1.5-A, yellow square). The two genes, homologous to *E. coli* extracytoplasmic function (ECF) sigma factors and sigma-modulators, are located upstream of HasR namely, *hasI* (an ECF sigma factor) (Figure 1.5-A, green oval) and *hasS* (HasI antisigma factor) which codes for HasI and HasS proteins. In *Serratia marcescens*, it has been shown by Rossi and coworkers [58] that positive and negative regulation of *hasI* and *hasS* genes work in cooperation with Fur-dependent mechanism to operate the *has* operon under iron-limiting conditions. When the HasR receptor is bound with the holo-hemophore (HasA), it activates the HasI sigma factor and inactivates the HasS, antisigma modulator, enabling the signal cascade. In contrast, in iron-replete conditions, the anti-sigma activity of HasS is blocked and HasI induces the transcription of HasS leading to the accumulation of inactive HasS molecules. As the external concentration of external heme source decreases or apo-hemophore is bound to the HasR, accumulated HasS blocks the expression of *hasR* [58, 59] (Figure 1.5).

1.4.3 Hemophore-dependent Heme Acquisition Systems

Hemophores HasA (**h**eme **a**cquisition **s**ystem) are small proteins which are secreted to the extracellular milieu to capture (free or protein bound) heme and deliver it to the specific outer membrane receptor, HasR. Hemophores belong to a family of highly conserved proteins with no

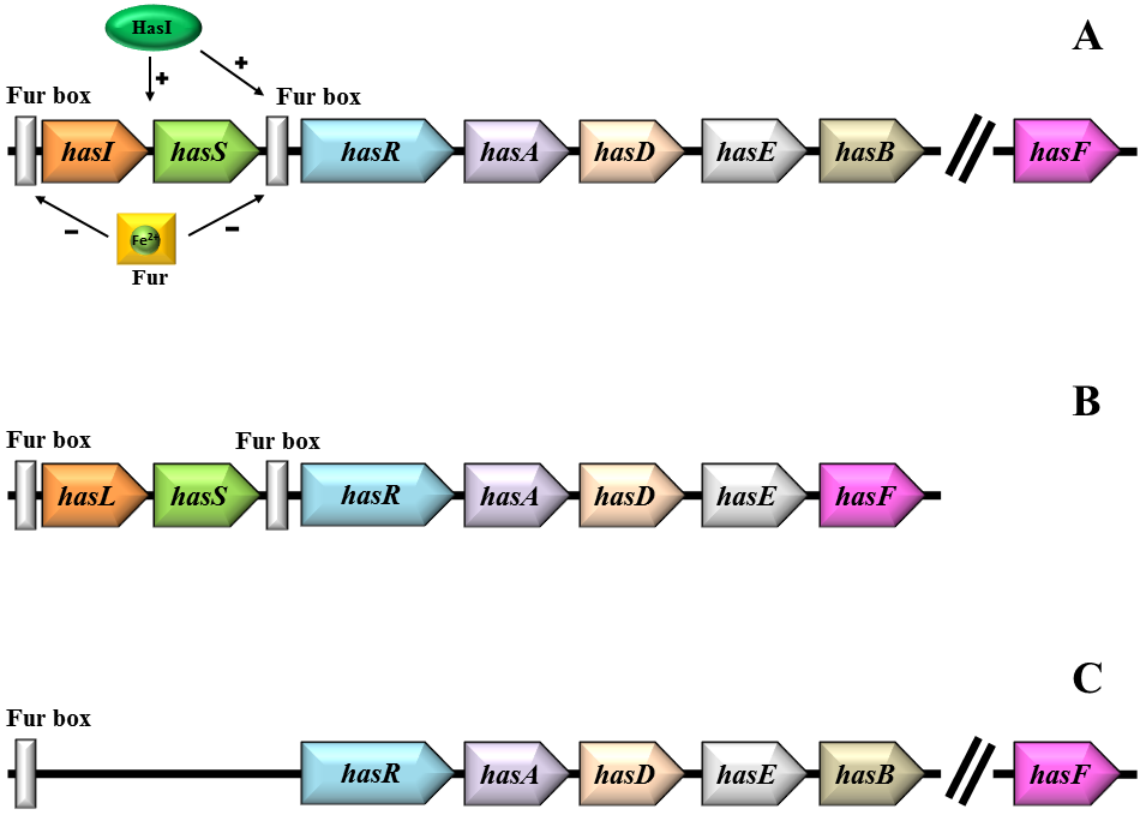


Figure 1.5. Genetic organization of *has* operon in (A) *Serratia marcescens*, (B) *Pseudomonas aeruginosa*, and (C) *Yersinia pestis*. Adapted and re-drawn: Figure 5-A from reference [58], Figure 5-C from reference [54].

sequence homology to other known proteins. Hemophores are secreted via the Sec-independent pathway also known as type I secretion pathway (TISS) which is one of the five secretion systems identified in Gram-negative bacteria [60-62]. TISS help transport small and large molecules from the cytoplasm to the extracellular medium. The architecture of the TISS consists of three components: (a) cytoplasmic membrane bound ATP-binding cassette (ABC) class of protein, comprised of a nucleotide binding domain (NBD), (b) an adaptor or a membrane fusion

protein (MFP) consisting of a cytoplasmic membrane anchor, (c) an outer membrane protein (OMP) belonging to TolC type receptors family (Figure 1.6).

This secretion apparatus is also known as ABC exporter and used by many different families of proteins for the secretion of specific proteins, including virulence factors, hydrolytic enzymes and hemophores. [60, 61].

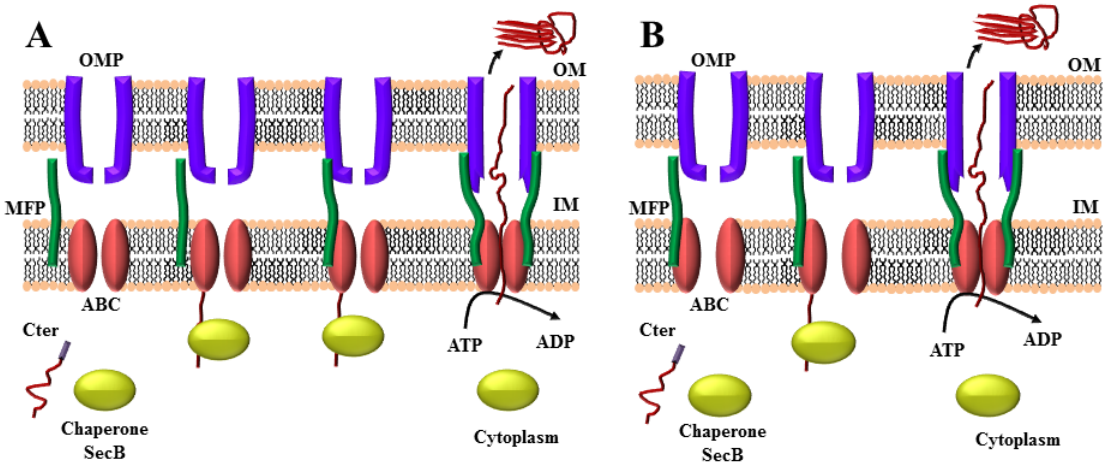


Figure 1.6. Schematic representation of the type I secretion pathway with the ABC/MFP complex either (A) not preformed or (B) preformed. C-terminus of the polypeptide interacts with the ABC protein and secretion of the polypeptide. Figure adapted and re-drawn from reference [60]. OMP: Outer membrane protein, MFP: Membrane fusion protein, IM: Inner membrane, OM: Outer membrane.

Hemophores (HasA) are secreted by ABC secretion pathway into the extracellular medium in a Sec-independent manner. Similar to other proteins secreted by this pathway, hemophores lack N-terminal secretion signal sequence. Instead a conserved secretion signal motif (Dxxx) is located within the last 50 amino acids residues near the C-terminal end [63] (Figure 1.7, red rectangle). The C-terminal signal peptide is important as it directs the secretion of the protein and specifically interacts with the ABC protein and modulates its ATPase activity [62] (Figure 1.6). Moreover, this interaction triggers the assembly of ABC, MFP and OMP

proteins into a functional complex. The specific interaction of C-terminal signal peptide with the ABC protein implies that the protein is fully synthesized before being exported out in the extracellular medium [62]. In case of hemophores, synthesis and secretion are coupled, a chaperone SecB binds and slows down folding of HasA, keeping the newly synthesized polypeptide chain in a conformation that is competent for secretion [61]. Two problems arise in the event of uncoupling between the synthesis and secretion; (i) HasA folds in the cytoplasm and binds heme from other cytoplasmic proteins, and (ii) folded HasA interacts with the transporter *in vivo*, and inhibits secretion of newly synthesized HasA molecules [64, 65].

*Haemophilus influenza*e, which lack the heme biosynthetic pathway, utilizes heme-hemopexin complex as a heme source. It secretes a truncated (21 residues of the signal peptide) hemophore HuxA which strongly binds the HPX with or without heme. The bound HPX complex to the HuxA is then transferred to the outer membrane receptor, HuxC, for internalization [55, 66].

1.4.4 Structural Characterization of Hemophore HasA

The apparatus consists of a secreted hemophore HasA and an outer membrane receptor HasR. The receptor HasR can bind heme directly but the efficiency of the heme uptake is increased ~ 100 fold in conjugation with hemophore HasA [50]. The HasA from *Serratia marcescens* (HasAs) was the first hemophore to be identified and thoroughly characterized from structural and functional point of view [67]. Recent biochemical and biophysical characterization of hemophores has been done for *Pseudomonas aeruginosa* (HasAp) [68-70] and *Yersinia pestis* ($HasA_{yp}$) [36].

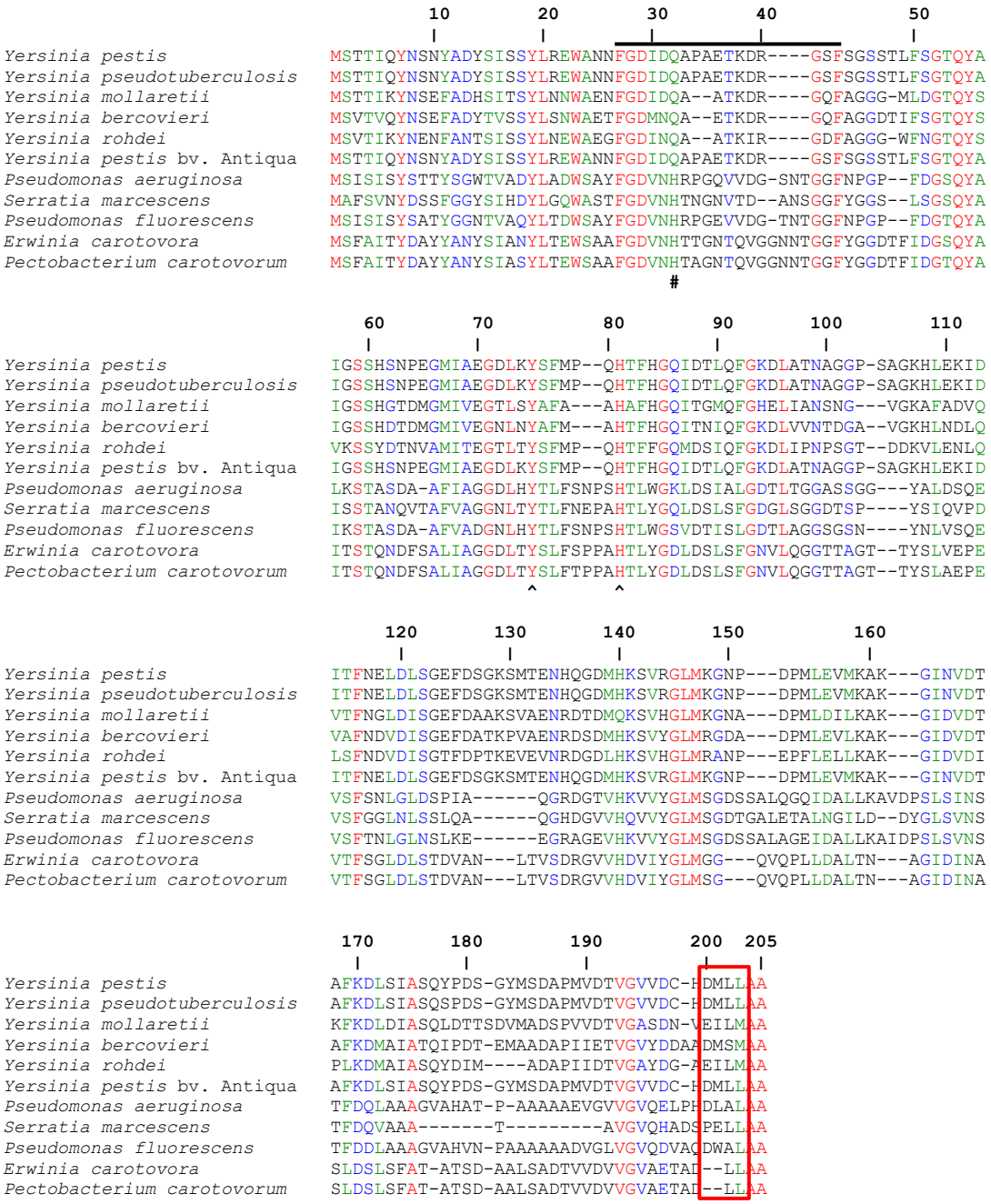


Figure 1.7. Full-length HasA sequences aligned against the sequence of HasA from *Yersinia pestis*. Residue 32 (H or Q) is highlighted by (#) and proximal loop residues Tyr75 and His83 are highlighted by (^). A horizontal line above the sequence highlights residues comprising the distal loop. Conserved residues across the alignment are in red, conservative substitutions in green and semi-conservative substitutions in blue; sequence numbering is according to the *Yersinia pestis* sequence [36].

1.4.4.1 Serratia marcescens hemophore: HasAs

The first available X-ray crystal structure of heme-bound (holo) hemophore was from *Serratia marcescens* (HasAs) [71]. Structurally and biochemically, it is the most well characterized hemophore. The monomeric protein is composed of 188 amino acid residues with a molecular mass of ~ 19 kDa. The crystal structure determined at 1.9 Å resolution [71] of HasAs (PDB code: 1B2V) revealed a novel protein fold that binds *b*-type heme in a 1:1 stoichiometry and with a high affinity ($K_d \sim 10^{-11}$ M) [72]. The overall protein structure, as mentioned by the authors [71], could be imagined as a shape of a fish, with the mouth of the fish representing the heme binding site. The protein is composed of α-, β-secondary structural elements, with one side comprised of four α-helices and the other, composed of seven antiparallel β-strands. Heme is coordinated by a unique pair of ligands, Tyr75 and His32, with each axial ligand harbored between the long loops, termed as the Y75 loop and the H32 loop (Figure 1.8-A and 1.8-B). Structurally, close to Tyr75 in the Y75 loop, His83 is present which forms a tight hydrogen bond (2.8 Å) with the Oη of Tyr75 via its Nδ1 NH group. It has been proposed that this hydrogen bond interaction modulates the Tyr-Fe bond, thereby increasing the nucleophilic character of Tyr75 [73, 74]. Similar structural features have been characterized for the holo hemophore from *P. aeruginosa* (HasAp) (Figure 1.8-C and 1.8-D) [68]. Similar to the holo-hemophore, solution NMR structure of apo-HasAs [72] and X-ray crystal structure of apo-HasAp [69] is available, which shows overall similar polypeptide fold compared to the holo-counterpart except for the H32 loop which is wide open and folds back onto the body of the protein. The axial ligand His32 is located ~ 30 Å away relative to the holo-protein (Figure 1.8-A and 1.8-C). Recent crystal structure of apo- and holo-hemophore from *Yersinia pestis* ($HasA_{yp}$)

presents a different conformation of the distal loop which is discussed in Chapter II of this dissertation.

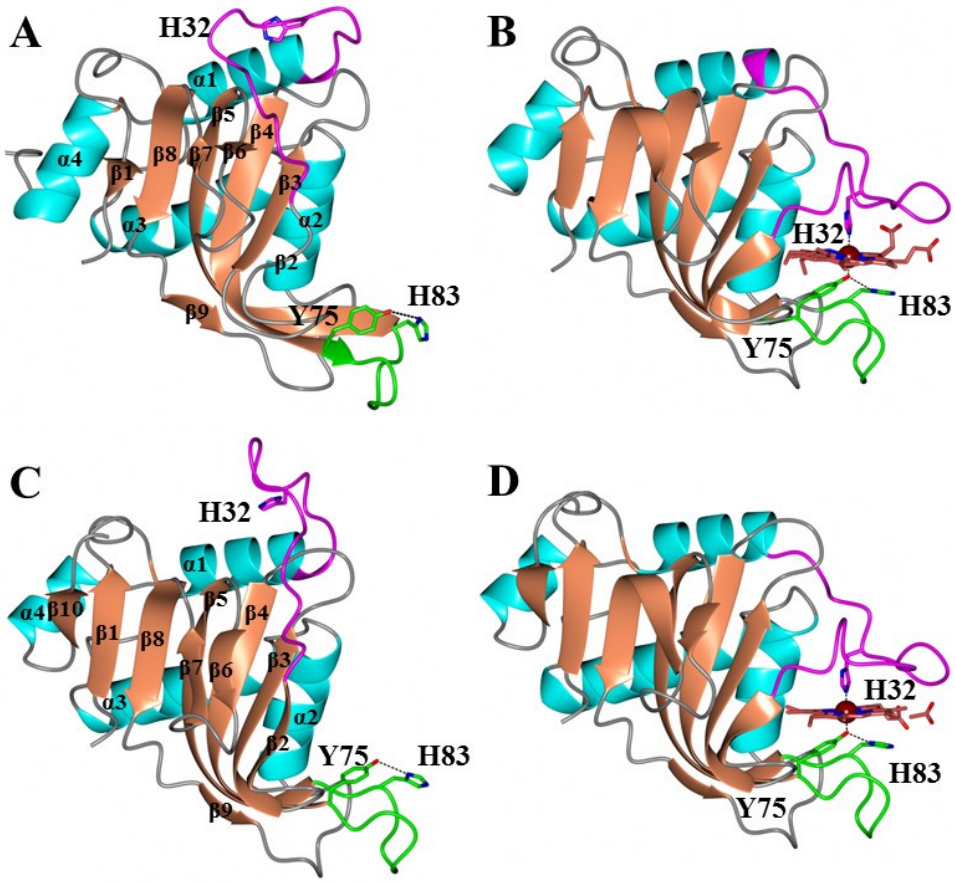


Figure 1.8. Structural comparison of apo- and holo-HasA from (A and B) *Serratia marcescens* and (C and D) from *Pseudomonas aeruginosa*. α-helices are colored cyan and β-strands in coral color. Distal (H32) loop harboring the His32 ligand is colored magenta, proximal (Y75) loop harboring the Tyr75 ligand is colored green. PDB codes: (A) apo-HasAs, 1Y2B, (B) holo-HasAs, 1B2V, (C) apo-HasAp, 3MOK and (D) holo-HasAp (3ELL).

1.4.5 Mechanism of Heme loading

Initially, Wolff and co-workers proposed an induced fit model involving a two-step hemin loading mechanism, where hemin first interacts and loads onto the Y75 loop followed by the

closure of the H32 loop thus coordinating the hemin-iron [72]. Gram-negative hemophore proteins having a His/Tyr ligand coordination to a ferric iron ion [75]. First evidence of Tyr ligation comes with the works of Lukat-Rogers and co-workers [75], upon reduction with dithionite in the presence of CO in holo-HasAs led to a TyrO-Fe(II)-CO complex formation with the displacement of the His32 residue from the distal loop. Further evidence comes from the crystal structure of HasAs in complex with its receptor (HasR), where the Tyr75 is coordinating the hemin-iron and the His32 is away from the hemin-binding conformation [76]. The above-mentioned two studies in HasAs suggested that His32 is a more labile ligand and is readily displaced [75]. A detailed structural investigation of the role of His32 side chain in hemin binding has been done using the H32A mutant of HasAp [69, 70]. The X-ray crystal structure of H32A holo-HasAp revealed a non-crystallographic dimeric molecule. Each of the two subunits coordinate the heme-iron ion by Tyr75 and the distal face of the heme molecules are stacked upon each other. The conformation of the A32 loop is identical to the H32 loop and the structure of each subunit is superimposable to the wt apo-HasAp structure (Figure 1.9). Kinetics studies done on the wt apo-HasAp using stopped-flow, EPR and resonance Raman, identifies biphasic hemin loading (i) where the first phase is completed within 20 ms and Tyr75 coordinates the hemin-iron ion from the Y75 loop and (ii) second phase completed in ~ 1 s with the closure of H32 loop and coordinating the heme-iron by His32 from the distal loop (Figure 1.10). Kinetic measurements done on H32A HasAp mutant revealed only the first phase corresponding to the hemin loading onto the Y75 loop [70]. Taken together, the structural and spectroscopic observations suggested that the hemin first loads onto the Y75 loop followed by the closure of H32 loop which is accompanied by the large conformational rearrangement [70].

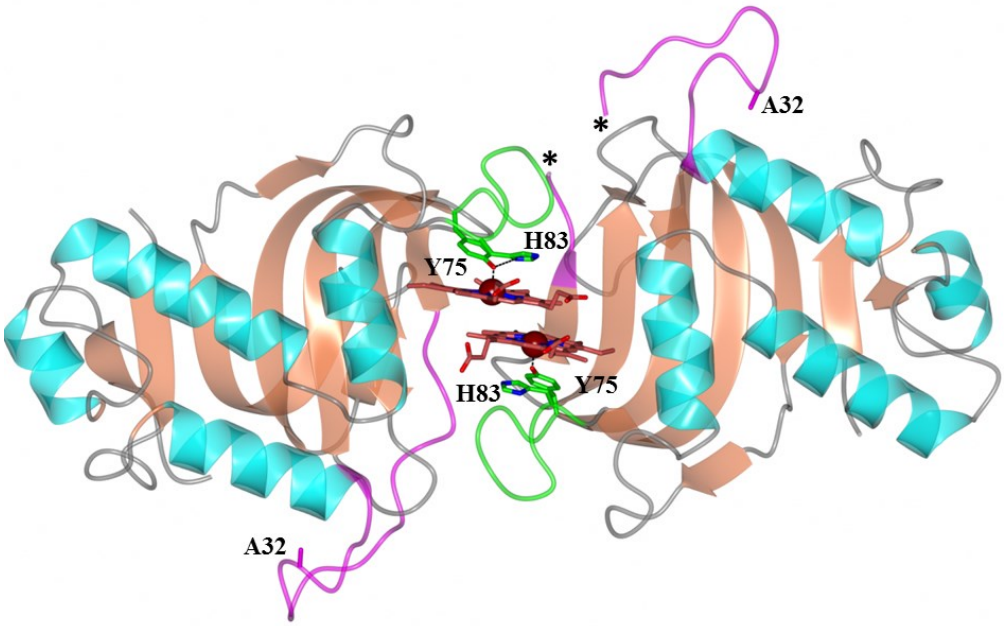


Figure 1.9. Crystal structure of H32A holo-HasAp dimer (PDB code: 3MOL) showing co-facial stacking of heme molecules. H32 loop is colored magenta and Y75 loop in green color, missing electron density in subunit B is located between the asterisks [69].

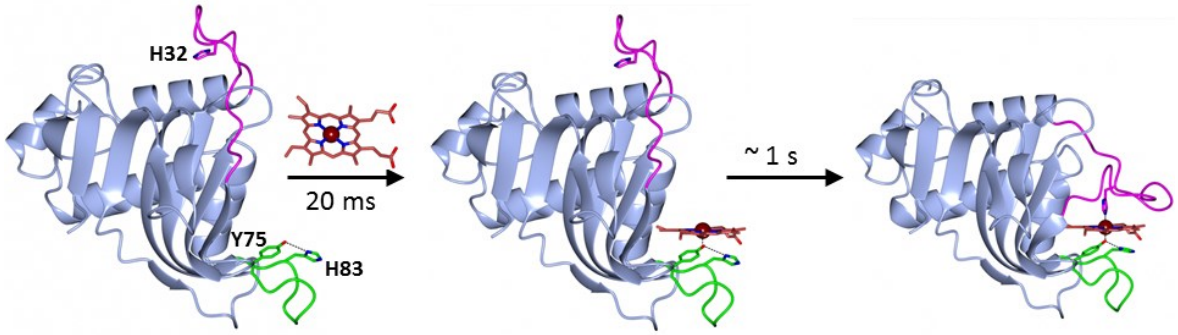


Figure 1.10. Schematic representation of heme loading mechanism of HasA using wt apo- and holo X-ray crystal structures from *P. aeruginosa*. Initial binding of heme is complete within 20 ms and is coordinated by Tyr75 followed by the closure of H32 loop thus completing the process in ~ 1 s. H32 loop is shown in magenta, Y75 loop in green [77]. Figure adapted from reference [31].

1.4.6 Interaction of HasA with the Receptor HasR

The detailed mechanism of heme delivery to the receptor HasR is not clearly understood. Co-crystal structures of holo-HasAs and apo-HasAs bound to HasR (PDB code: 3CSL and 3CSN) and holo-HasA bound to the I167G mutant of HasR (PDB code: 3DDR) are available, and provide some insight for the delivery of heme from holo-HasA to HasR (Figure 1.11). Protein-protein interactions have been suggested as the driving force for the heme transfer from the holo-HasAs to the receptor HasR, which induces conformational changes leading to the weakening/ or breakage of the Tyr75-Fe coordination bond and moving of the heme from HasAs to HasR coordinated by His189 and His609 residues [61, 76].

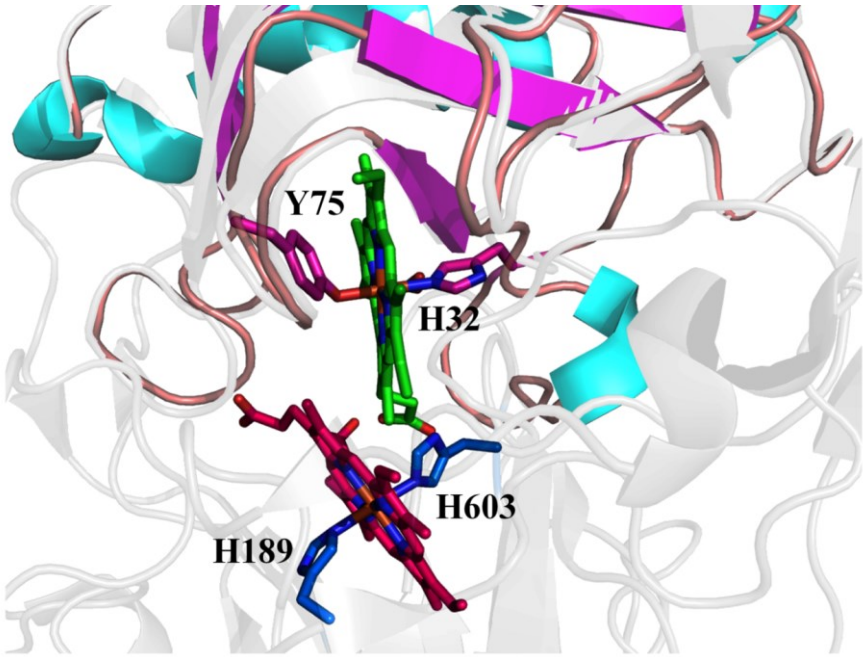


Figure 1.11. Superimposed crystal structures of holo-HasA (PDB code: 1B2V) and of HasA-hemin-HasR complex (PDB code: 3CSL) from *Serratia marcescens*. The structure of the HasA-hemin-HasR complex is shown in grey at 80% transparency. The HasR heme ligands His189 and His603 are shown in blue. Figure adapted from reference [31].

1.4.7 Hemin in Periplasm and Further Translocation

Once in the periplasm, periplasmic binding proteins (PBP) serve a crucial function of capturing the heme from the outer membrane receptor, HasR for further transport across the periplasmic space (Figure 1.3-A). Crystal structures are available for the PBPs of *Pseudomonas aeruginosa* (PhuT) [45], *Shigella dysenteriae* (ShuT) [45] and *Yersinia pestis* (HmuT) [78]. Structurally, the N- and C-terminal domains are connected by a long α-helix and the heme loads between the two domains coordinated by a Tyr residue (figure 1.12) except for HmuT in which two molecules of heme are bound to the structure coordinated by Tyr70 and His167 residues (Figure 1.12-C). The structural information provides an insight into the PBP heme-binding protein which mediates heme delivery through the ABC transporter across the inner membrane.

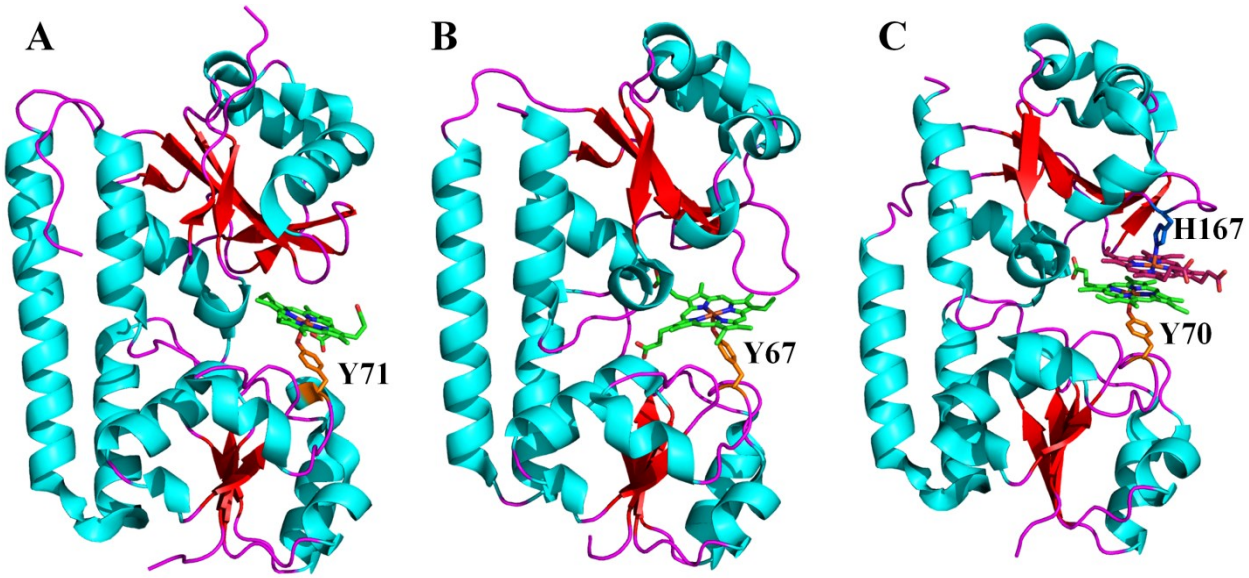
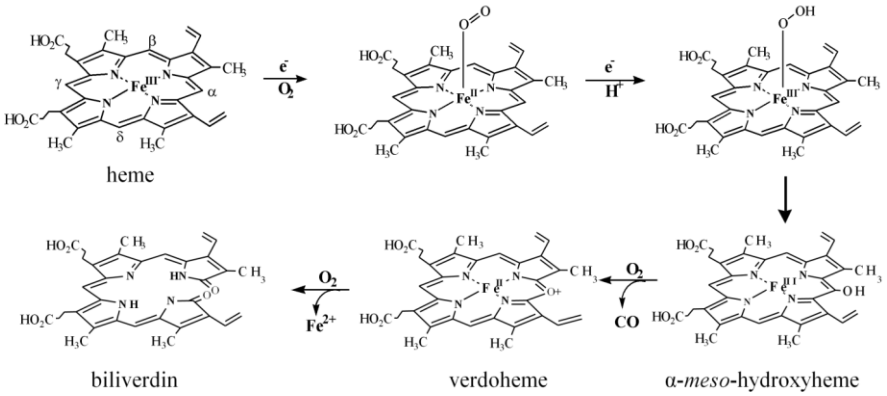


Figure 1.12. Crystal structures of heme binding PBPs (A) PhuT (PDB code: 2R79), (B) ShuT (PDB code: 2RG7) and, (C) HmuT (PDB code: 3NU1). The structure of HmuT (C) has two heme molecules coordinated by Tyr70 and His167 ligands. Figure adapted from reference [31].

1.4.8 Iron Release and Fate of Heme in the Cytosol

Once the macrocycle is in the cytoplasm, chaperone PhuS sequesters the heme and transfers it to the heme degrading enzyme (heme oxygenase, HO) (Figure 1.3). HO enzymes catalyze the breakage of the heme macrocycle releasing the iron for the metabolic needs or for storage in the bacteria. The HO mediated degradation of heme to release iron is summarized in scheme 1. Briefly, the process begins with the reduction of Fe^{3+} iron ion to Fe^{2+} which enables an oxygen molecule to bind at the distal site, forming an oxyferrous complex (Fe^{2+}-O_2). The oxyferrous complex upon accepting a second electron and a proton is converted into an activated ferric hydroperoxy (Fe^{3+}-OOH) oxidizing species. Recent findings in *Pseudomonas aeruginosa* demonstrate that the electron donor is a ferredoxin $NADP^+$ reductase (*pa*-FPR), a NADPH-dependent flavoenzyme [31, 79]. The Fe^{3+}-OOH intermediate in HO reacts with α-*meso* carbon in heme to produce ferric α-*meso*-hydroxyheme that undergoes a rapid O_2-dependent elimination of carbon monoxide (CO) to form verdoheme. Verdoheme is subsequently oxidized to Fe^{3+}-biliverdin by an oxygen molecule breaking the macrocycle to release iron ion [31].



Scheme 1. Schematic representation of the heme degradation carried out by the heme oxygenase (HO) enzymes. Scheme adapted from reference [31].

1.5 Research Problem and Rationale

The ubiquitous role of iron is important for many physiological functions. Despite its importance, the redox pair of Fe^{2+}/ Fe^{3+} poses unique challenges to the living cells. The insolubility of the ferric form (Fe^{3+}) at physiological pH and the toxicity of the ferrous form (Fe^{2+}) generating free radicals via a Fenton-type reaction, are overcome by sequestering iron in high affinity heme binding proteins in vertebrate cells. Similar to iron, heme is also toxic to cells and is quickly sequestered by hemopexin and serum albumin heme binding proteins, keeping the free iron levels below 10^{-18} M in the plasma.

Low concentrations of free or bound iron in blood plasma poses a significant challenge for the pathogenic bacteria. The need to thrive in a hostile environment has led Gram-negative pathogens to evolve various iron/ heme scavenging strategies. These strategies can be broadly classified into (a) secretion of small molecular weight molecules called "siderophores" which bind ferric-iron with a high affinity, (b) direct binding of the iron-complexed proteins to the outer membrane receptors and (c) secretion of heme-binding proteins called "hemophores" that bind the heme and subsequently deliver it to their cognate outer membrane receptor. It has been established that iron acquisition is critical to bacterial survival during the course of infection. Hemophore dependent heme acquisition system has been structurally characterized in *Serratia marcescens* [71, 80], *Pseudomonas aeruginosa* [53, 68, 69, 81] and *Yersinia pestis* [36].

Yersinia pestis is a Gram-negative rod-shaped bacterium belonging to the *Enterobacteriaceae* family. It is a causative agent of bubonic plague and has a dual, mammal/ flea, life cycle. Similar to other Gram-negative pathogenic bacteria, *Yersinia pestis* also requires iron for its metabolic needs. An iron deficient environment in the mammalian hosts poses a

challenge to acquire and utilize iron for its metabolic needs. *Y. pestis* uses heme and a wide variety of host heme containing protein complexes including hemoglobin, hemopexin, hemoglobin-haptoglobin, myoglobin and serum albumin as sources to iron.

1.5.1 Organization of Chapters in this Dissertation

Chapter 2: Structurally characterized hemophores have an unusual set of heme binding ligand (His/Tyr) which bind and coordinates the heme between the proximal (Y75) and the distal (H32) loops. With the first X-ray crystal structure from *S. marcescens* and then from *P. aeruginosa*, this coordination pair was thought to operate in all Gram-negative pathogens with an induced fit binding of the heme macrocycle sandwiched between the His32 and Tyr75 residues. A closer look at the multiple sequence alignment (Figure 7) of hemophores from different Gram-negative pathogens reveals that the Y75 loop residues, Tyr75 and its hydrogen bond partner His83 (His81 in $HasA_{yp}$), are conserved across the alignment, whereas the His32 is not. Instead in *Yersinia* species, there is Gln32 present instead of His32 in the distal loop. Given the structural similarity between the known hemophore *Serratia marcescens* (HasAs) and *Pseudomonas aeruginosa* (HasAp), hypotheses were formulated for *Yersinia pestis*: (a) the heme iron would be coordinated by a novel set of ligands (Gln32/Tyr75), (b) there will be structural compensation for His32 to Gln32 residue change and (c) the kinetics of heme loading be similar to that of HasAp or completely different.

In order to investigate above hypotheses, we used chromatographic methods to isolate and purify the hemophore from *Yersinia pestis* (KIM10+). Crystals were grown for both apo- and holo-$HasA_{yp}$ proteins which were then characterized by X-ray crystallography to elucidate the structural features. The detailed results and discussions are presented in Chapter II of this dissertation.

Chapter 3: X-ray crystal structure of wt apo-HasAp structure shows that the H32 loop is wide open and the His32 is located approximately ~ 30 Å away from the heme binding site. Similar observation was complemented from the solution NMR study of the apo-protein. These two analyses suggest that the H32 loop is conformationally stable and do not show flexibility in the apo-form. The stability of the H32 loop in the apo-protein could be visualized as a "zipper-like" network of interactions holding the H32 loop in the wide open conformation. Upon heme loading, the intramolecular interaction triggers the H32 loop closure, suggesting a large conformational rearrangement in the H32 loop thus, coordinating the heme-iron by His32. Combining the available structural information from *Serratia marcescens*, *Pseudomonas aeruginosa* and *Yersinia pestis* hemophores, the conformational changes were mapped onto the apo-HasAp protein structure to understand the H32 loop "rigid" open structure. Mapping the structural changes shows the presence of Arg33 in a very strategic location over the body of the hemophore HasAp. Hypotheses were formulated suggesting that Arg33 is the key to anchoring the open conformation of the His32 loop. To test the hypotheses, we replaced Arg33 with Ala33. The mutated protein was purified and then screened for crystal growth. Uniformly labeled ^{15}N-mutated proteins were prepared, which were investigated using solution NMR spectroscopy. Chapter III discusses the detailed structural investigation of R33A HasAp mutant proteins.

Chapter 4: The amino acid sequence of known hemophores (Figure 7) suggests a strong conservation of Tyr75 and His83 (His81 in $HasA_{yp}$) residues. From the available crystal structures (HasAp, $HasA_{yp}$ and HasAs), the location of these two residues on the Y75 loop is near to each other (Figure 8 A-D). Tyr75 (O_η) engages in a hydrogen bond interaction with the His83 (N_δ) (His81 in $HasA_{yp}$) and the interaction is thought to increase the phenolate character of

Try75. It has been proposed that this hydrogen bond interaction between the Tyr75 and His83 (His81 in $HasA_{yp}$) is essential for proximal heme coordination by Tyr75 [68, 73, 74]. Given the structural conservation of Tyr75 and His83 and the proposed requirement for the Tyr75-His83 hydrogen bond in heme coordination by Tyr75, we carried out structural and spectroscopic investigation of the Y75 loop conserved residues in the *P. aeruginosa*. To investigate this hydrogen bond requirement, we replaced Tyr75 and His83 for Alanine. Detailed structural and kinetic investigations of the Y75A and H83A mutants and discussion of the observations are summarized in Chapter IV.

Chapter 5: The importance of current findings in the context of past and present knowledge is discussed. Future directions and avenues for further research in this area are suggested along with the dissertation summary.

1.6 Bibliography

1. Lambert PA. Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. *Journal of the Royal Society of Medicine*. 2002;95 Suppl 41:22-6.
2. http://www.polyphor.com/products/pseudomonas-infections/infections-caused-by-pseudomonas-aeruginosa.
3. Skaar EP. The battle for iron between bacterial pathogens and their vertebrate hosts. *PLoS Pathog*. 2010;6(8):e1000949.
4. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial Iron Homeostasis. *FEMS Microbiology Reviews*. 2003;27:215-37.
5. Nobles CL, Maresso AW. The theft of host heme by Gram-positive pathogenic bacteria. *Metallomics : integrated biometal science*. 2011 Aug;3(8):788-96.
6. Wandersman C, Delepelaire P. Haemophore functions revisited. *Molecular microbiology*. 2012 Aug;85(4):618-31.
7. Pierre JL, Fontecave M. Iron and activated oxygen species in biology: the basic chemistry. *Biometals*. 1999 Sep;12(3):195-9.
8. Wandersman C, Delepelaire P. Bacterial iron sources: from siderophores to hemophores. *Annual review of microbiology*. 2004;58:611-47.
9. Faraldo-Gomez JD, Sansom MS. Acquisition of siderophores in gram-negative bacteria. *Nature reviews Molecular cell biology*. 2003 Feb;4(2):105-16.
10. Schalk IJ, Guillon L. Fate of ferrisiderophores after import across bacterial outer membranes: different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways. *Amino acids*. 2013 May;44(5):1267-77.

11. Braun V, Killmann H. Bacterial solutions to the iron-supply problem. *Trends Biochem Sci*. 1999 Mar;24(3):104-9.

12. Cornelis P, Dingemans J. Pseudomonas aeruginosa adapts its iron uptake strategies in function of the type of infections. *Frontiers in cellular and infection microbiology*. 2013;3:75.

13. Wolz C, Hohloch K, Ocaktan A, Poole K, Evans RW, Rochel N, Albrecht-Gary AM, Abdallah MA, Doring G. Iron release from transferrin by pyoverdin and elastase from Pseudomonas aeruginosa. *Infect Immun*. 1994 Sep;62(9):4021-7.

14. Wilderman PJ, Vasil AI, Johnson Z, Wilson MJ, Cunliffe HE, Lamont IL, Vasil ML. Characterization of an endoprotease (PrpL) encoded by a PvdS-regulated gene in Pseudomonas aeruginosa. *Infect Immun*. 2001 Sep;69(9):5385-94.

15. Bjorn MJ, Iglewski BH, Ives SK, Sadoff JC, Vasil ML. Effect of iron on yields of exotoxin A in cultures of Pseudomonas aeruginosa PA-103. *Infect Immun*. 1978 Mar;19(3):785-91.

16. Calderwood SB, Mekalanos JJ. Iron regulation of Shiga-like toxin expression in Escherichia coli is mediated by the fur locus. *J Bacteriol*. 1987 Oct;169(10):4759-64.

17. Coulanges V, Andre P, Ziegler O, Buchheit L, Vidon DJ. Utilization of iron-catecholamine complexes involving ferric reductase activity in Listeria monocytogenes. *Infect Immun*. 1997 Jul;65(7):2778-85.

18. Andrews SC, Robinson AK, Rodriguez-Quinones F. Bacterial iron homeostasis. *FEMS Microbiol Rev*. 2003 Jun;27(2-3):215-37.

19. Brot N, Goodwin J. Regulation of 2,3-dihydroxybenzoylserine synthetase by iron. *J Biol Chem*. 1968 Feb 10;243(3):510-3.

20. Bryce GF, Brot N. Iron transport in Escherichia coli and its relation to the repression of 2,3-dihydroxy-N-benzoyl-L-serine synthetase. *Arch Biochem Biophys*. 1971 Feb;142(2):399-406.

21. Troxell B, Hassan HM. Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Frontiers in cellular and infection microbiology*. 2013;3:59.

22. Cornelis P, Bodilis J. A survey of TonB-dependent receptors in fluorescent pseudomonads. *Environmental microbiology reports*. 2009 Aug;1(4):256-62.

23. Cornelis P. Iron uptake and metabolism in pseudomonads. *Appl Microbiol Biotechnol*. 2010 May;86(6):1637-45.

24. Ochsner UA, Vasil AI, Vasil ML. Role of the ferric uptake regulator of Pseudomonas aeruginosa in the regulation of siderophores and exotoxin A expression: purification and activity on iron-regulated promoters. *J Bacteriol*. 1995 Dec;177(24):7194-201.

25. Dumas Z, Ross-Gillespie A, Kummerli R. Switching between apparently redundant iron-uptake mechanisms benefits bacteria in changeable environments. *Proceedings Biological sciences / The Royal Society*. 2013 Aug 7;280(1764):20131055.

26. Perry RD, Balbo PB, Jones HA, Fetherston JD, DeMoll E. Yersiniabactin from Yersinia pestis: biochemical characterization of the siderophore and its role in iron transport and regulation. *Microbiology*. 1999 May;145 (Pt 5):1181-90.

27. Krewulak KD, Vogel HJ. Structural biology of bacterial iron uptake. *Biochimica et biophysica acta*. 2008 Sep;1778(9):1781-804.

28. Anzaldi LL, Skaar EP. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infect Immun*. 2010 Dec;78(12):4977-89.

29. Ponka P. Cell Biology of Heme. *Am J Med Sci*. 1999;318:241-56.

30. Ascenzi P, Bocedi A, Visca P, Altruda F, Tolosano E, Beringhelli T, Fasano M. Hemoglobin and heme scavenging. *IUBMB life*. 2005 Nov;57(11):749-59.

31. Benson DR, Rivera M. Heme Uptake and Metabolism in Bacteria. Met Ions Life Sci. 122013. p. 279-332.

32. Wilks A, Burkhard KA. Heme and virulence: how bacterial pathogens regulate, transport and utilize heme. *Nat Prod Rep*. 2007 Jun;24(3):511-22.

33. Adams PA, Berman MC. Kinetics and mechanism of the interaction between human serum albumin and monomeric haemin. *Biochem J*. 1980 Oct 1;191(1):95-102.

34. Hrkal Z, Muller-Eberhard U. Partial characterization of the heme-binding serum glycoproteins rabbit and human hemopexin. *Biochemistry*. 1971 May 11;10(10):1746-50.

35. Morgan WT, Muller-Eberhard U. Interactions of porphyrins with rabbit hemopexin. *J Biol Chem*. 1972 Nov 25;247(22):7181-7.

36. Kumar R, Lovell S, Matsumura H, Battaile KP, Moenne-Loccoz P, Rivera M. The hemophore HasA from Yersinia pestis (HasAyp) coordinates hemin with a single residue, Tyr75, and with minimal conformational change. *Biochemistry*. 2013 Apr 23;52(16):2705-7.

37. Smith AD, Wilks A. Extracellular heme uptake and the challenges of bacterial cell membranes. *Current topics in membranes*. 2012;69:359-92.

38. Skaar EP, Gaspar AH, Schneewind O. IsdG and IsdI, Heme-Degrading Enzymes in the Cytoplasm of *Staphylococcus aureus*. *J Biol Chem*. 2004;279:436-43.

39. Wandersman C, Stojiljkovic I. Bacterial heme sources: the role of heme, hemoprotein receptors and hemophores. *Current opinion in microbiology*. 2000 Apr;3(2):215-20.

40. Mills M, Payne SM. Identification of shuA, the gene encoding the heme receptor of Shigella dysenteriae, and analysis of invasion and intracellular multiplication of a shuA mutant. *Infect Immun*. 1997 Dec;65(12):5358-63.

41. Eakanunkul S, Lukat-Rodgers GS, Sumithran S, Ghosh A, Rodgers KR, Dawson JH, Wilks A. Characterization of the periplasmic heme-binding protein shut from the heme uptake system of Shigella dysenteriae. *Biochemistry*. 2005 Oct 4;44(39):13179-91.

42. Burkhard KA, Wilks A. 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*. 2007 May 18;282(20):15126-36.

43. Torres AG, Payne SM. Haem iron-transport system in enterohaemorrhagic Escherichia coli O157:H7. *Mol Microbiol*. 1997 Feb;23(4):825-33.

44. Ochsner UA, Johnson Z, Vasil ML. Genetics and regulation of two distinct haem-uptake systems, phu and has, in Pseudomonas aeruginosa. *Microbiology*. 2000 Jan;146 (Pt 1):185-98.

45. Ho WW, Li H, Eakanunkul S, Tong Y, Wilks A, Guo M, Poulos TL. Holo- and apo-bound structures of bacterial periplasmic heme-binding proteins. *The Journal of biological chemistry*. 2007 Dec 7;282(49):35796-802.

46. Lansky IB, Lukat-Rodgers GS, Block D, Rodgers KR, Ratliff M, Wilks A. The cytoplasmic heme-binding protein (PhuS) from the heme uptake system of Pseudomonas aeruginosa is an intracellular heme-trafficking protein to the delta-regioselective heme oxygenase. *J Biol Chem*. 2006 May 12;281(19):13652-62.

47. Hornung JM, Jones HA, Perry RD. The hmu locus of Yersinia pestis is essential for utilization of free haemin and haem--protein complexes as iron sources. *Mol Microbiol*. 1996 May;20(4):725-39.

48. Stojiljkovic I, Hantke K. Hemin uptake system of *Yersinia enterocolitica*: Similarities with other TonB-Dependent Systems in Gram-Negative Bacteria. *EMBO J*. 1992;11:4359-67.

49. Tong Y, Guo M. Cloning and characterization of a novel periplasmic heme-transport protein from the human pathogen Pseudomonas aeruginosa. *J Biol Inorg Chem*. 2007 Aug;12(6):735-50.

50. Ghigo JM, Letoffe S, Wandersman C. A new type of hemophore-dependent heme acquisition system of Serratia marcescens reconstituted in Escherichia coli. *Journal of bacteriology*. 1997 Jun;179(11):3572-9.

51. Letoffe S, Ghigo JM, Wandersman C. Iron acquisition from heme and hemoglobin by a Serratia marcescens extracellular protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1994 Oct 11;91(21):9876-80.

52. Letoffe S, Redeker V, Wandersman C. Isolation and characterization of an extracellular haem-binding protein from Pseudomonas aeruginosa that shares function and sequence similarities with the Serratia marcescens HasA haemophore. *Molecular microbiology*. 1998 Jun;28(6):1223-34.

53. Letoffe S, Omori K, Wandersman C. Functional characterization of the HasA(PF) hemophore and its truncated and chimeric variants: determination of a region involved in binding to the hemophore receptor. *J Bacteriol*. 2000 Aug;182(16):4401-5.

54. Rossi MS, Fetherston JD, Letoffe S, Carniel E, Perry RD, Ghigo JM. Identification and characterization of the hemophore-dependent heme acquisition system of Yersinia pestis. *Infect Immun*. 2001 Nov;69(11):6707-17.

55. Hanson MS, Slaughter C, Hansen EJ. The hbpA gene of Haemophilus influenzae type b encodes a heme-binding lipoprotein conserved among heme-dependent Haemophilus species. *Infect Immun*. 1992 Jun;60(6):2257-66.

56. Cope LD, Thomas SE, Hrkal Z, Hansen EJ. Binding of heme-hemopexin complexes by soluble HxuA protein allows utilization of this complexed heme by Haemophilus influenzae. *Infect Immun*. 1998 Sep;66(9):4511-6.

57. Gao JL, Nguyen KA, Hunter N. Characterization of a hemophore-like protein from Porphyromonas gingivalis. *J Biol Chem*. 2010 Dec 17;285(51):40028-38.

58. Rossi MS, Paquelin A, Ghigo JM, Wandersman C. Haemophore-mediated signal transduction across the bacterial cell envelope in Serratia marcescens: the inducer and the transported substrate are different molecules. *Mol Microbiol*. 2003 Jun;48(6):1467-80.

59. Biville F, Cwerman H, Letoffe S, Rossi MS, Drouet V, Ghigo JM, Wandersman C. Haemophore-mediated signalling in Serratia marcescens: a new mode of regulation for an extra cytoplasmic function (ECF) sigma factor involved in haem acquisition. *Mol Microbiol*. 2004 Aug;53(4):1267-77.

60. Delepelaire P. Type I secretion in gram-negative bacteria. *Biochim Biophys Acta*. 2004 Nov 11;1694(1-3):149-61.

61. Cescau S, Cwerman H, Letoffe S, Delepelaire P, Wandersman C, Biville F. Heme acquisition by hemophores. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*. 2007 Jun;20(3-4):603-13.

62. Sapriel G, Wandersman C, Delepelaire P. The N terminus of the HasA protein and the SecB chaperone cooperate in the efficient targeting and secretion of HasA via the ATP-binding cassette transporter. *J Biol Chem*. 2002 Feb 22;277(8):6726-32.

63. Ghigo JM, Wandersman C. A carboxyl-terminal four-amino acid motif is required for secretion of the metalloprotease PrtG through the Erwinia chrysanthemi protease secretion pathway. *J Biol Chem*. 1994 Mar 25;269(12):8979-85.

64. Debarbieux L, Wandersman C. Folded HasA inhibits its own secretion through its ABC exporter. *The EMBO journal*. 2001 Sep 3;20(17):4657-63.

65. Wolff N, Sapriel G, Bodenreider C, Chaffotte A, Delepelaire P. Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate. *J Biol Chem*. 2003 Oct 3;278(40):38247-53.

66. Hanson MS, Hansen EJ. Molecular cloning, partial purification, and characterization of a haemin-binding lipoprotein from Haemophilus influenzae type b. *Mol Microbiol*. 1991 Feb;5(2):267-78.

67. Létoffé S, Ghigo JM, Wandersman C. Iron Acquisition from Heme and Hemoglobin by *Serratia marcescens* Extracellular Protein. *Proc Natl Acad Sci USA*. 1994;91:9876-80.

68. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P, Rivera M. Structural characterization of the hemophore HasAp from Pseudomonas aeruginosa: NMR spectroscopy reveals protein-protein interactions between Holo-HasAp and hemoglobin. *Biochemistry*. 2009 Jan 13;48(1):96-109.

69. Jepkorir G, Rodriguez JC, Rui H, Im W, Lovell S, Battaile KP, Alontaga AY, Yukl ET, Moenne-Loccoz P, Rivera M. Structural, NMR spectroscopic, and computational

investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa. *Journal of the American Chemical Society*. 2010 Jul 21;132(28):9857-72.

70. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moenne-Loccoz P. Kinetic and spectroscopic studies of hemin acquisition in the hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 Aug 10;49(31):6646-54.

71. Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. The crystal structure of HasA, a hemophore secreted by Serratia marcescens. *Nature structural biology*. 1999 Jun;6(6):516-20.

72. Wolff N, Izadi-Pruneyre N, Couprie J, Habeck M, Linge J, Rieping W, Wandersman C, Nilges M, Delepierre M, Lecroisey A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*. 2008 Feb 15;376(2):517-25.

73. Caillet-Saguy C, Delepierre M, Lecroisey A, Bertini I, Piccioli M, Turano P. Direct-detected 13C NMR to investigate the iron(III) hemophore HasA. *J Am Chem Soc*. 2006 Jan 11;128(1):150-8.

74. Wolff N, Deniau C, Letoffe S, Simenel C, Kumar V, Stojiljkovic I, Wandersman C, Delepierre M, Lecroisey A. Histidine pK(a) shifts and changes of tautomeric states induced by the binding of gallium-protoporphyrin IX in the hemophore HasA(SM). *Protein Sci*. 2002 Apr;11(4):757-65.

75. Lukat-Rodgers GS, Rodgers KR, Caillet-Saguy C, Izadi-Pruneyre N, Lecroisey A. Novel heme ligand displacement by CO in the soluble hemophore HasA and its proximal ligand mutants: implications for heme uptake and release. *Biochemistry*. 2008 Feb 19;47(7):2087-98.

76. Krieg S, Huche F, Diederichs K, Izadi-Pruneyre N, Lecroisey A, Wandersman C, Delepelaire P, Welte W. Heme uptake across the outer membrane as revealed by crystal structures of the receptor-hemophore complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2009 Jan 27;106(4):1045-50.

77. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moënne-Loccoz P. Kinetic and Spectroscopic Studies of Hemin Acquisition in the Hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 2010/08/10;49(31):6646-54.

78. Mattle D, Zeltina A, Woo JS, Goetz BA, Locher KP. Two stacked heme molecules in the binding pocket of the periplasmic heme-binding protein HmuT from Yersinia pestis. *Journal of molecular biology*. 2010 Nov 26;404(2):220-31.

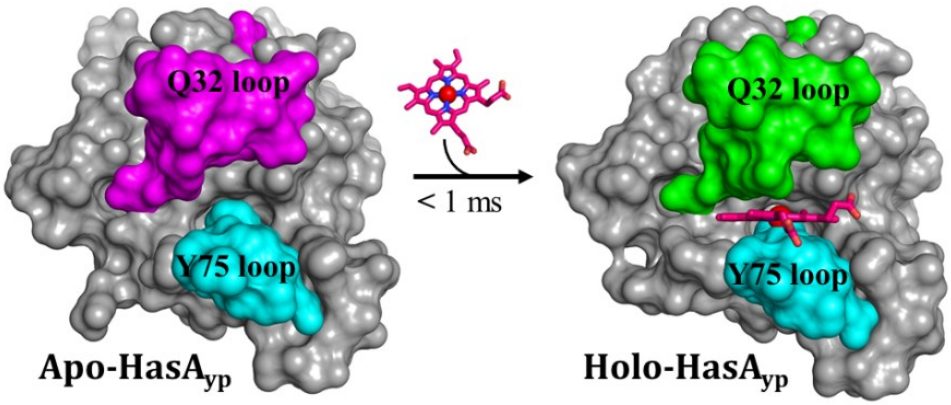
79. Wang A, Zeng Y, Han H, Weeratunga S, Morgan BN, Moënne-Loccoz P, Schönbrunn E, Rivera M. Biochemical and Structural Characterization of *Pseudomonas aeruginosa* Bfd and FPR: Ferredoxin $NADP^+$ Reductase and Not Ferredoxin is the Redox Partner of Heme Oxygenase under Iron-Starvation Conditions. *Biochemistry*. 2007;46:12198-211.

80. Létoffé S, Ghigo JM, Wandersman C. Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J Bacteriol*. 1994;176:5372-7.

81. Létoffé S, Redeker V, Wandersman C. Isolation and Characterization of an Extracellular Haem-Binding Protein from *Pseudomonas aeruginosa* that shares function and sequence similarities with the *Serratia Marcescens* HasA Haemophore. *Mol Microbiol*. 1998;28:1223-**4.**

Chapter 2

The Hemophore HasA from *Yersinia pestis* ($HasA_{yp}$) Coordinates Hemin with a Single Residue, Tyr75, and with Minimal Conformational Change



Summary

Under iron-limiting conditions several Gram-negative pathogens secrete a hemophore (HasA) to scavenge hemin from various iron binding/ containing proteins in the vertebrate host and deliver it to the specific outer membrane receptor (HasR) for internalization. Structurally characterized hemophores from *Serratia marcescens* (HasAs) and *Pseudomonas aeruginosa* (HasAp) bind hemin between two loops (Y75 and H32), which harbor the axial ligands His32 and Tyr75. In both the organisms, the hemophore has similar apo- and holo structures except the H32 loop which shows a large conformational rearrangement and relocates His32 ~ 30 Å from an open (apo) to a closed (holo) upon binding of the heme. Hemin binding to the Y75 loop triggers closing of the H32 loop and enables binding of His32. Amino acid sequence alignment of HasA proteins from *Yersinia* species show conservation of Tyr75 and His83 in the Y75 loop, but no conservation of the His32 residue, instead contain a Gln at position 32. In order to understand the structural compensation in the absence of His32, we investigated and carried out the structural characterization of the hemophores from Yersinia pestis ($HasA_{yp}$) in the heme-free (apo-) and heme-bound (holo-) forms. Surprisingly, the Q32 loop in apo-$HasA_{yp}$ is already in the closed conformation but no residue from the Q32 loop binds hemin in holo-$HasA_{yp}$. Moreover, Tyr75 is the only endogenous protein provided hemin ligand in holo-$HasA_{yp}$. In agreement with the minimal reorganization between the apo- and holo-structures, hemin binding to apo-$HasA_{yp}$ occurs on a sub-millisecond timescale undetected by conventional stopped-flow measurements.

2.1 Introduction

Although iron is important for many crucial biological functions, its chemical properties present unique challenges to living cells, which have to overcome the insolubility of Fe(III) and the toxicity of Fe(II) by sequestering iron in heme, iron-sulfur clusters or iron binding proteins. The very low concentrations of free iron in host cells pose significant challenges to pathogenic bacteria, which have evolved efficient strategies to scavenge iron, including secretion of hemophores, siderophores, hemolysins, proteases and cytotoxins [1, 2]. Given that ~70% of total iron is bound to hemoglobin, heme is an important iron source. Thus, a strategy used by bacteria to acquire heme is the deployment of hemophores, which are small proteins secreted in the extracellular milieu of the vertebrate hosts to capture free or protein-bound heme and deliver it to the outer membrane receptor of bacteria. The heme is then transferred from the outer membrane receptor to the cytosol via specific interactions with the proteins in the periplasm and the inner membrane [2, 3], where the macrocycle is degraded by heme-degrading enzymes to release the iron [1, 4, 5]. The released iron then can be either used for the metabolic purpose or stored in iron storage proteins like BfrB in the bacteria [6]. The HasA-type hemophore was first identified in *Serratia marcescens* [7] and then shown to be conserved in several Gram negative pathogens including, *Pseudomonas aeruginosa*, *Pseudomonas fluorescens, Yersinia pestis*, *Yersinia pseudotuberculosis*, *Erwinia carotovora* and *Pectobacterium carotovorum* [8-12]. Hemophores from *S. marcescens* (HasAs) [13, 14] and *P. aeruginosa* (HasAp) [8, 15] have been structurally characterized in their apo- and heme-bound (holo) forms and are found to be nearly identical. The X-ray crystal structure reveals a novel protein fold composed of a α+β type fold with the wall of α-helices on one side and wall of β-strand on the another. The hemophore protein binds heme in a 1:1 stoichiometry and the hemin iron in holo-HasAp (and HasAs) is coordinated by

His32 and Tyr75. Each axial ligand is harbored in a loop, termed either the H32 or Y75 loop. From the solution NMR studies of apo-HasAs [14] and X-ray crystal structure of apo-HasAp [8], the main difference between the apo- and holo-structures is a large rearrangement of the H32 loop, which relocates His32 ~30 Å upon loading of the heme (Figure 2.1).

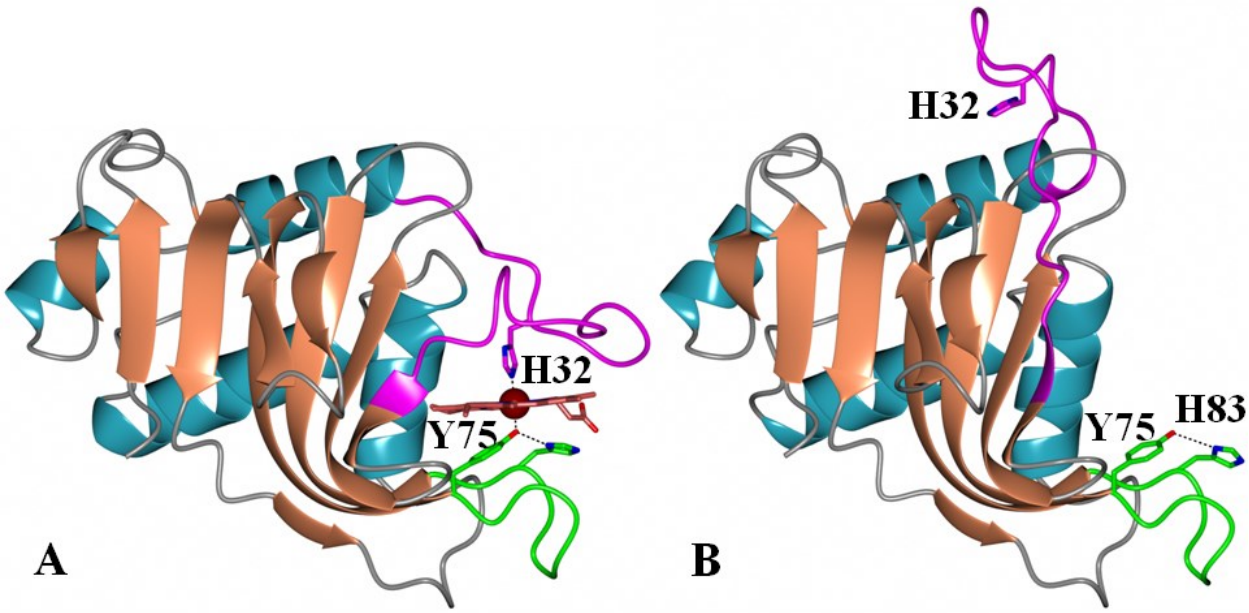


Figure 2.1. Structure of (A) holo-HasAp (PDB code: 3ELL) and (B) apo-HasAp (PDB code: 3MOK) showing the proximal (Y75) and distal (H32) ligands. The Y75 loop is shown in green, the H32 loop is in magenta and the heme is in red. Heme loads onto the Y75 loop; triggers closure of the H32 loop, and enables the His32 to coordinate the heme iron. Figure adapted from reference [16].

To probe which axial ligand (His32 or Tyr75) from the hemophore first comes in contact with the heme, structural and spectroscopic studies were carried out with the WT and H32A HasAp [8, 17]. A detailed spectroscopic investigation identifies biphasic hemin loading: (i) first

phase where hemin loads onto the Y75 loop within a few milliseconds and is likely rapidly coordinated by Tyr75, followed by a (ii) second phase where the coordination by His32 takes place, which is significantly slower (hundreds of milliseconds to second scale) [18]. In the X-ray crystal structure of H32A holo-HasAp in complex with imidazole, hemin-iron is coordinated by Tyr75 and imidazole, the H32 loop is near the distal heme face and adopts a conformation very similar to that of WT holo-HasAp; NMR studies suggest that this conformation is maintained in solution [8]. These findings led to the conclusion that hemin loading onto the Y75 loop triggers closing of the H32 loop. Results from targeted molecular dynamic simulations allowed identification of motions that are likely important for transmitting the presence of heme in the Y75 loop to the H32 loop in order to initiate its closing [8].

Given the unusual His/Tyr coordination of the hemin-iron and the induced fit closing of the H32 loop upon hemin loading onto the Y75 loop of HasAp or HasAs, it is intriguing that His32 is not conserved amongst HasA proteins. Multiple sequence alignment of hemophore sequences from different Gram-negative pathogens reveals that the Y75 loop residues, Tyr75 and His83 (His81 in $HasA_{yp}$), is conserved across the alignment whereas His32 is not conserved. Instead in *Yersinia* species, (Figure 2.2) there is a Gln at position 32 and do not have a His residue close in the sequence that could coordinate the hemin iron. These observations suggest that if the hemophore structures of *Yersinia* species are similar to HasAp and HasAs, then the hemin iron would be axially coordinated by an unprecedented set of ligands, Gln32 and Tyr75. In order to understand how the structures of hemophores compensate for the absence of H32, we carried out the structural characterization of the hemophore from *Yersinia pestis* KIM10+ ($HasA_{yp}$) in its heme-free (apo-) and heme-bound (holo-) forms.

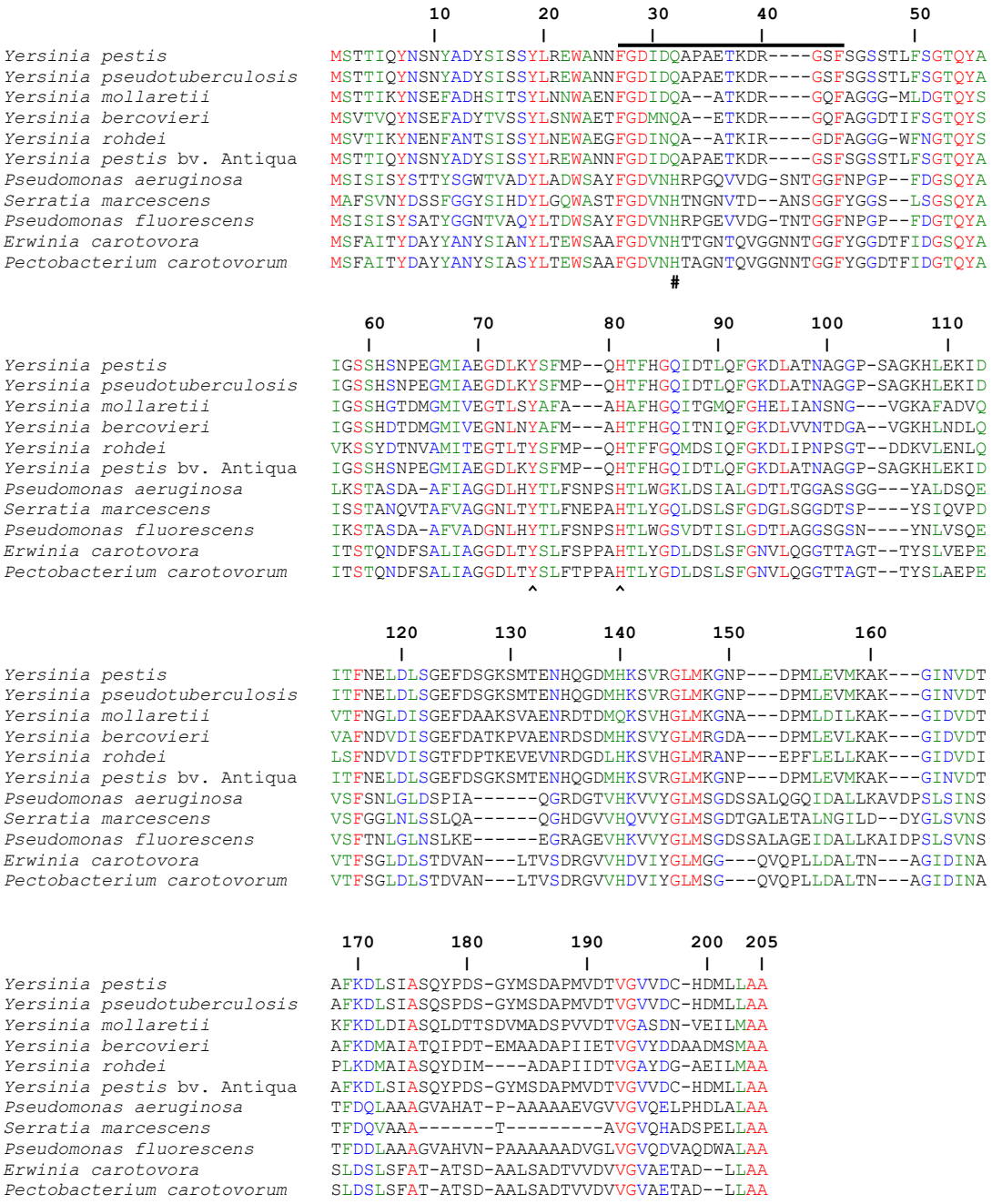


Figure 2.2. Full-length HasA sequences aligned against the sequence of HasA from *Yersinia pestis*. Residue 32 (H or Q) is highlighted by (#) and proximal loop residues Tyr75 and His83 are highlighted by (^). A horizontal line above the sequence highlights residues comprising the distal loop. Conserved residues across the alignment are in red, conservative substitutions in green and semi-conservative substitutions in blue; sequence numbering is according to the *Yersinia pestis* sequence. Sequence alignment done using CLUSTAL W [19], figure adapted from reference [16].

2.2 Experimental Methods

2.2.1 Cloning of the $HasA_{yp}$ Gene

A synthetic gene coding for full-length *Yersinia pestis* HasA ($HasA_{yp}$) was synthesized and sub-cloned into the pET11a expression vector by GenScript (Piscataway, NJ). Silent mutations to introduce codons favored by *Escherichia coli* were engineered into the $HasA_{yp}$ gene [20]. Restriction sites *Nde*I and *Bam*HI were introduced at the 5' and 3' ends of the gene, respectively (Table 2.1), to facilitate sub-cloning into the pET11a vector. The recombinant plasmid harboring the $HasA_{yp}$ gene was then transformed into the *E. coli* BL21-GOLD (DE3) competent cells (Stratagene, La Jolla, CA). A gene coding for truncated $HasA_{yp}$ (full-length – 12 residues) was constructed by appropriately placing a stop codon in the recombinant pET11a plasmid harboring the gene coding for full-length $HasA_{yp}$ using the QuickChange mutagenesis kit from Stratagene (La Jolla, CA). The oligonucleotides were synthesized by Integrated DNA Technologies, Inc. and were used without further purification. The primers used were 5'-GCCCCGATGGTTGATACGGTTTAAGTTGTTGATTGTCACGATATG-3' and 5'-CATATCGTGACAATCAACAACTTAAACCGTATCAACCATCGGGGC-3'; the underlined codons represent mismatches that introduce a stop codon after Val 193 (see Table 2.1). The mutation was confirmed by sequencing (ACGT, Inc., IL) and the recombinant DNA plasmid was transformed into *Escherichia coli* BL21-GOLD (DE3) competent cells for subsequent protein expression.

5'	**NdeI**															
GGC	**CAT**	**ATG**	TCT	ACC	ACC	ATC	CAA	TAC	AAC	AGC	AAC	TAT	GCG	GAC	TAC	TCT
		M	S	T	T	I	Q	Y	N	S	N	Y	A	D	Y	S
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ATC	TCC	TCG	TAC	CTG	CGT	GAA	TGG	GCA	AAC	AAC	TTC	GGC	GAT	ATT	GAT	CAG
I	S	S	Y	L	R	E	W	A	N	N	F	G	D	I	D	Q
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
GCG	CCG	GCC	GAA	ACG	AAA	GAC	CGT	GGC	TCA	TTT	TCG	GGT	AGC	TCT	ACC	CTG
A	P	A	E	T	K	D	R	G	S	F	S	G	S	S	T	L
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
TTC	AGC	GGC	ACG	CAA	TAT	GCA	ATT	GGT	AGT	TCC	CAT	TCT	AAC	CCG	GAA	GGC
F	S	G	T	Q	Y	A	I	G	S	S	H	S	N	P	E	G
50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
ATG	ATC	GCT	GAA	GGT	GAT	CTG	AAA	TAC	TCT	TTT	ATG	CCG	CAG	CAT	ACC	TTC
M	I	A	E	G	D	L	K	Y	S	F	M	P	Q	H	T	F
67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83
CAC	GGC	CAG	ATC	GAT	ACG	CTG	CAA	TTT	GGT	AAA	GAC	CTG	GCA	ACC	AAT	GCT
H	G	Q	I	D	T	L	Q	F	G	K	D	L	A	T	N	A
84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
GGC	GGT	CCG	AGT	GCA	GGC	AAA	CAC	CTG	GAA	AAA	ATT	GAT	ATC	ACG	TTT	AAC
G	G	P	S	A	G	K	H	L	E	K	I	D	I	T	F	N
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117
GAA	CTG	GAT	CTG	AGC	GGC	GAA	TTC	GAC	AGC	GGT	AAA	TCT	ATG	ACC	GAA	AAT
E	L	D	L	S	G	E	F	D	S	G	K	S	M	T	E	N
118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134
CAT	CAG	GGT	GAT	ATG	CAC	AAA	AGC	GTC	CGC	GGC	CTG	ATG	AAA	GGT	AAC	CCG
H	Q	G	D	M	H	K	S	V	R	G	L	M	K	G	N	P
135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151
GAC	CCG	ATG	CTG	GAA	GTG	ATG	AAA	GCC	AAA	GGC	ATT	AAT	GTT	GAT	ACC	GCG
D	P	M	L	E	V	M	K	A	K	G	I	N	V	D	T	A
152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168
TTC	AAA	GAC	CTG	AGT	ATC	GCC	TCC	CAA	TAT	CCG	GAT	TCA	GGT	TAC	ATG	TCG
F	K	D	L	S	I	A	S	Q	Y	P	D	S	G	Y	M	S
169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185
GAC	GCC	CCG	ATG	GTT	GAT	ACG	GTT	GGT	GTT	GTT	GAT	TGT	CAC	GAT	ATG	CTG
D	A	P	M	V	D	T	V	G	V	V	D	C	H	D	M	L
186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202
				BamHI		**3'**		^								
CTG	GCT	GCC	TAA	**GGA**	**TCC**	GGG										
L	A	A	-													
203	204	205														

Table 2.1. DNA and corresponding amino acid sequence of full-length $HasA_{yp}$. *Nde*I and *Bam*HI restriction endonuclease sites were introduced at the 5' and 3' end for sub-cloning. The carrot symbol (^) indicates the place where a stop codon was introduced to express full length minus 12 residues (FL-12) $HasA_{yp}$. Figure adapted from reference [16].

2.2.2 Protein Expression and Purification

$HasA_{yp}$ was expressed using previously described methods and was isolated using a protocol reported for the purification of apo-HasAp [8] with a few modifications: supernatant from cell lysis was loaded onto a Q-sepharose Fast Flow column (2.6 cm i.d. x 15 cm, GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 7.6) at 4 °C. The column was then washed with 2 column volumes of the same buffer and the protein eluted with the same buffer using a linear NaCl gradient (0-400 mM). Fractions containing the $HasA_{yp}$ protein, as determined by SDS-PAGE, were pooled and dialyzed against 50 mM sodium phosphate buffer (pH 7.0) containing 900 mM ammonium sulfate before loading onto a Phenyl Sepharose 6 Fast Flow (GE Healthcare) column (2.6 cm i.d. x 12 cm) pre-equilibrated with the same buffer at 25 °C. The protein was eluted in stepwise manner: (a) 50 mM sodium phosphate/750 mM ammonium sulfate (pH 7.0) was used to elute weakly bound proteins, including holo-$HasA_{yp}$ and (b) a linear gradient of sodium phosphate (50 - 20 mM)/ammonium sulfate (750 - 0 mM), pH 7.0, was used to elute apo-$HasA_{yp}$. Fractions containing apoprotein were pooled and concentrated to a volume of ~ 1 mL using 10 kDa molecular weight cut-off (MWCO) Amicon ultracentrifuge filters (Millipore, MA) before loading onto a Sephadex G-75 (2.6 cm i.d. x 90 cm, GE Healthcare) size exclusion column pre-equilibrated and eluted with sodium phosphate (μ = 0.1, pH 7.8) buffer.

Apo-$HasA_{yp}$ purified to homogeneity was analyzed by electrospray ionization mass spectrometry. The mass spectra indicated that the full-length protein is cleaved at the C-terminus during purification, with the shortest protein corresponding to full-length – 12 residues, not including the initiator Met (Figure 2.3-A). C-terminal cleavage of hemophores is common feature and may have functional relevance [12, 21]. Both HasAs and HasAp are secreted in the

extracellular milieu where it undergoes the C'-terminus proteolytic cleavage that removes the last 12-20 amino acid residues. In case of HasAs, a single C'-terminus cleavage removes the last 12 residues, whereas in HasAp, last 21 residues are cleaved [12, 22]. A stop codon was introduced after Val 193 (Table 2.1) to express truncated apo-$HasA_{yp}$ consisting of full-length – 12 residues (FL-12). The mass spectrum (Figure 2.3-B) shows that FL-12, devoid of the initial initiator Met, can be purified to homogeneity.

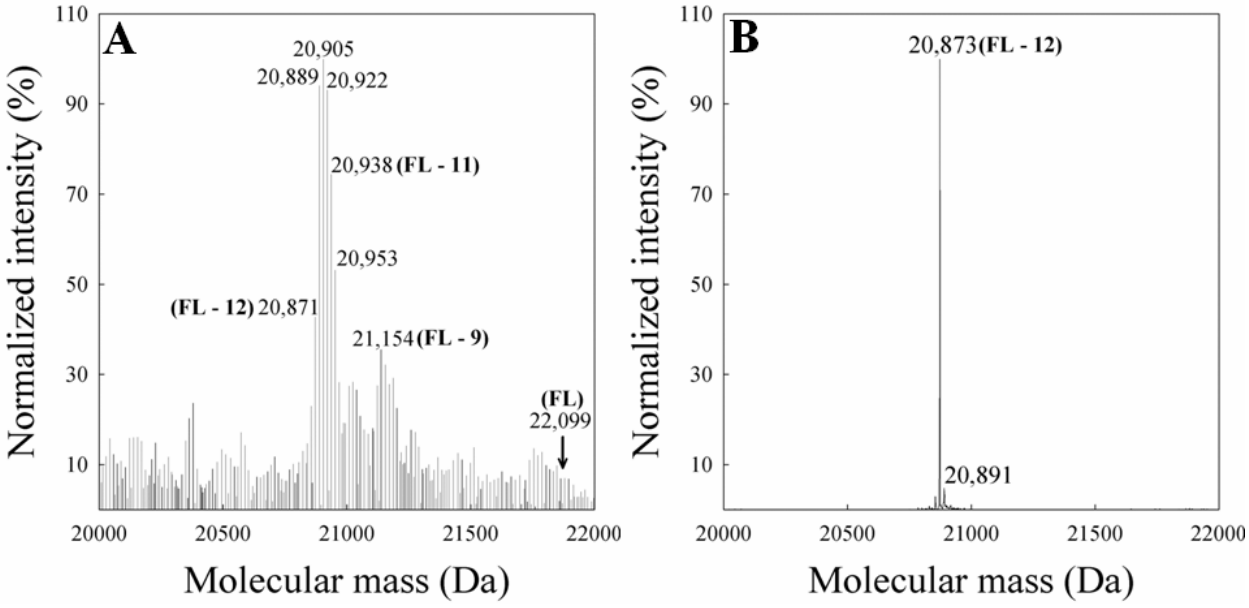


Figure 2.3. (A) Mass spectroscopic analysis of protein obtained upon expression and purification of full-length apo-$HasA_{yp}$ (expected MW = 22,099 Da) indicated that the protein is cleaved at the C-terminus, with the shortest protein corresponding to full-length – 12 (FL-12) residues, not including the initiator methionine. (B) Mass spectroscopic analysis of protein obtained upon expression and purification of FL-12 apo-$HasA_{yp}$ showed that the truncated protein devoid of the initiator methionine (expected MW = 20,874 Da) can be purified to homogeneity (experimental MW = 20,873). Figure adapted form reference [16].

To prepare holo-$HasA_{yp}$, a solution of apo-$HasA_{yp}$ (FL-12) in sodium phosphate buffer (μ = 0.1, pH 7.8) was titrated with a solution of hemin (1 mM) in dimethyl sulfoxide. The resultant solution was incubated overnight at 4 °C, concentrated to ~ 1 mL using 10 kDa MWCO Amicon ultracentrifuge filters and loaded onto a Sephadex G-75 (2.6 cm i.d., x 90 cm) size exclusion column pre-equilibrated and eluted with sodium phosphate (μ = 0.1, pH 7.8) buffer. The homogeneity of the protein was assessed by SDS-PAGE and mass spectrometry. The UV-Vis spectrum of $HasA_{yp}$ (Figure 2.4) exhibits a Soret maximum at 403 nm (ε_{403} = 116.8 mM^{-1} cm^{-1}) and bands at 498, 535 and 620 nm.

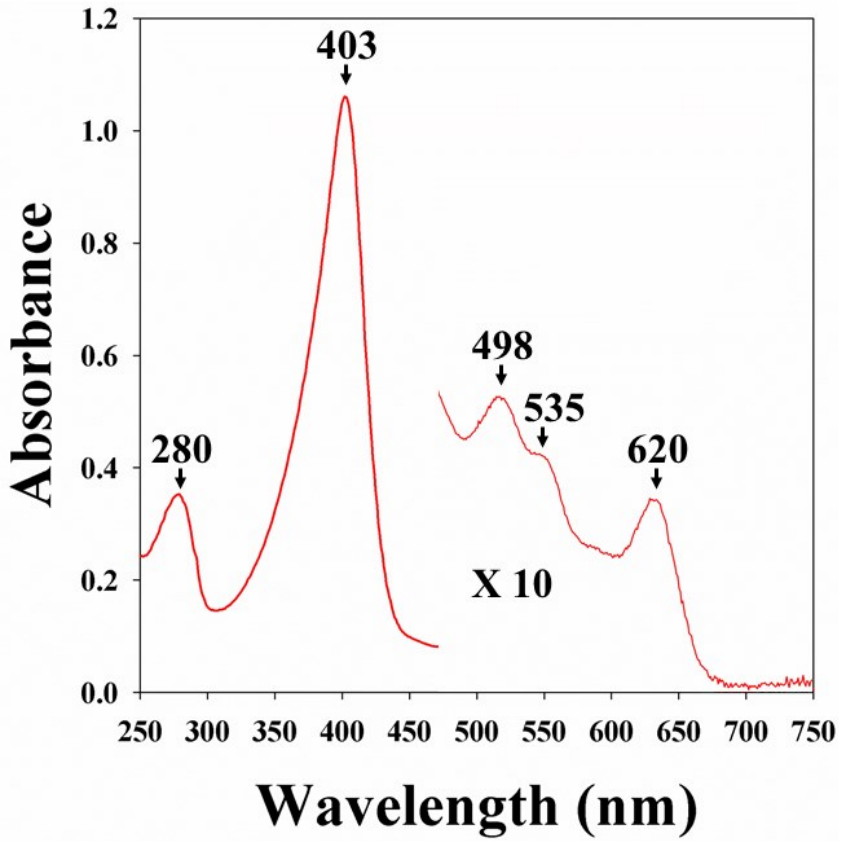


Figure 2.4. Electronic absorption spectrum of heme-reconstituted holo-$HasA_{yp}$.

2.2.3 Crystallization

Crystallization experiments were carried out with apo- and holo-$HasA_{yp}$ (FL-12) using the sitting drop vapor diffusion method and Compact Jr. Crystallization plates (Emerald BioSystems). Equal volumes (0.75 µL) of protein and crystallization solution were mixed and equilibrated against 75 µL of reservoir solution.

2.2.3.1 Apo-$HasA_{yp}$ Crystallization

Crystals showing tetragonal morphology were obtained within one week from a solution of apo-$HasA_{yp}$ (63 mg/mL) in sodium phosphate buffer (µ = 0.1, pH 7.8) and the Index HT screen (Hampton Research) (appendix I) condition H2 (0.2 M potassium sodium tartarate tetrahydrate, 20% (w/v) PEG 3350) (Figure 2.5-A). These crystals were cryo-protected by transferring to a solution containing 80% crystallization solution and 20% (v/v) PEG 400 for approximately 30 s before freezing in liquid nitrogen. Crystals with hexagonal morphology were obtained upon mixing the same protein solution with the Index HT screen (Hampton Research) condition C9 (1.1 M sodium malonate pH 7.0, 0.1 HEPES pH 7.0 and 0.5% (v/v) Jeffamine ED-2001 pH 7.0) (Figure 2.5-B). These crystals were transferred to a solution containing 80% crystallization solution and 20% (v/v) PEG 400 for approximately 30 s before freezing in liquid nitrogen.

2.2.3.2 Holo-$HasA_{yp}$ Crystallization

A cluster of thin needles grew within 3-4 days from 32 mg/mL holo-$HasA_{yp}$ in sodium phosphate buffer (µ = 0.1, pH 7.8) and the Wizard III screen (Emerald BioSystems) condition 3 (20% (w/v) PEG 3350, 200 mM magnesium formate dihydrate). These crystals diffracted to low resolution. Therefore, refinement of the crystallization conditions was carried out using the

pHaT Buffer (Emerald Biosystems) (appendix II) screen. Thick, needle like crystals were obtained within 1-2 days using pHat Buffer Screen condition C10 (0.5 M CHES/ NaOH, pH 8.8) (Figure 2.5-C). The crystals were cryo-protected by transferring to a solution containing 80% crystallization solution and 20% DMSO for approximately 30 s prior to freezing in liquid nitrogen.

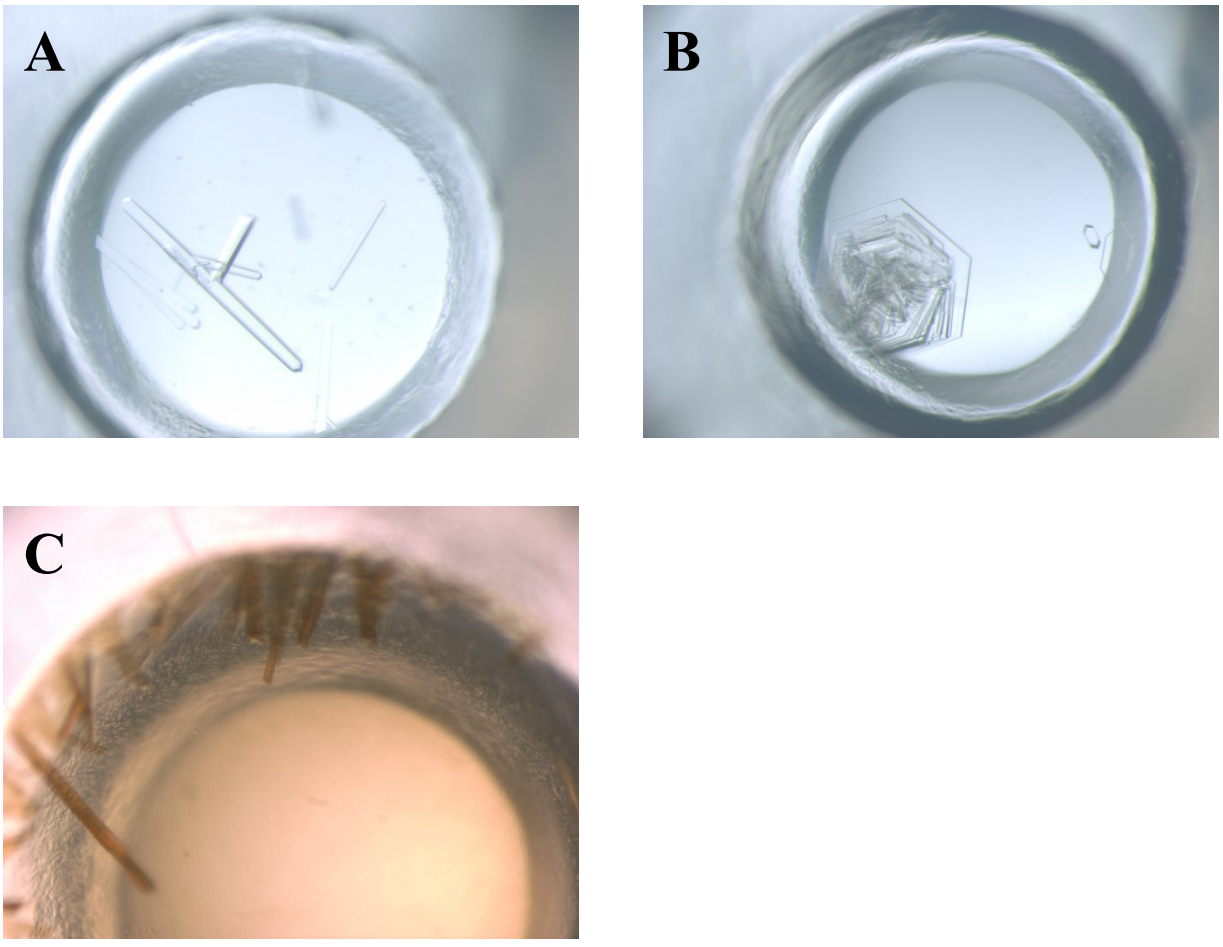


Figure 2.5. Crystal images of (A) tetragonal crystal form, (B) hexagonal crystal form of apo-$HasA_{yp}$, and (C) crystal of holo-$HasA_{yp}$ formed in a compact Jr. sitting drop vapor diffusion plates.

2.3.4 X-ray Data Collection, Structure Solution and Refinement

Data for apo-HasA$_{yp}^{tet}$, apo-HasA$_{yp}^{hex}$ and holo-HasA$_{yp}$ were collected at 100 K at the Advanced Photon Source (APS) IMCA-CAT, beamline 17-ID using a Dectris Pilatus 6M pixel array detector. The diffraction data were integrated with XDS [23] via the XDSAPP [24] interface and the Laue class and data scaling were performed with Aimless [25]. Structure solution of the apo-protein was obtained by molecular replacement with Phaser [26] via the Phenix [27] interface using the holo-HasAs structure (PDB: 1DK0) as a search model. Automated model building and structure refinement was conducted using Phenix and manual model building was carried out with Coot [28]. The final apo-HasA$_{yp}^{tet}$ structure was used as the search model for subsequent molecular replacement searches for the apo-HasA$_{yp}^{hex}$ and holo-HasA$_{yp}$ data. Crystals of apo-HasA$_{yp}^{tet}$ were indexed in a tetragonal *P* lattice with the most probable Laue class 4/*mmm* and space groups $P4_1\ 2_1\ 2$ or $P4_3\ 2_1\ 2$. Molecular replacement searches were conducted in all space groups with 422 point symmetry and the top solution was obtained in $P4_1\ 2_1\ 2$ with one molecule in the asymmetric unit. The model was refined with anisotropic atomic displacement parameters for all atoms. A sodium ion was modeled near the C-terminus which is coordinated to the backbone carbonyls of L172, A175 and F169. For apo-HasA$_{yp}^{hex}$ the crystals were indexed in a hexagonal lattice *P* with the most probable Laue class 6/*m* and space group $P6_3$. Indeed, the top molecular replacement solution was obtained in $P6_3$ although space groups with 6 point symmetry were tested. The electron density maps contained features consistent with a tris molecule (from the buffer) on a crystallographic 3-fold as well as four malonate and three polyethylene glycol (PEG) molecules (from crystallization solution). TLS refinement [29] was incorporated in the latter stages of structure refinement. Holo-HasA$_{yp}$ crystals were indexed in a *C*-centered lattice (space group *C*2). Molecular replacement was

conducted in this space group searching for 10 molecules in the asymmetric unit which were positioned by Phaser. Following subsequent rounds of refinement, large difference electron density peaks (F_o-F_c) greater than 3σ were observed on the distal sites of the heme molecules. These sites were originally assigned as water molecules, however residual positive electron density above the 3σ contour level was observed around these water molecules following refinement. Chloride ions were ultimately assigned at these sites based on the electron density, the coordination environment (positively charged Arg residues) and expected non-bonded distances to the heme Fe-atoms [30]. The average distance between the Cl^- and heme-Fe atoms is 2.82 Å. Table 2.2 summarizes the crystallographic details for the three refined structures.

Table 2.2. X-ray diffraction and refinement statistics for $HasA_{yp}$ structures

	Apo-$HasA_{yp}^{tet}$	**Apo-$HasA_{yp}^{hex}$**	**Holo-$HasA_{yp}$**
Data Collection			
Unit-cell parameters (Å, °)	*a*=70.34, *c*=74.84	*a*=75.07, *c*=129.30	*a*=139.40, *b*=69.56, *c*=232.30, *β*=100.1
Space group	$P4_12_12$	$P6_3$	$C2$
Resolution (Å)[1]	41.42-1.10 (1.12-1.10)	45.83-1.60 (1.63-1.60)	47.05-2.20 (2.24-2.20)
Wavelength (Å)	1.0000	1.0000	1.0000
Temperature (K)	100	100	100
Observed reflections	638,353	552,152	379,425
Unique reflections	76,552	54,313	109,337
$\langle I/\sigma(I)\rangle$[1]	21.8 (2.9)	18.2 (2.1)	9.9 (2.0)
Completeness (%)[1]	100 (100)	100 (100)	98.2 (97.7)

Multiplicity[1]	8.3 (8.2)	10.2 (10.3)	3.5 (3.6)
R_{merge} (%)[1, 2]	4.2 (75.5)	7.6 (1.257)	8.7 (66.8)
R_{meas} (%)[1, 4]	4.5 (80.5)	8.1 (1.323)	10.3 (78.4)
R_{pim} (%)[1, 4]	1.5 (27.6)	2.5 (41.2)	5.5 (40.8)
$CC_{1/2}$ [1, 5]	1.000 (0.802)	0.999 (0.688)	0.996 (0.686)
Refinement			
Resolution (Å)	33.04-1.10	37.54-1.60	38.62-2.20
Reflections (working/test)	72,634 / 3,844	51,507 / 2,761	103,819 / 5,497
R_{factor} / R_{free} (%)[3]	13.9 / 14.6	16.1 / 19.0	16.6 / 22.3
No. of atoms (Protein/Heme/Water)	1,458/-/220	2,754 / - / 259	13,624 / 430 / 720
Model Quality			
R.m.s deviations			
Bond lengths (Å)	0.010	0.009	0.012
Bond angles (^{o})	1.214	1.046	1.158
Average *B*-factor (Å^2)			
All Atoms	16.9	25.9	31.8
Protein	15.2	25.1	31.6
Heme	-	-	35.4
Water	28.5	32.4	33.7
Coordinate error, maximum likelihood (Å)	0.08	0.16	0.29
Ramachandran Plot			
Most favored (%)	98.45	98.35	98.08
Additionally allowed (%)	1.55	1.38	1.92

1) Values in parenthesis are for the highest resolution shell.

2) $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the ith reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.

3) $R_{factor} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$; Rfree is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.

4) R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}[25, 31]. R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}[32, 33].

5) $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data [34, 35].

2.2.5 Spectroscopic and Kinetic Analyses of $HasA_{yp}$

Kinetic measurements (EPR, Resonance Raman and Stopped-flow experiments) were done by our collaborators, Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University at Portland, Oregon, USA and details can be found in reference [16]. Results relevant to the analysis are presented in this dissertation with due acknowledgement.

2.3 Results and Discussion

2.3.1 The Q-32-bearing loop in apo-$HasA_{yp}$ is in the closed conformation.

2.3.1.1 Apo-HasAyp: Tetragonal crystal form

Crystals for apo-$HasA_{yp}$ were obtained from 70 unique conditions upon setting up for the screening. Diffraction data were collected for the tetragonal and hexagonal crystal forms as

discussed in the Experimental section. The tetragonal crystal form (apo-$HasA_{yp}^{tet}$), which has one molecule in the asymmetric unit, was refined to a resolution of 1.1 Å (Table 2.2). Electron density was observed from S2 to M184, except for T48 and L49, which were not modeled. Apo-$HasA_{yp}$ exhibits the $\alpha + \beta$ fold characteristic of HasAp and HasAs [13, 15], which is composed of a "β-sheet wall" of nine anti-parallel β-strands connected by hairpins and an "α-helix wall" composed of three α-helices and a 3_{10}-helix (Figure 2.6-A). The structure of the Y75 loop, extending from Tyr75 to Phe83, is identical to the Y75 loops in HasAp and HasAs, with Tyr75 (O_η) hydrogen bonded (2.6 Å) to His81 (N_δ).

2.3.1.2 Apo-$HasA_{yp}$: Hexagonal crystal form

The hexagonal crystal form ($HasA_{yp}^{hex}$), which exhibits two molecules in the asymmetric unit and was refined to 1.6 Å resolutions (Figure 2.6-B). Both molecules in the asymmetric unit are nearly identical (C_α-RMSD = 0.76 Å), exhibit well defined electron density for residues S2 to D180 and contain a PEG molecule between the Q32 and Y75 loops. Superposing the $HasA_{yp}^{tet}$ structure with the structures of molecules A and B of $HasA_{yp}^{hex}$ reveals near identical architectures (Figure 2.6-C), except for the apparent absence of PEG or other exogenous molecules between the Q32 and Y75 loops of apo-$HasA_{yp}^{tet}$. Residues Ser2 to Asp180 from the three apo-structures were used to calculate C_α-RMSD values using the secondary structure matching algorithm in the program Superpose [36] via the CCP4 interface [37]. Interestingly, the loop bearing Gln32, extending from Asn26 to Ser42, adopts a different conformation from that seen in apo-HasAp and apo-HasAs (Figure 2.6-D).

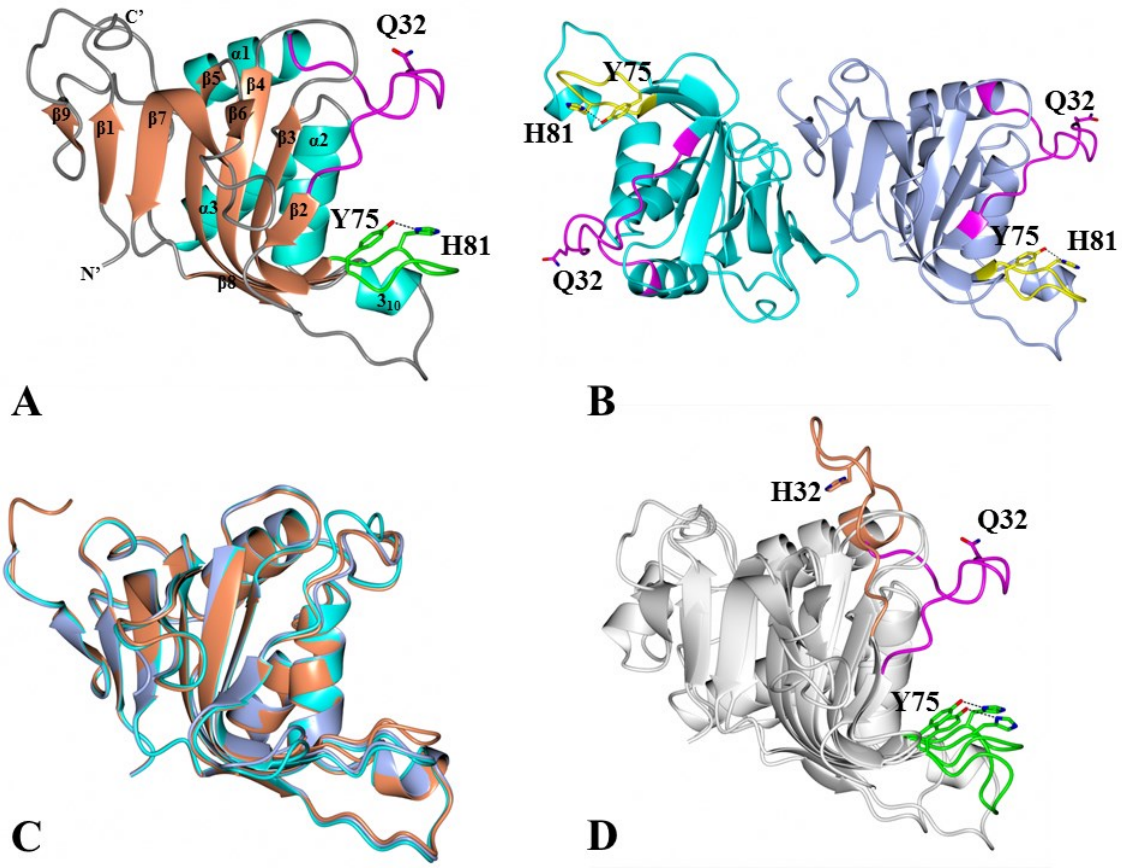


Figure 2.6. Crystal structure of (A) apo-HasA$_{yp}^{tet}$ (PDB code: 4JER) with the Q32 loop shown in magenta and the Y75 loop in green; Gln32, Tyr75 and His81 are shown in sticks. (B) apo-HasA$_{yp}^{hex}$ (PDB code: 4JES), molecule A is shown in light blue and molecule B is shown in cyan. In both the molecules the Y75 loop is in yellow with Tyr75 and His81 rendered in sticks and the Q32 loop in magenta with Gln32 in sticks. (C) Superimposed crystal structures of apo-HasA$_{yp}^{tet}$ (coral), apo-HasA$_{yp}^{hex}$, molecule A (blue) and apo-HasA$_{yp}^{hex}$, molecule B (cyan). (D) Superimposed crystal structure of apo-HasAp (PDB code: 3MOK) and apo-HasA$_{yp}^{tet}$ (PDB code: 4JER) where, the H32 loop is shown in coral in apo-HasAp and the Q32 loop in apo-HasA$_{yp}^{tet}$ in magenta. Figure adapted from reference [16].

2.3.2 The structures of apo- and holo-HasA$_{yp}$ are nearly identical.

The crystal structure of holo-HasA$_{yp}$ was solved to a resolution of 2.2 Å. Ten molecules were present in the asymmetric unit cell with a well-defined electron density traced from Ser2 to

Asp180 for each of the molecules; superposing all ten chains results in a C_α-RMSD of 0.50 Å. Figure 2.7-A illustrates the surprising observation that unlike HasAp or HasAs, the Q32- loop in $HasA_{yp}$ does not change conformation upon binding hemin (C_α-RMSD = 0.56 Å). As in the previously characterized hemophores, the hemin iron in $HasA_{yp}$ is coordinated by Tyr75, which also forms a hydrogen bond (2.7 Å) with the N_δ of His81 (His83 in HasAs and HasAp). Despite the similar coordinating environment of the proximal binding site, the environment of the distal site in $HasA_{yp}$ is distinct (Figure 2.7-B) in that the 6th coordination site is not occupied by a protein-provided ligand.

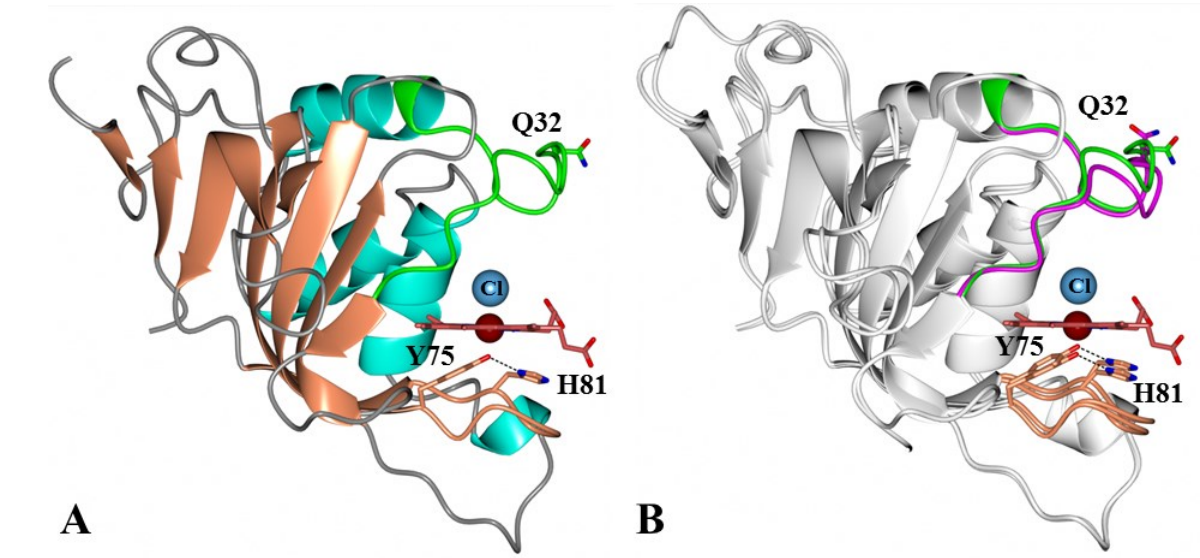


Figure 2.7. (A) Crystal structure of holo-$HasA_{yp}$ (PDB code: 4JET). Q32 loop is shown in green, the Y75 loop in coral and the chloride ion in ice-blue color. The hemin iron is coordinated by only one protein provided ligand, Tyr75, and (B) superimposition of apo-$HasA_{yp}$ (magenta) with holo-$HasA_{yp}$ (green), shows minimum structural rearrangement upon hemin binding [16].

Note that Gln32, which we hypothesized may coordinate the hemin iron, is located approximately 14 Å away (C_α) from the heme iron, in a position very similar to the one it occupies in the apo-form (Figure 2.8). Gln32 is clearly not part of the heme-binding pocket. The most significant change in the Q32 loop upon hemin binding is the relocation of the Arg40 side chain (Figure 2.9). This small reorganization protects the hemin distal face immediately after it binds to $HasA_{yp}$, a situation that contrasts with apo-HasAp or apo-HasAs, where the hemin distal face is exposed to the aqueous environment prior to closing of the H32 loop.

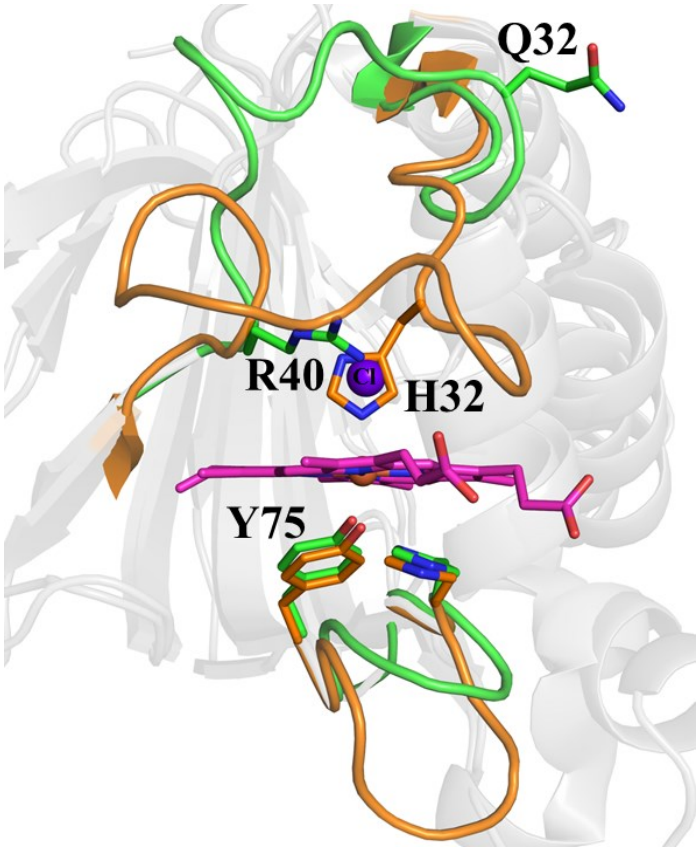


Figure 2.8. Zoomed-in view comparing the heme-binding pockets of holo-$HasA_{yp}$ (green) and holo-HasAp (orange); a chloride ion (purple sphere) in the distal pocket of $HasA_{yp}$ is in the position occupied by the side chain of H32 in holo-HasAp and holo-HasAs. Structures were superimposed using the program Superpose [36].

Electron density immediately above the hemin iron was best modeled as a chloride ion (Figure 2.9). The modeled Cl^- is not within binding distance of Fe (III) (Fe-Cl = 2.8 Å).

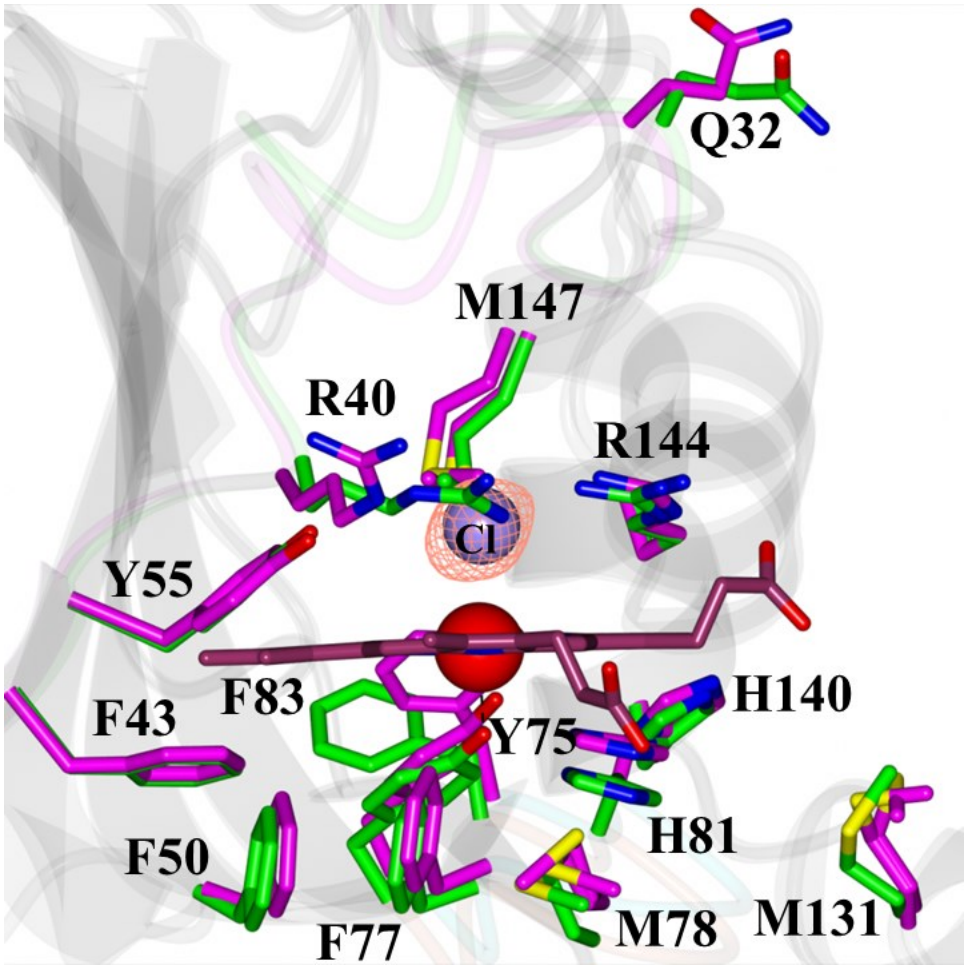


Figure 2.9. View of the heme-binding pocket in the superimposed structures of apo-$HasA_{yp}^{tet}$ and holo-$HasA_{yp}$ illustrate the minor structural differences between apo- (magenta) and holo- (green) $HasA_{yp}$. The distal site in the holo protein is coordinated by a chloride ion; the $F_o - F_c$ omit map contoured at 3σ is shown in orange mesh representation. Figure adapted from reference [16].

Kinetic measurements (resonance Raman and EPR) show that in solution the heme iron (III) in holo-$HasA_{yp}$ adopts a 6-coordinate high-spin state [16] (Figure 2.10). Although attempts

to detect Fe (III)-Cl and Fe (III)-OH stretching frequencies were unsuccessful we presume that the sixth ligand might be a solvent molecule and/or a loosely bound chloride ion. Because HasA sequences with Gln32 contain Arg40 but those with His32 do not (Figure 2.9), we hypothesize that Arg40 may stabilize the chloride ion observed in the holo-$HasA_{yp}$ structure.

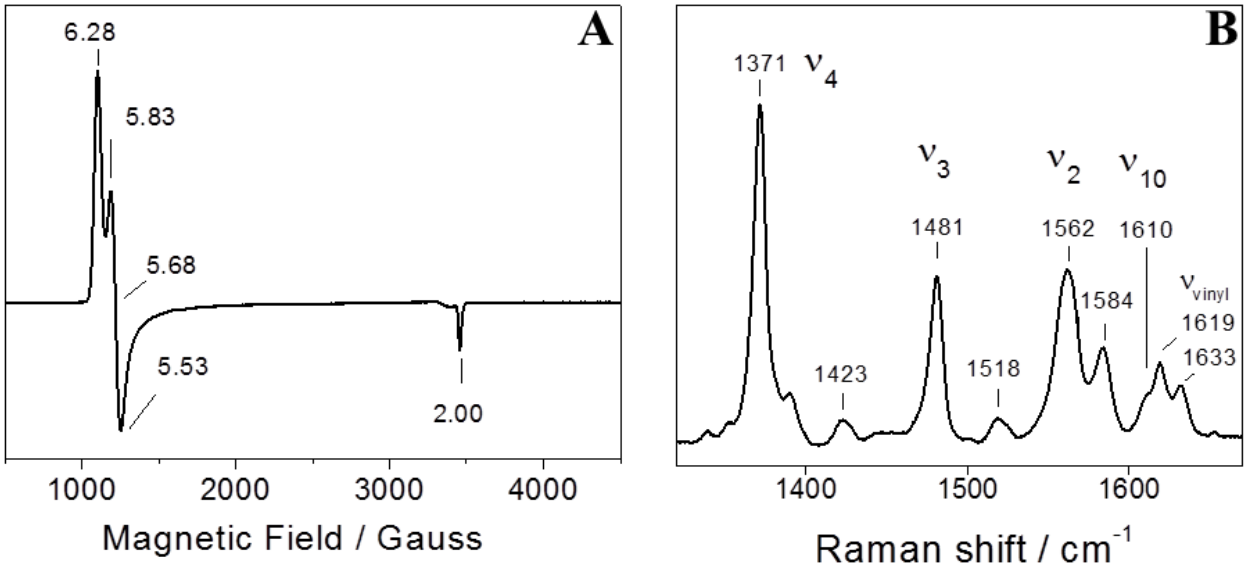


Figure 2.10. (A) X-band EPR spectrum of holo-$HasA_{yp}$ at 300 μM, *g* values of rhombic high-spin ferric heme signals are given (B) RR spectrum of holo-$HasA_{yp}$ at 300 μM, prominent porphyrin skeletal modes supporting the 6-coordinate high-spin state and the two vinyl stretching mode are labeled. Spectroscopic measurements were done by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU), Portland, Oregon, USA [16].

A closer look at the amino acid sequence alignment of the known hemophores suggests that the Q32 loop of $HasA_{yp}$ is 3-residues shorter than the H32 loop of HasAp and HasAs (Figure 2.2). This absence of three residues in the $HasA_{yp}$ sequence may decrease the conformational flexibility of the Q32 loop compared to the H32 loop in the HasAp. Structurally the three extra

residues (Ser41Asn42Thr43) in HasAp are present near the 'hinge 2' (N-terminus of β2 strand) region (Figure 2.11).

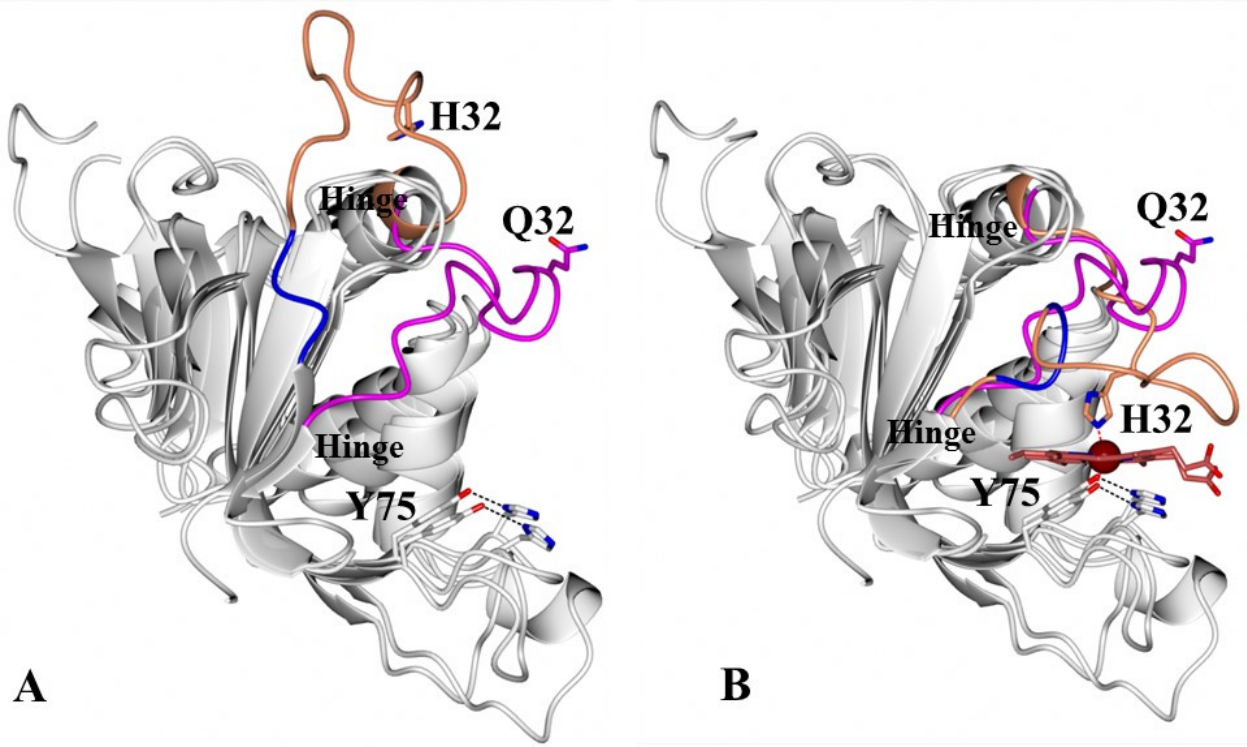


Figure 2.11. Superimposed crystal structures of: (A) apo-HasAp (PDB code: 3MOK) and apo-$HasA_{yp}^{tet}$ (PDB code: 4JER). (B) holo-HasAp (PDB code: 3ELL) and holo-$HasA_{yp}$ (PDB code: 4JET). The Q32 loop of $HasA_{yp}^{tet}$ is shown in magenta, and the H32 loop of HasAp in coral, with the three extra residues (Ser41Asn42Thr43) highlighted in blue. Figure adapted from reference [16].

In contrast to the large differences observed in the H32 loop, the structure of the Y75 loop is conserved amongst the three hemophores. As pointed out previously [8] and comparing the proximal loop residues in the structures of HasAp, $HasA_{yp}$ and HasAs, it is clear that the Y75 loop forms a conserved hydrophobic surface for hemin to interact with the hemophore (Figure 2.12). Accordingly, upon hemin binding to $HasA_{yp}$ causes minimal changes to the Y75 loop; and

the only largest change observed is a rearrangement of the Phe83 side chain to accommodate the incoming hemin (Figure 2.9).

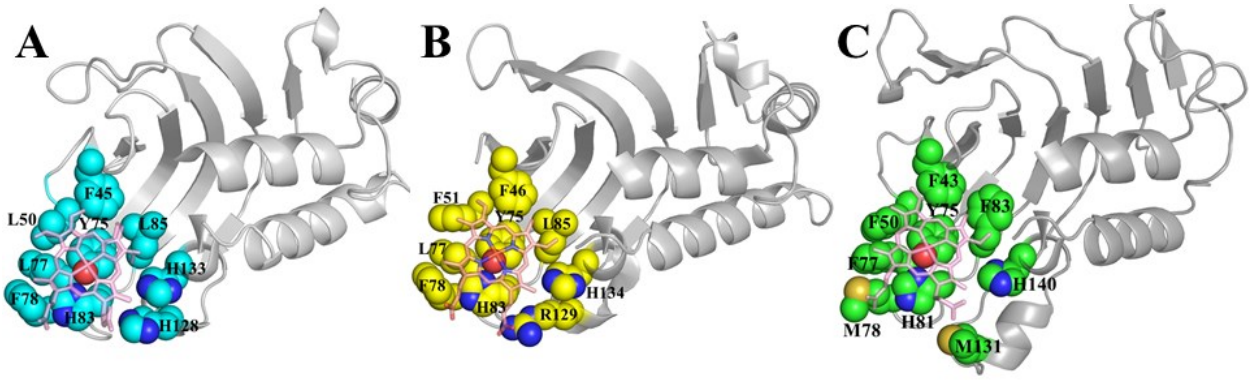


Figure 2.12. Top view of the hemin-binding pocket. Residues that form a hydrophobic (sticky) platform for macrocycle capture onto the Y75 loop and residues that interact with hemin propionates are highlighted in sphere rendering. (A) holo-HasAs (PDB code: 1B2Y) (B) holo-HasAp (PDB code: 3ELL) and holo-$HasA_{yp}$ (PDB code: 4JET). Figure adapted from reference [16].

2.3.3 $HasA_{yp}$ loads hemin from solution faster than HasAp

Kinetic measurements (Stopped-flow spectroscopy), done by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU), Portland, Oregon, USA, reveal that hemin capture by apo-$HasA_{yp}$ is complete within the millisecond dead time of the apparatus and is thus much faster than in apo-HasAp (Figure 2.13). These observations support the idea that apo-$HasA_{yp}$ adopts a closed configuration in solution and that minimal reorganization occurs upon binding hemin (Figure 2.14 A-D). This distinctive behavior of $HasA_{yp}$ is likely to also affect hemin transfer to the receptor protein, which in *S. marcescens* is thought to modulate hemin binding affinity to HasAs via interaction with the H32 loop [38].

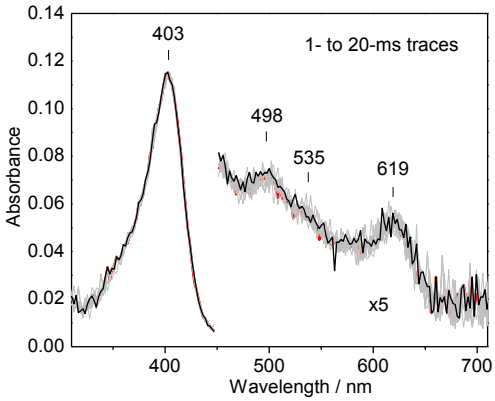


Figure 2.13. Stopped-flow absorption spectra obtained upon reacting 5 μM hemin with 30 μM apo-HasA$_{yp}$ at 4 °C (red trace, 1 ms; grey traces, 2 to 19 ms; back trace, 20 ms). The capture of hemin and formation of holo-HasA$_{yp}$ is completed within the millisecond mixing time of the instrument. Spectroscopic measurements are done by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University, Portland, Oregon, USA [16].

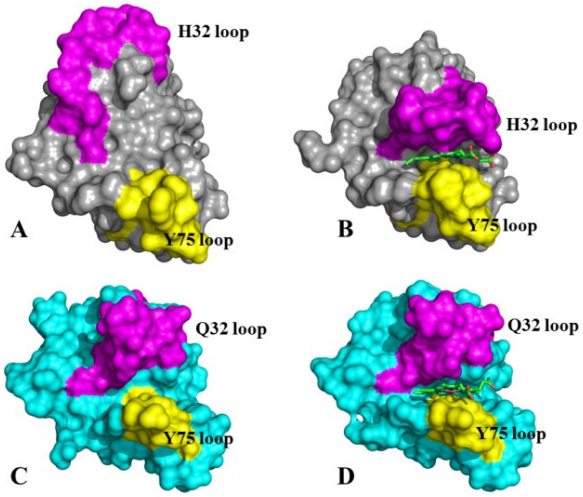


Figure 2.14. Comparison of (A) and (B) H32 loop reorganization in HasAp upon heme binding to that of (C) and (D) Q32 loop of HasA$_{yp}$. Heme is shown in sticks representation. PDB codes: (A) apo-HasAp, 3MOK, (B) holo-HasAp, 3ELL, (C) apo-HasA$_{yp}$, 4JER, and (D) holo-HasA$_{yp}$, 4JET.

2.3.4 Spontaneous cleaving of the C-terminus in FL-12

Mass spectral analysis of apo-FL-12 stored at -20 °C for several months showed that FL-12 had been cleaved predominantly at Ser183, producing a protein that corresponds to full-length minus 21 residues (FL-21). Additional investigation showed that apo-FL-12 is nearly quantitatively converted to apo-FL-21 upon incubation at 4 °C for a week. In comparison, holo-FL-12 is stable to C-terminal cleavage even if the protein is incubated at 4 °C for two weeks. These observations, which indicate that the apo-FL-12 protein has a propensity to undergo relatively slow additional cleavage of its C terminal domain to produce apo-FL-21, prompted us to carry out mass spectrometric analysis of protein in the single crystals obtained from the wells that produced the samples used for X-ray diffraction. Consistent with the observations in solution, it was found that apo-FL-12 obtained from the single crystals was cleaved at Ser185 (FL-21), whereas holo-FL-12 obtained from the corresponding single crystals remained as the FL-12 protein. C-terminal cleavage of hemophores is common [39, 40] and may have functional significance. These observations are in agreement with reports indicating that HasAp secreted to the extracellular environment undergo proteolytic cleavage of their C-terminal residues [22]. The most abundant form of HasAp secreted by *P. aeruginosa* is the truncated protein missing 15-21 C-terminal residues [22]. In contrast, the most abundant form of HasAp secrete by quorum sensing-impaired mutants of *P. aeruginosa* is full-length HasAp [22].

2.4 Conclusion

In this chapter, we characterized a novel hemophore from *Yersinia pestis*, which show completely different structural features in the distal (Q32) loop compared to the HasAp and

HasAs. The Q32 loop does not show a large conformational rearrangement. Instead it is already in a closed conformation and minimum structural changes occur upon hemin binding.

2.5 Bibliography

1. Benson DR, Rivera M. Heme Uptake and Metabolism in Bacteria. Met Ions Life Sci. 122013. p. 279-332.
2. Wandersman C, Delepelaire P. Haemophore functions revisited. *Molecular microbiology*. 2012 Aug;85(4):618-31.
3. Tullius MV, Harmston CA, Owens CP, Chim N, Morse RP, McMath LM, Iniguez A, Kimmey JM, Sawaya MR, Whitelegge JP, Horwitz MA, Goulding CW. Discovery and characterization of a unique mycobacterial heme acquisition system. *Proceedings of the National Academy of Sciences of the United States of America*. 2011 Mar 22;108(12):5051-6.
4. Smith AD, Wilks A. Extracellular heme uptake and the challenges of bacterial cell membranes. *Current topics in membranes*. 2012;69:359-92.
5. Skaar EP, Gaspar AH, Schneewind O. IsdG and IsdI, Heme-Degrading Enzymes in the Cytoplasm of *Staphylococcus aureus*. *J Biol Chem*. 2004;279:436-43.
6. Weeratunga S, Lovell S, Yao H, Battaile KP, Fischer CJ, Gee CE, Rivera M. Structural Studies of Bacterioferritin B (BfrB) from *Pseudomonas aeruginosa* Suggest a Gating Mechanism for Iron Uptake via the Ferroxidase Center. *Biochemistry*. 2010;49:1160-75.
7. Létoffé S, Ghigo JM, Wandersman C. Iron Acquisition from Heme and Hemoglobin by *Serratia marcescens* Extracellular Protein. *Proc Natl Acad Sci USA*. 1994;91:9876-80.
8. Jepkorir G, Rodriguez JC, Rui H, Im W, Lovell S, Battaile KP, Alontaga AY, Yukl ET, Moenne-Loccoz P, Rivera M. Structural, NMR spectroscopic, and computational investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa. *Journal of the American Chemical Society*. 2010 Jul 21;132(28):9857-72.

9. Cescau S, Cwerman H, Letoffe S, Delepelaire P, Wandersman C, Biville F. Heme acquisition by hemophores. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*. 2007 Jun;20(3-4):603-13.

10. Rossi MS, Fetherston JD, Letoffe S, Carniel E, Perry RD, Ghigo JM. Identification and characterization of the hemophore-dependent heme acquisition system of Yersinia pestis. *Infect Immun*. 2001 Nov;69(11):6707-17.

11. Ochsner UA, Johnson Z, Vasil ML. Genetics and regulation of two distinct haem-uptake systems, phu and has, in Pseudomonas aeruginosa. *Microbiology*. 2000 Jan;146 (Pt 1):185-98.

12. Letoffe S, Redeker V, Wandersman C. Isolation and characterization of an extracellular haem-binding protein from Pseudomonas aeruginosa that shares function and sequence similarities with the Serratia marcescens HasA haemophore. *Molecular microbiology*. 1998 Jun;28(6):1223-34.

13. Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. The crystal structure of HasA, a hemophore secreted by Serratia marcescens. *Nature structural biology*. 1999 Jun;6(6):516-20.

14. Wolff N, Izadi-Pruneyre N, Couprie J, Habeck M, Linge J, Rieping W, Wandersman C, Nilges M, Delepierre M, Lecroisey A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*. 2008 Feb 15;376(2):517-25.

15. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P, Rivera M. Structural characterization of the hemophore HasAp from Pseudomonas aeruginosa: NMR spectroscopy reveals protein-protein

interactions between Holo-HasAp and hemoglobin. *Biochemistry*. 2009 Jan 13;48(1):96-109.

16. Kumar R, Lovell S, Matsumura H, Battaile KP, Moenne-Loccoz P, Rivera M. The hemophore HasA from Yersinia pestis (HasAyp) coordinates hemin with a single residue, Tyr75, and with minimal conformational change. *Biochemistry*. 2013 Apr 23;52(16):2705-7.

17. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moenne-Loccoz P. Kinetic and spectroscopic studies of hemin acquisition in the hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 Aug 10;49(31):6646-54.

18. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moënne-Loccoz P. Kinetic and Spectroscopic Studies of Hemin Acquisition in the Hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 2010/08/10;49(31):6646-54.

19. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*. 1994 Nov 11;22(22):4673-80.

20. Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution*. 1985 Jan;2(1):13-34.

21. Letoffe S, Ghigo JM, Wandersman C. Secretion of the Serratia marcescens HasA protein by an ABC transporter. *J Bacteriol*. 1994 Sep;176(17):5372-7.

22. Arevalo-Ferro C, Hentzer M, Reil G, Gorg A, Kjelleberg S, Givskov M, Riedel K, Eberl L. Identification of quorum-sensing regulated proteins in the opportunistic pathogen

Pseudomonas aeruginosa by proteomics. *Environmental microbiology*. 2003 Dec;5(12):1350-69.

23. Kabsch W. Xds. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):125-32.

24. Krug M, Weiss MS, Heinemann U, Mueller U. XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *J Appl Cryst*. 2012;45:568-72.

25. Evans PR. An Introduction to Data Reduction: Space-Group Determination, scaling and intentisy statistics. *Acta Cryst*. 2011;D67:282-92.

26. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *Phaser* crystallographic software. *J Appl Cryst*. 2007;40:658-74.

27. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):213-21.

28. Emsley P, Lohkamp B, Scott WG, Cowan K. Features and Development of Coot. *Acta Cryst*. 2010;D66:12-21.

29. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr*. 2006 Apr;62(Pt 4):439-50.

30. Wong SG, Abdulqadir R, Le Brun NE, Moore GR, Mauk AG. Fe-Haem Bound to Escherichia coli Bacterioferritin Accelerates Iron Core Formation by an Electron Transfer Mechanism. *Biochm J*. 2012;444:553-60.

31. Evans P. Scaling and assessment of data quality. *Acta crystallographica Section D, Biological crystallography*. 2006 Jan;62(Pt 1):72-82.

32. Diederichs K, Karplus PA. Improved R-factors for Diffraction Data Analysis in Macromolecular Crystallography. *Nat Struct Biol*. 1997;4:269-75.

33. Weiss MS. Global indicators of X-ray data quality. *J Appl Cryst*. 2001;34:130-5.

34. Karplus PA, Diederichs K. Linking crystallographic model and data quality. *Science*. 2012 May 25;336(6084):1030-3.

35. Evans P. Biochemistry. Resolving some old problems in protein crystallography. *Science*. 2012 May 25;336(6084):986-7.

36. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica Section D, Biological crystallography*. 2004 Dec;60(Pt 12 Pt 1):2256-68.

37. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography*. 2011 Apr;67(Pt 4):235-42.

38. Krieg S, Huche F, Diederichs K, Izadi-Pruneyre N, Lecroisey A, Wandersman C, Delepelaire P, Welte W. Heme uptake across the outer membrane as revealed by crystal structures of the receptor-hemophore complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2009 Jan 27;106(4):1045-50.

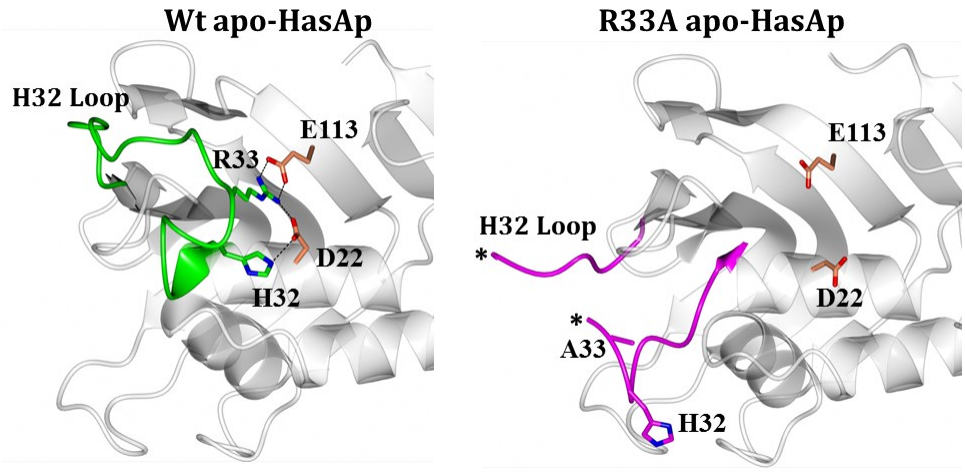
39. Létoffé S, Redeker V, Wandersman C. Isolation and Characterization of an Extracellular Haem-Binding Protein from *Pseudomonas aeruginosa* that shares function and sequence

similarities with the *Serratia Marcescens* HasA Haemophore. *Mol Microbiol*. 1998;28:1223-4.

40. Letoffe S, Ghigo JM, Wandersman C. Iron acquisition from heme and hemoglobin by a Serratia marcescens extracellular protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1994 Oct 11;91(21):9876-80.

Chapter 3

Replacing Arginine 33 for Alanine in the Hemophore HasA from *Pseudomonas aeruginosa* Causes Closure of the H32 Loop in the Apo-Protein



Summary

Under iron limiting conditions several Gram-negative bacteria secrete a hemophore that binds heme in the vertebrate host and shuttles it to specific outer membrane receptor. Structural and biochemical characterization of hemophores from *Serratia marcescens* (HasAs), *Pseudomonas aeruginosa* (HasAp) and *Yersinia pestis* ($HasA_{yp}$) are available in heme-bound (holo) and heme-free (apo) forms. Heme is bound between two extended loops, the Y75 loop harboring the Tyr75 ligand and the H32 loop harboring the His32 ligand. Heme first loads onto the Y75 loop where it is coordinated by Tyr75 which triggers a large conformational rearrangement in the H32 loop relocating His32 side chain ~ 30Å, from an open to close form thus, coordinating the heme-iron ion by His32. X-ray crystal structure and solution NMR studies have established that the H32 loop in wt and mutant apo-HasAp is quite stable and folded back over the body of the protein. Given the overall protein fold similarity in HasAp, HasAs and $HasA_{yp}$, we analyzed the stabilizing interactions in the H32 loop in wt apo-HasAp. To understand the open H32 loop structure in apo-HasAp, we replaced arginine at position 33 with an alanine (R33A) in *Pseudomonas aeruginosa*. Results obtained from solution NMR studies suggest a completely different organization of the H32 loop in R33A apo-protein. Further, X-ray crystal structures of R33A apo-HasAp reveal that replacement of arginine 33 for alanine shifts the H32 loop from an open to a closed conformation similar to that seen in wt holo-HasAp. Additional evidence from molecular dynamics simulation studies also suggest the closed loop conformation of the H32 loop in the R33A apo-protein.

3.1 Introduction

Iron and iron-binding proteins are essential components required for physiologically important functions like biosynthesis of DNA, gene regulation, electron transport chains of respiratory enzymes, and degradation of xenobiotics [1, 2]. In biological fluids the concentration of free-iron is kept at vanishingly low levels (~ 10^{-18} M) by sequestration of the micronutrient in iron-binding proteins such as transferrin and lactoferrin [3]. The low concentration of free iron in vertebrate hosts poses a challenge for pathogenic bacteria. Therefore, bacteria have developed sophisticated mechanisms to scavenge iron from its host [4].

One such mechanism employed by some Gram-negative pathogens is the secretion of small (~ 19 kDa) hemophores known as **h**eme **a**cquisition **s**ystem A (HasA) proteins into the vertebrate host upon infection. Secreted hemophore proteins capture heme from damaged or lysed heme containing proteins and deliver it to a specific outer membrane receptor (HasR) on the bacteria [4]. Extensive structural and biochemical studies of hemophore proteins from Gram-negatives such as *Serratia marcescens* (HasAs) [5, 6], *Pseudomonas aeruginosa* (HasAp) [4, 7] and *Yersinia pestis* ($HasA_{yp}$) [8] have been reported. The X-ray crystal structures of heme-bound (holo) hemophores from *S. marcescens* (HasAs) [5] and *P. aeruginosa* (HasAp) [7] revealed a unique fold, with one side of the protein packed with "α-helical wall" and another with antiparallel β-sheets [5]. The two sides of the secondary structural elements (α-helix and β-strands) are connected by two long loops. The loop which harbors the proximal ligand (Tyr75) is called the Y75 loop and the other, which harbors the distal ligand (His32), is called the H32 loop. The heme macrocycle is bound between the two loops and the heme-iron is coordinated by the axial ligands, His32 and Tyr75 [5, 7]. Structural characterization of the heme-free (apo) hemophore in solution using NMR (HasAs; [6]) and X-ray crystallography (HasAp [4]), showed

similar secondary structural features as that of the holo-counterpart, except for the H32 loop. In the apo-protein, the H32 loop is in an "open" conformation and is folded back over the body of the protein (Figure 3.1-A). Upon heme loading, the H32 loop undergoes a large conformational rearrangement which relocates the His32 ligand ~ 30 Å, from the "open" to a "closed" state, thus enabling coordination of the heme-iron by His32 from the distal site [4, 9] (Figure 3.1-B). Amino acid sequence analysis (Figure 3.2) indicates that Tyr75 is conserved among the HasA sequences but that there is no conservation of His32 [4, 8]; instead in *Yersinia* species, Gln32 is present at the equivalent position. Structural studies on apo- and holo-$HasA_{yp}$ [8] indicate that Tyr75 coordinates the heme-iron from the proximal site similar to HasAp and HasAs but there is no conformational rearrangement of the Q32 loop upon heme loading. Instead, the Q32-bearing loop is already in closed conformation and undergoes minimum structural changes upon heme binding [8].

Hemophores (HasAp, $HasA_{yp}$ and HasAs) belong to three different species of Gram-negative bacteria, yet they all show similar protein fold and function in heme uptake and delivery to the outer membrane receptor in bacteria. Although they have similar function, it is interesting that the heme coordination is different in HasAp (and HasAs), in which there is large rearrangement of the H32 loop upon heme loading, from $HasA_{yp}$, where only one ligand (Tyr75) coordinates the heme-iron ion and no rearrangements occurs in the Q32 loop [4, 6, 8]. Solution NMR and X-ray crystallography provide evidence that the H32 loop in the wt apo-HasAp is conformationally stable and do not show flexibility in solution [4, 6, 8]. The stability of the long H32 loop in the apo-protein can be visualized as a "zipper-like" network of interactions holding the loop in the open conformation (Figure 3.8-A). Combining the available structural information from hemophores (HasAp and HasAs), we asked what interactions stabilize the H32

loop. To address this question, detailed comparative analysis of interactions along the H32 loop in the wt apo-HasAp was carried out, which showed that arginine at position 33 (Arg33) is anchored to the body of the protein by a network of salt bridge and H-bond interactions with Glu113 and Asp22. Consequently we replaced Arg33 for Alanine (R33A) in HasAp. Results from solution NMR spectroscopy and X-ray crystal structures of R33A apo-HasAp indicate that the single residue replacement shifts the H32 loop in the apo-protein to a conformation very similar that seen in the H32 loop of wt holo-HasAp. Additional complementary insight into the H32 loop transition from open to closed form is stems from molecular dynamics simulations studies, which map the relocation of the H32 loop.

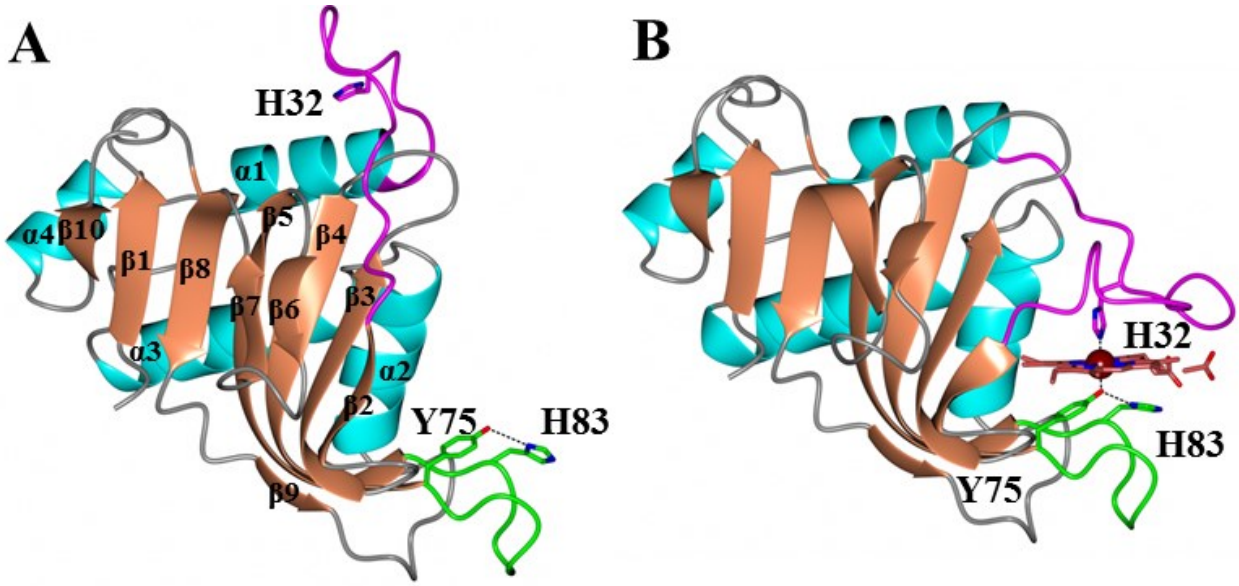


Figure 3.1. Crystal structure of (A) apo-HasAp (PDB code: 3MOK) and (B) holo-HasAp (PDB code: 3ELL) showing the Y75 loop (green) and H32 loop (magenta). Heme is shown in crimson red. Heme loads onto the Y75 loop which triggers the closure of the H32 loop, thus coordinating the heme iron by His32 [4, 7, 9].

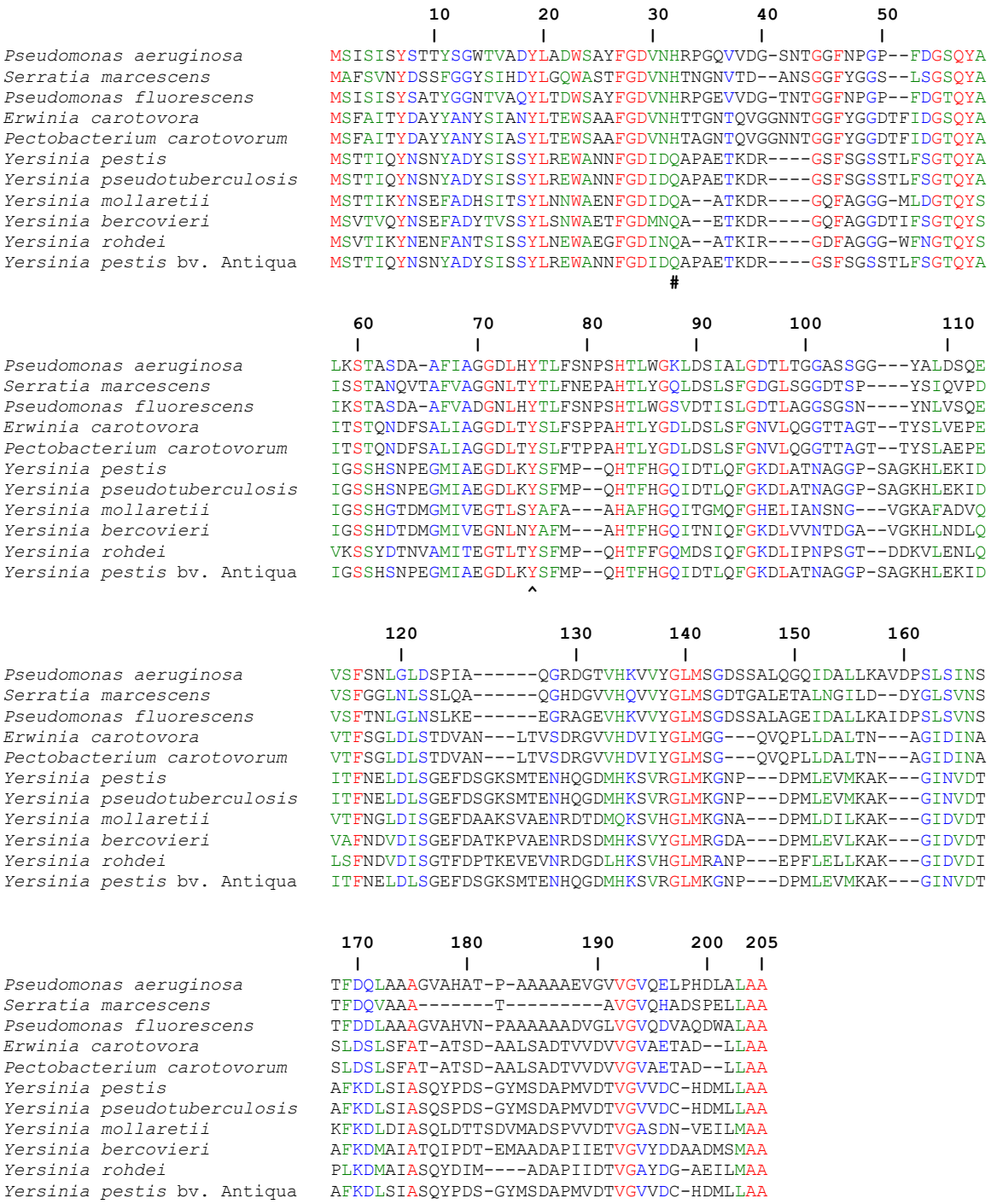


Figure 3.2. Multiple sequence alignment of full-length hemophore sequences aligned against the sequence of HasA from *Pseudomonas aeruginosa*. Residue 32 (H or Q) is highlighted by (#) and proximal loop residue Tyr75 is highlighted by (^). Conserved residues across the alignment are in red, conservative substitutions in green and semi-conservative substitutions in blue; sequence numbering is according to the *P. aeruginosa* sequence. Sequence alignment is done using CLUSTAL W [10].

3.2 Experimental Procedures

3.2.1 Cloning and Site-directed mutagenesis

Truncated form of HasAp missing the last 21 C-terminal residues was used to construct the R33A mutant [7]. The primers were synthesized by Integrated DNA Technologies, Inc. and were used with the QuickChange site-directed mutagenesis kit (Stratagene; La Jolla, CA). Primers used for the R33A mutation were 5`-TATTTTGGCGATGTGAACCATGCGCCGGGCCAGGTGGTGGATGGC-3` and 5`-GCCATCCACCACCTGGCCCGGCGCATGGTTCACATCGCCAAAATA-3`, where the underlined codons represent target substitutions. The recombinant gene harboring the mutation was transformed into XL1-Blue Competent cells (Stratagene) for amplification, and the DNA sequence was verified by ACGTInc (Wheeling, Ilinois). The recombinant DNA plasmid with the mutation was transformed into *E. coli* BL21-GOLD (DE3) competent cells for subsequent protein expression.

3.2.2 Protein Expression and Purification

Previously reported protocols [4] were used to express, purify and reconstitute with heme the R33A HasAp protein. Briefly, overnight culture of a single colony of freshly transformed cells in 50 mL of Luria-Bertani (LB) medium (100 µg/mL ampicillin) was used to inoculate 1 L of fresh M9 minimal medium [11]. The resultant culture was incubated at 37 °C with continuous shaking at 235 rpm to an OD_{600} of 0.8 – 0.9. Subsequently, temperature was lowered to 30 °C and protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The cells were cultured for 5 h and were harvested by centrifugation at 4800 rpm and stored at -20 °C. The harvested cell pellets were resuspended in

20 mM Tris.HCl (pH 7.6), 3 mL/g cell paste, and disrupted by sonication in the presence of DNase (Sigma-Aldrich, St. Louis, MO), and the cell debris were separated by centrifugation at 19, 500 rpm at 4 °C for 40 min.

3.2.3 Purification of Apo-protein

The clarified supernatant from cell lysis was loaded onto a Q-sepharose Fast Flow column (2.6 cm i.d. x 15 cm, GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 7.6) at 4 °C. The column was then washed with 2 column volumes of the same buffer and the protein eluted with the same buffer using a linear NaCl gradient (0-600 mM). Fractions containing the HasAp protein, as determined by SDS-PAGE, were pooled and dialyzed against 50 mM sodium phosphate buffer (pH 7.0) containing 0.7 M ammonium sulfate before loading onto a Phenyl Sepharose 6 Fast Flow (GE Healthcare) column (2.6 cm i.d. x 12 cm) pre-equilibrated with the same buffer at 25 °C. The protein was eluted in stepwise manner: (a) 50 mM sodium phosphate/0.5 M ammonium sulfate (pH 7.0) was used to elute weakly bound proteins, including holo-HasAp and (b) a linear gradient of sodium phosphate (50 - 20 mM) (pH 7.0)/ammonium sulfate (500 -0 mM) was used to elute apo-HasAp. Fractions containing apoprotein were pooled and concentrated to a volume of ~ 1 mL using 10 kDa molecular weight cut-off (MWCO) Amicon ultracentrifuge filters (Millipore, MA) before loading onto a Sephadex G-75 (2.6 cm i.d. x 90 cm, GE Healthcare) size exclusion column pre-equilibrated and eluted with sodium phosphate (μ=0.1, pH 7.8) buffer.

3.2.4 Preparation of R33A holo-HasAp

To prepare R33A holo-HasAp, solution of apo-protein in sodium phosphate buffer (μ = 0.1, pH 7.8) was titrated with a 1 mM hemin solution in dimethyl sulfoxide under constant

stirring. The resultant heme-reconstituted solution was incubated overnight at 4 °C, concentrated to ~ 1 mL using a 10 kDa molecular weight cut-off Amicon ultracentrifuge filter devices, and passed through a G-75 Sephadex (GE Healthcare) column (2.6 cm i.d. x 90 cm; GE Healthcare) pre-equilibrated and eluted with sodium phosphate (μ = 0.1, pH 7.8) buffer.

3.2.5 Crystallization and Data Collection

Purified apo- and holo-proteins were used for crystal screening using the sitting drop vapor diffusion method in Compact Jr. Crystallization plates (Emerald Biosciences, WA). Equal volumes (0.5 μL) of each, protein and crystallization solution were mixed and equilibrated against 100 μL of the reservoir volume.

3.2.5.1 R33A apo-HasAp

R33A apo-HasAp was concentrated to 50 mg/ml in 100 mM sodium phosphate buffer (pH 7.8), and the solution used to screen for crystal growth at 18 °C. Prism shaped (Figure 3.3-A) crystals of apo R33A-HasAp were obtained in about 43 days from the Crystal Screen (Hampton Research) condition # G10 (0.05 M $Cd(SO_4)_2$ hydrate, 0.1 M HEPES pH 7.5, 1.0 M sodium acetate trihydrate) (appendix III). Crystals were cryo-protected and transferred to a solution containing 80% crystallization solution and 20% glycerol before freezing in liquid nitrogen for data collection.

3.2.5.2 R33A holo-HasAp

Plate shaped (Figure 3.3-B) protein crystals of R33A holo-HasAp (20 mg/ml in 100 mM sodium phosphate pH 7.8) were obtained within 72 h from Emerald Biosciences Wizard III & IV condition # 32 (2100 mM DL malic acid, pH 7) (appendix IV). The crystals were then cryo-

protected by transferring to a solution containing 80% crystallization solution and 20% glycerol before freezing in liquid nitrogen for data collection.

Initial collection of diffraction data and model building were done in-house using a Rigaku RU-H3R rotating anode generator equipped with Osmic Blue focusing mirror and an R-axis IV^{2+} image plate detector. High-resolution diffraction data were acquired at 100K at the Advanced Photon Source (APS) IMCA-CAT beamline 17ID using a Dectris Pilatus 6M pixel array detector.

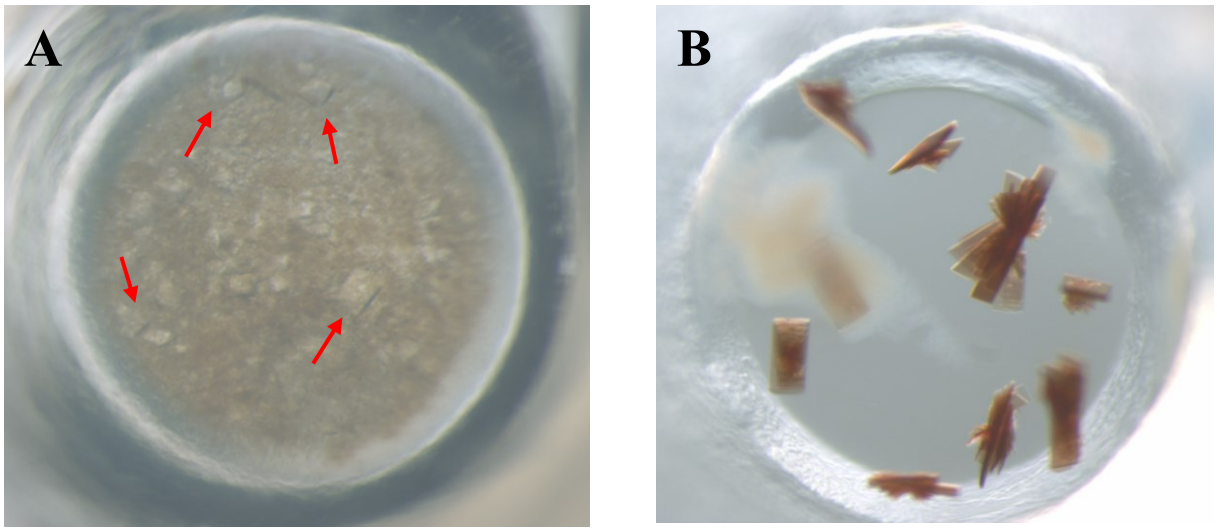


Figure 3.3. (A) Prism shaped crystals of R33A apo-HasAp obtained from Hampton Research crystal screen condition G10 (0.05 M $Cd(SO_4)_2$ hydrate, 0.1 M HEPES pH 7.5, 1.0 M sodium acetate trihydrate), red arrow points to the crystals in the well, and (B) plate shaped crystals of R33A holo-HasAp obtained from Emerald Biosciences wizard III and IV condition 32 (2100 mM DL malic acid, pH 7).

3.2.6 Structure Solution and Refinement

Diffraction intensities were integrated and scaled using the XDS [12] and Scala [13] packages, respectively and the Laue class was checked using Pointless [14]. Structure solutions for apo- and holo-proteins were obtained by molecular replacement with Phaser [15] via the Phenix [16] interface, using the coordinates from the structure of wt HasAp (PDB code: 3ELL). Structure refinement and manual model building for apo- and holo-proteins were performed with Phenix [16] and Coot [17] respectively. Structure validation was carried out using Molprobity [18] and figures were prepared using CCP4mg [19] and PyMol [20]. X-ray data collection and refinement statistics are summarized in Table 3.2.

3.2.6.1 Structural Refinement: R33A apo-HasAp

Initial structure solution and refinement were done using in-house diffraction data collected with Cu-K_α radiation. Crystals were indexed in Primitive Tetragonal Bravais lattice with Laue class of 4/*mmm* and space groups $P4_12_12$ or $P4_32_12$. All space groups with 422 point symmetry were searched during the molecular replacement. Top solution was obtained in $P4_12_12$. The Matthews – cell content analysis module [21, 22] via CCP4 [19] package calculated two molecules in the asymmetric unit. This information was used while performing molecular replacement via Phenix [16] package, which calculated one molecule in the asymmetric unit as the correct solution with ~ 70% solvent content. During the refinement, large anomalous difference electron density peaks ($F_o - F_c$) greater than 20σ were observed that were assigned as cadmium ions from the crystallization solution. Consequently, data for the structure refinement were collected at two different wavelengths at the APS synchrotron facility. The one collected at a higher wavelength (λ = 1.7396 Å, 7,150.1 eV) yielded strong anomalous electron density and another collected at lower wavelength (λ = 1.0000 Å, 12,394 eV) yielded an appreciable

anomalous signal for cadmium ions. Using the two diffraction data set, anomalous Fourier difference electron density maps were calculated using CCP4 [16] and three cadmium ions were assigned during the structure refinement (Figure 3.4). Table 3.1 lists the peak height of each of the three cadmium ions from the anomalous Fourier difference electron density maps.

Table 3.1. Anomalous peak heights for three cadmium ions in R33A apo-HasAp.

	Anomalous peak height		
Wavelength	Cadmium ion # 1	Cadmium ion # 2	Cadmium ion # 3
1.0000 Å	27.79	26.13	10.50
1.7396 Å	47.74	43.36	17.76

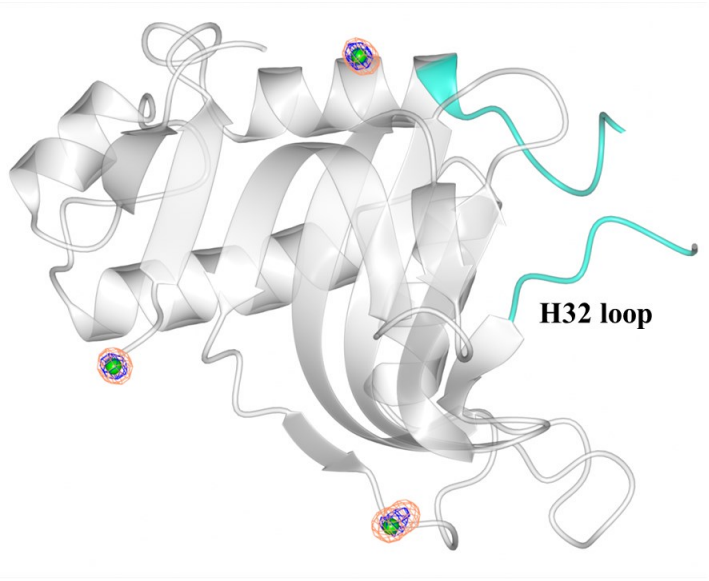


Figure 3.4. Crystal structure of R33A apo-HasAp showing the position of cadmium ions (green spheres). Anomalous difference Fourier electron density map contoured at 15σ is shown in blue (λ = 1.0000 Å) and coral (λ = 1.7396 Å). H32 loop is shown in cyan color.

3.2.6.2 Structural Refinement: R33A holo-HasAp

Crystals for R33A holo-HasAp protein were indexed in a Bravais lattice of Primitive Orthorhombic with a probable Laue class of *mmm* and space groups of $P2_12_12_1$ or $P2_12_12$. Molecular replacement searches, using Phenix, were conducted in all space groups with 222 point symmetry and the top solution was obtained in $P2_12_12_1$ with one molecule in the asymmetric unit. Refinement and subsequent refinement was conducted with anisotropic atomic displacement parameter for all atoms. Three sodium ions were modeled to the final structure. Table 3.2 summarizes the crystallographic details for the structures.

3.2.7 NMR Spectroscopy

I would like to thank and acknowledge Dr. Huili Yao, Research Associate, Rivera Lab, University of Kansas, Lawrence, KS, USA for conducting the NMR experiments for this work. Results relevant to the analysis are presented in this dissertation with due acknowledgement.

Uniformly labeled U-^{15}N-R33A HasAp were prepared with minor modification of the above described protocol: Overnight culture was resuspended in M9 media with 1.0 g of $^{15}NH_4Cl$. IPTG (final concentration of 1 mM) was added when the culture reached an OD_{600} ~ 0.9, cells were harvested and the protein was purified as described above.

3.2.8 Molecular Dynamics Simulations

I would like to thank and acknowledge Dr. Wonpil Im and Dr. Qi, Yifei, Center for Bioinformatics and Department of Molecular Biosciences, University of Kansas, Lawrence, KS, USA, for performing the molecular dynamics simulations for this work. Results relevant to the analysis are presented in this dissertation with due acknowledgement.

The simulation systems were built using the *Quick MD Simulator* module in CHARMM-GUI [23]. The protein was solvated in a 81 Å × 81 Å × 81 Å TIP3 water box and buffered with 0.15 M NaCl. For both wild type and R33A apo-protein, three independent simulation runs, each with a different initial velocity, were generated using NAMD [24]. The simulations were performed with a time step of 2 fs in NPT ensemble with the temperature maintained at 300 K using Langevin dynamics, and pressure maintained at 1.01325 bar with the Nose-Hoover piston method [25, 26]. The van der Waals interactions were switched off at 10-12 Å [27], and the electrostatic interactions were calculated using the particle-mesh Ewald method [28].

Table 3.2. X-ray diffraction and refinement statistics for R33A HasAp

	R33A apo-HasAp	**R33A holo-HasAp**
Data Collection		
Unit-cell parameters (Å, °)	*a*=92.78, *c*=95.71	*a*= 34.17, *b*= 47.17, *c*=101.37
Space group	$P4_12_12$	$P2_12_12_1$
Resolution (Å)[1]	66.62-1.94 (2.05-1.94)	47.16-1.05 (1.07-1.05)
Wavelength (Å)	1.0000	1.0000
Temperature (K)	100	100
Observed reflections	406,207	485,013

Unique reflections	31,392	75,527
$\langle I/\sigma(I)\rangle$[1]	21.1 (2.4)	15.5 (3.3)
Completeness (%)[1]	100 (99.8)	97.8 (94.4)
Multiplicity[1]	12.9 (13.1)	6.4 (6.2)
R_{merge} (%)[1, 2]	8.0 (147.6)	5.9 (51.1)
R_{meas} (%)[1, 4]	8.3 (153.6)	7.0 (61.1)
R_{pim} (%)[1, 4]	2.3 (42.0)	2.8 (24.1)
$CC_{1/2}$ [1, 5]	1.000 (0.928)	0.998 (0.861)
Refinement		
Resolution (Å)	33.31-1.94	34.53-1.05
Reflections (working/test)	31,257 / 1,578	75,450 / 3,776
R_{factor} / R_{free} (%)[3]	19.08 / 22.30	13.42 / 14.68
No. of atoms (Protein/Heme/Water)	2,488/-/59	2696 / 146 / 204
Model Quality		
R.m.s deviations		
Bond lengths (Å)	0.021	0.008
Bond angles (^{o})	1.152	1.080

Average *B*-factor (Å^2)		
All Atoms	52.99	11.33
Protein	52.95	10.65
Heme	-	10.16
Cadmium	54.13	-
Water	52.91	20.87
Coordinate error, maximum likelihood (Å)	0.22	0.06
Ramachandran Plot		
Most favored (%)	93.75	98.45
Additionally allowed (%)	6.25	1.55

1) Values in parenthesis are for the highest resolution shell.

2) $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i*th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.

3) $R_{\text{factor}} = \Sigma_{hkl} ||F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)|| / \Sigma_{hkl} |F_{\text{obs}}(hkl)|$; Rfree is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.

4) R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}. [13, 29] R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}[30, 31].

3.3 Results

3.3.1 Purification of R33A apo- and holo-HasAp

A Previously reported protocol [4], which utilizes hydrophobic affinity chromatography to separate the apo- from the holo-protein, was used to purify R33A apo-HasAp. R33A holo-protein was prepared from the purified apo-protein by heme reconstitution in vitro, as described in the experimental section. Protein homogeneity was checked using SDS PAGE which showed a single band corresponding to a molecular mass of ~ 19 kDa. The experimental molecular mass determined by electrospray ionization mass spectrometry is 18,673 Da., which is in good agreement with the molecular mass calculated from the amino acid sequence, 18,674 Da, without the initiator methionine.

3.3.2 UV-vis Spectroscopic Characterization of R33A holo-HasAp

Similar to the wt holo-HasAp [7], R33A mutants exist as 6-coordinate high spin/low spin (6c HS/LS) equilibrium mixtures. The UV-vis absorption spectrum exhibits a Soret peak at 407 nm (ε_{407} = 144.2 mM^{-1} cm^{-1}) with the visible bands at 495, 540, 577, and 616 nm. The Soret and 616 nm bands are indicative of high-spin heme iron whereas the bands at 577 and 540 nm are indicative of the low-spin species (Figure 3.5).

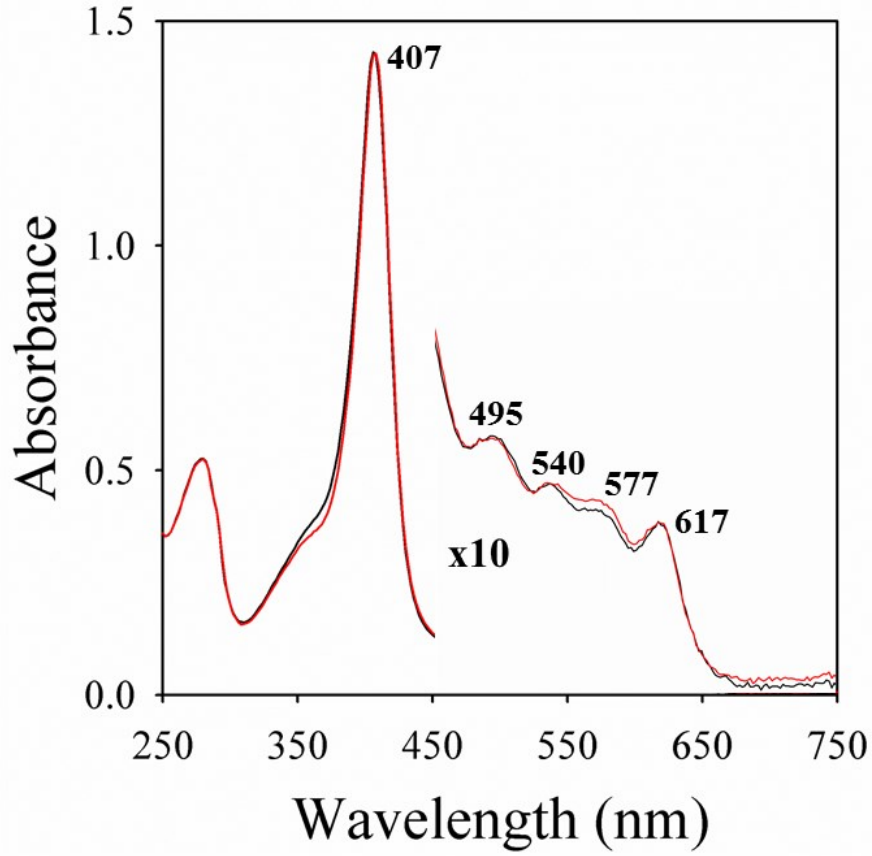


Figure 3.5. Superimposed electronic absorption spectrum of heme reconstituted wt (red) and R33A (black) holo-HasAp.

3.3.3 NMR Spectroscopic Characterization of the R33A Apo- and Holo-HasAp in Solution

Previous solution NMR studies done using mutant apo-HasAp do not show changes in the HSQC spectrum compared to the HSQC spectrum of the wt providing strong evidence for the stability of the H32 loop in the apo-protein [4, 7]. In order to investigate the topological location of the H32 loop in the R33A mutant protein, we conducted solution NMR experiments. To this end, we prepared uniformly labeled U-^{15}N-R33A apo- and holo-HasAp proteins according to the method described in the Experimental section. Figure 3.6 presents the superimposed ^{15}N-HSQC

spectrum of R33A apo-and holo-proteins with the corresponding wt proteins. To our surprise, we observed that the superimposed HSQC spectrum of the R33A apo-protein shows several cross peaks that do not match with the cross peaks in the wt apo-protein (Figure 3.6-A). This observation suggested that the H32 loop may no longer be organized as observed in the wt apo-protein. HSQC spectra taken at different temperatures (15 °C and 25 °C) also do not strongly affect or shift the cross peaks in the mutant apo-protein, suggesting that the observable resonances originate from well-structured portions of the protein. Furthermore, when we reconstitute the R33A apo-protein with hemin, the HSQC spectrum of the resultant holo-protein becomes superposable with that of holo wt HasAp, suggesting that upon heme binding the loop gets organized as observed in the wt holo-protein (Figure 3.6-B). These results strongly suggested that a portion of the apo-R33A mutant is distinct from its wt apo counterpart. We surmised that the most likely difference is in the H32 loop, and therefore we decided to investigate the mutant proteins using X-ray crystallography. The results from these experiments are discussed in the next section.

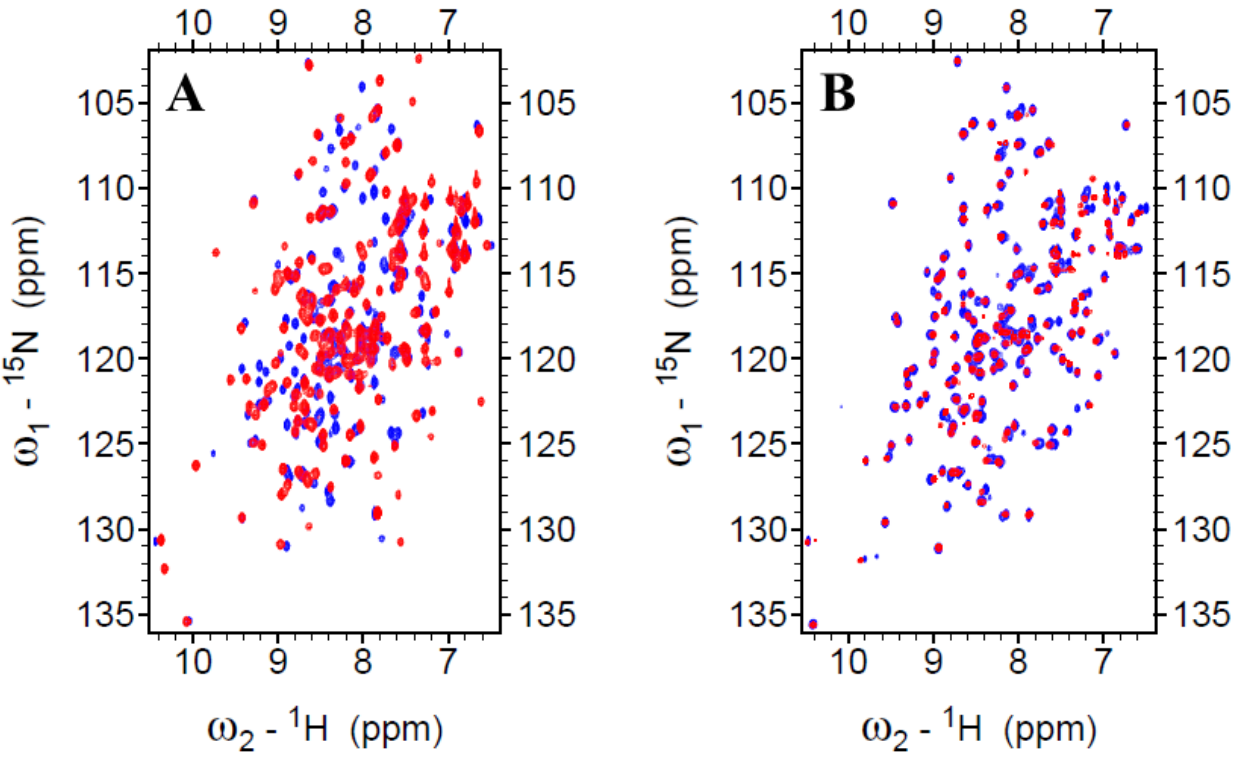


Figure 3.6. Superimposed ^{1}H, ^{15}N-HSQC spectra of (A) wt apo-HasAp (red) with R33A apo-HasAp (blue) and, (B) wt holo-HasAp (red) with R33A holo-HasAp. Spectra were taken at 32 °C using a 600MHz Bruker spectrometer. Acquisition parameters are as follows: 2048 (^{1}H) × 256 (^{15}N) complex points; 10.7 kHz (^{1}H) and 2.4 kHz (^{15}N) spectral widths; 32 scans per increment; d1=1sec. NMR spectroscopic measurements were done by Dr. Huili Yao, Research associate, Rivera Lab, University of Kansas, Lawrence, KS, USA.

3.3.4 Structural Characterization of R33A Apo-HasAp

3.3.4.1 Different Conformation of the H32 Loop in the Crystal Structure of R33A Apo-HasAp

The X-ray crystal structure of R33A apo-HasAp having one molecule in the asymmetric unit was refined to a resolution of 1.94 Å (Table 3.2). Electron density was traced from Ser2 to Ala184, except for five residues (Gly35 to Asp39) located in the His32 loop, for which no

electron density was observed. Overall, the polypeptide fold of R33A apo-HasAp is very similar to the wt (apo or holo) protein except for the H32 loop (Figure 3.7-A). The mutant apo-protein has α+β arrangement similar to that of the wt (apo- or holo) HasAp, which is composed of a "α-helical wall" on one side and the "β-sheet wall" on the other [4, 7]. The proximal, Y75 loop is similar to the wt apo-HasAp in conformation, with identical Tyr75 and His83 conformation. The most surprising structural feature of the R33A apo-HasAp structure is the H32 loop, which adopts a significantly different conformation compared to wt apo-HasAp (Figure 3.7-B). In fact, the conformation of the H32 loop is very similar to the "closed" conformation observed in the crystal structure of wt holo-HasAp (Figure 3.7-C). Due to missing electron density from Gly35 to Asp39 (marked by an asterisks, Figure 3.7-A), the His32 loop is not completely defined but the segments of the loop with clear electron density show similar conformation adopted near the C-terminus end of the α1 helix (Gly28), and the N-terminus end of the β2 strand (Gly45) in the wt holo-HasAp. Structural superimposition of R33A apo-HasAp with wt holo-HasAp reveals near identical conformation of the observable portion of the H32 loop in the R33A mutant, near the loop hinges located at Gly28 (C'-end of α1) and at Gly45 (beginning of β2) (Figure 3.7-C). Similar observations of closed conformation of the H32 loop in apo-R33A HasAp is seen when the structure is superimposed with apo- and holo-$HasA_{yp}$, in which the Q32 loop is already in the closed conformation [8] (Figure 3.7-D).

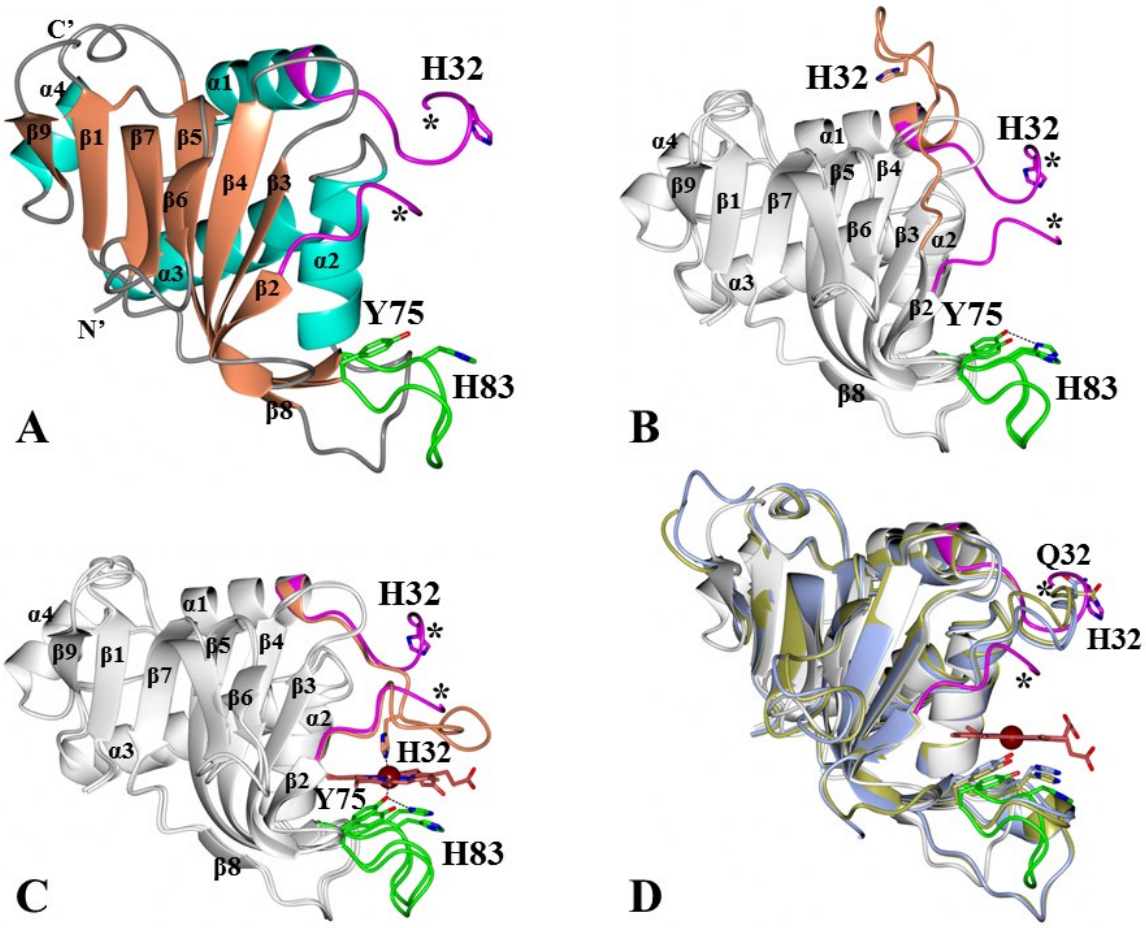


Figure 3.7. (A) Crystal structure of R33A apo-HasAp. Superimposed R33A apo-HasAp with (B) wt apo-HasAp (PDB code: 3MOK), (C) wt holo-HasAp (PDB code: 3ELL) and (D) apo-$HasA_{yp}$ (ice-blue color, PDB code: 4JER), holo-$HasA_{yp}$ (gold color, PDB code: 4JET). The H32 loop in R33A apo-HasAp is shown in magenta and wt apo- and holo-HasAp is shown in coral color. Missing electron density is shown by asterisks in R33A apo-HasAp.

In wt apo-HasAp, the H32 loop is held in an open position by various backbone and side-chain hydrogen bonding interactions, which can be thought of as a "zipper" that runs along the loop (Table 3.3 and Figure 3.8). The X-ray crystal structure as well as solution NMR studies [4, 7] suggest that the H32 loop in wt HasAp is stable and does not show dynamic behavior. It is therefore interesting that despite the zipper-like set of interactions stabilizing the loop against the

body of the protein, replacement of a single residue, Arg33 allows the loop to detach from the body of the protein and adopt the closed loop conformation. This is due to important salt bridge interactions between the Arg33 (NH1, NH2) and Glu113 (Oε1, Oε2) from β7 strand, hydrogen bond interaction between Arg33 (NH2, Nε) and Asp22 (Oδ2) from α1 helix. Moreover, the aliphatic portion of the Arg33 side chain interacts hydrophobically with the side chain of Tyr26 (Figure 3.8-A). In the wt apo-protein, Glu113 and Asp96 have their charges stabilized by salt bridge and H-bond interactions with the Arg33 and Gln36 side-chains. In the mutant apo-protein, these stabilizing interactions are lost. In the absence of Arg33, the H32 loop moves towards its closed conformation and the negative charge of Asp96 and Gln113 are stabilized by coordinating to a cation, in this case Cd ion (from the crystallization solution) (Figure 3.8-B).

Table 3.3. List of interaction in the H32 loop with the body of the protein

H32 loop residues	Distance (Å)	Interacting residues
Asp29 (N)	3.2	Tyr26 (O)
His32 (Nε)	2.8	Asp22 (Oδ)
Arg33 (NH)	2.8	Glu113(Oε)
Arg33 (NH)	3.0	Asp22 (Oδ)
Val37(O)	3.1	Asp64 (Oδ)
Asp39 (N)	2.9	Asp64 (Oδ)
Gly40 (N)	2.9	Asp64 (Oδ)
Asn42 (Nδ)	2.8	Gly100 (O)
Thr43 (O)	3.2	Lys59 (N)
Gly45 (N)	2.8	Ala57 (O)

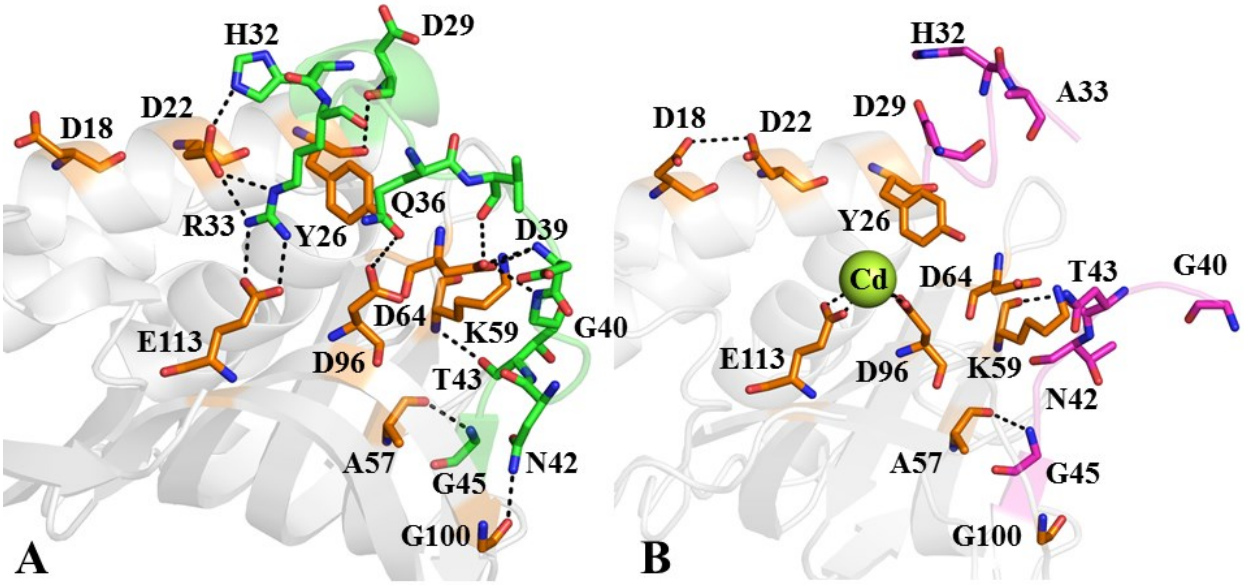


Figure 3.8. Network of hydrogen bond interaction in the H32 loop in (A) wt apo-HasAp (PDB code: 3MOK) and (B) R33A apo-HasAp. Residues from protein body are shown in orange color and H32 loop residues are colored green in wt and magenta in R33A, cadmium ion is shown as sphere (limon color).

3.3.4.2 Similar Conformation of the H32 Loop in the Crystal Structure of R33A Holo-HasAp

The crystal structure of R33A holo-HasAp having one molecule in the asymmetric unit was refined using diffraction data to 1.05 Å resolution. The overall mutant holo-protein structure is nearly identical to that of wt holo-protein, with an C_{α}-RMSD [32] of 0.22 Å. The heme iron is coordinated proximally by the Tyr75 and distally by His32 residues (Figure 3.9-A). The Y75 loop in the mutant protein adopts similar conformation as of wt holo-protein. In the H32 loop, residues Val38 to S41 adopt different conformation compared to the wt. It is important to note that in the X-ray crystal structure of R33A apo-protein these residues were disorganized and therefore not modeled due to absence of electron density. However, despite the minor

conformational changes in the H32 loop, His32 imidazole side-chain (Nδ atom) still coordinates the heme-iron ion and occupies nearly same conformation as in the wt holo-protein.

An important difference between the wt and the mutant holo-structure is the orientation of the heme molecule. In the wt holo-protein and in other mutants whose structures have been elucidated (Y75A, H83A), heme binds to the hemophore in only one orientation [4, 7, 33]. This is the first example of a mutant HasAp which binds heme in two orientations, which differ by a 180° rotation about the α-γ meso axis (Figure 3.9-B). The F_o-F_c electron density omit map contoured at 3σ cutoff shows the two orientation of the heme molecule (Figure 3.9-B).

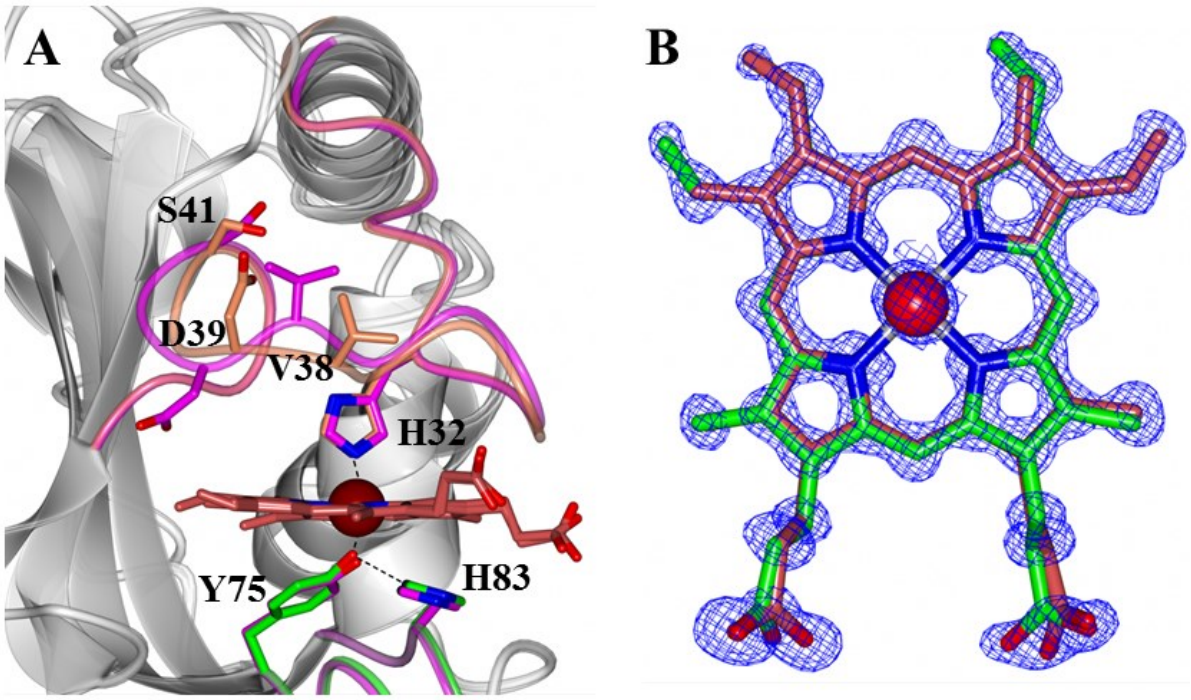


Figure 3.9. (A) Superimposed structures of wt holo-HasAp, (coral) (PDB code: 3ELL) and R33A holo-HasAp (magenta) and (B) F_o-F_c electron density omit map contoured at 3σ showing two orientation of the heme molecule (crimson red and green).

3.4 Discussion

3.4.1 Wt holo-HasAp and R33A holo-HasAp have similar UV-vis Spectra

R33A HasAp was expressed and purified to homogeneity. The R33A holo-HasAp protein was prepared from the purified apo-form by an in vitro reconstitution with hemin. The purity and the molecular weight of the proteins were assessed with the help of SDS polyacrylamide gel electrophoresis and ESI/MS. Similar to the wt holo-HasAp, UV-vis spectroscopy suggests R33A holo-HasAp exists in solution as a mixture of LS and HS species, suggesting no change in the structure of the mutant holo protein or coordination of the heme.

3.4.2 New Set of Cross Peaks Observed in the HSQC Spectrum of R33A Apo-HasAp compared to Wt Apo-HasAp

The most interesting feature of the R33A apo-protein is the conformation of the H32 loop. It is surprising that a single residue (Arg33) replacement can bring severe changes to the H32 loop in the apo-protein structure. This is also manifested in the new set of cross peaks observed in the HSQC spectrum of R33A apo-HasAp relative to the HSQC spectrum of the wt apo-protein (see Figure 3.6-A). In the wt apo-HasAp, the H32 loop is held open by various interactions along the loop that have been compared to a “zipper (see Table 3.3 and Figure 3.8). The stability of the H32 loop is evident from the solution NMR studies, where the HSQC of wt apo-HasAp as well as the mutant apo-HasAp, all show no evidence of conformational disorder in solution, even in HSQC spectra taken at different temperatures [4]. Further, electron density defining the His32 loop in the crystal structure of wt apo-HasAp is well defined and does not show indications of conformational disorder. Once the hemin is loaded onto the Y75 loop of wt HasAp, the proximal side is rapidly coordinated by the Tyr75 residue and an intramolecular

cascade of rearrangements occurs within the protein molecule which eventually breaks the various interactions in the H32 loop and enables the closure of the H32 loop, facilitating binding of His32 to the distal site of the hemin-iron.

3.4.3 Crystal Structure Shows the H32 Loop in the Mutant Apo-protein is in its Closed Conformation

Crystal structure of R33A apo- and holo-HasAp reveals an overall similar arrangement of secondary structural elements (α-helices and β-strands) of the polypeptide compared to the wt protein. The data also shows that replacing arginine at position 33 for alanine results in a structure where the H32 loop is closed, or at least in a conformation that is very similar to the conformation observed in the wt holo-HasAp (see Figure 3.7-C). Traceable electron density for the H32 loop in the crystal structure of R33A apo-protein clearly reveals the position of the hinge regions and inner portions of the loop (except for Gly35 to Asp39), which are superimposable to equivalent sections in the His32 loop of wt holo-HasAp (see Figure 3.7-B and C).

3.4.4 Molecular Dynamics Simulation Results Show Closed Conformation of H32 Loop in Mutant Apo-Protein

Molecular dynamics (MD) simulation experiments were performed to provide additional insight into the dynamic motion of the H32 loop. Figure 3.10 shows the results from the three independent MD (R33A1, R33A2 and R33A3) experiments, which corroborate the propensity of the H32 loop in R33A apo-HasAp to adopt the closed conformation. Analyzing the root mean square deviation (RMSD) as a function of time during the MD simulation of apo R33A- and apo wt apo-HasAp provides two important observations: (a) similar to solution NMR and crystal structure studies, the His32 loop in the wt apo-protein remains stable throughout the length of the

simulation, whereas (b) the loop in the apo R33A mutant starts to begin its transition at different time points, depending on the system. However, all systems converge on a closed loop at the end of the simulations (Figure 3.10-A).

Superimposing the crystal structures of wt (green) and R33A (magenta) apo-protein with R33A1 (cyan) (Figure 3.10-B) reveals very similar organization of the H32 loop in the mutant apo-protein and R33A1. Similarly, superimposition with crystal structure of the wt holo-protein (orange in Figure 3.10-C) also identifies similarity in the organization of the H32 loop of R33A1. Comparing the crystal structure of R33A apo-HasAp with the simulation structures (R33A1, R33A2 and R33A3) provides a snapshot of the H32 loop conformation (Figure 3.10 D-F). In the crystal structure of R33A apo-HasAp, no electron density could be modeled for residues Gly35 – Asp39 in the H32 loop, but from the R33A1 simulation structure those residues can be estimated and compared with the wt holo-HasAp (Figure 3.10-C). Assuming that the conformation of the H32 loop is similar to the R33A1 (Figure 3.10-C), we could speculate that the closed loop conformation would affect the heme loading in the mutant apo-protein. MD simulation results provide additional evidence that replacing Arg33 to alanine triggers a switch for the open H32 loop to closed conformation.

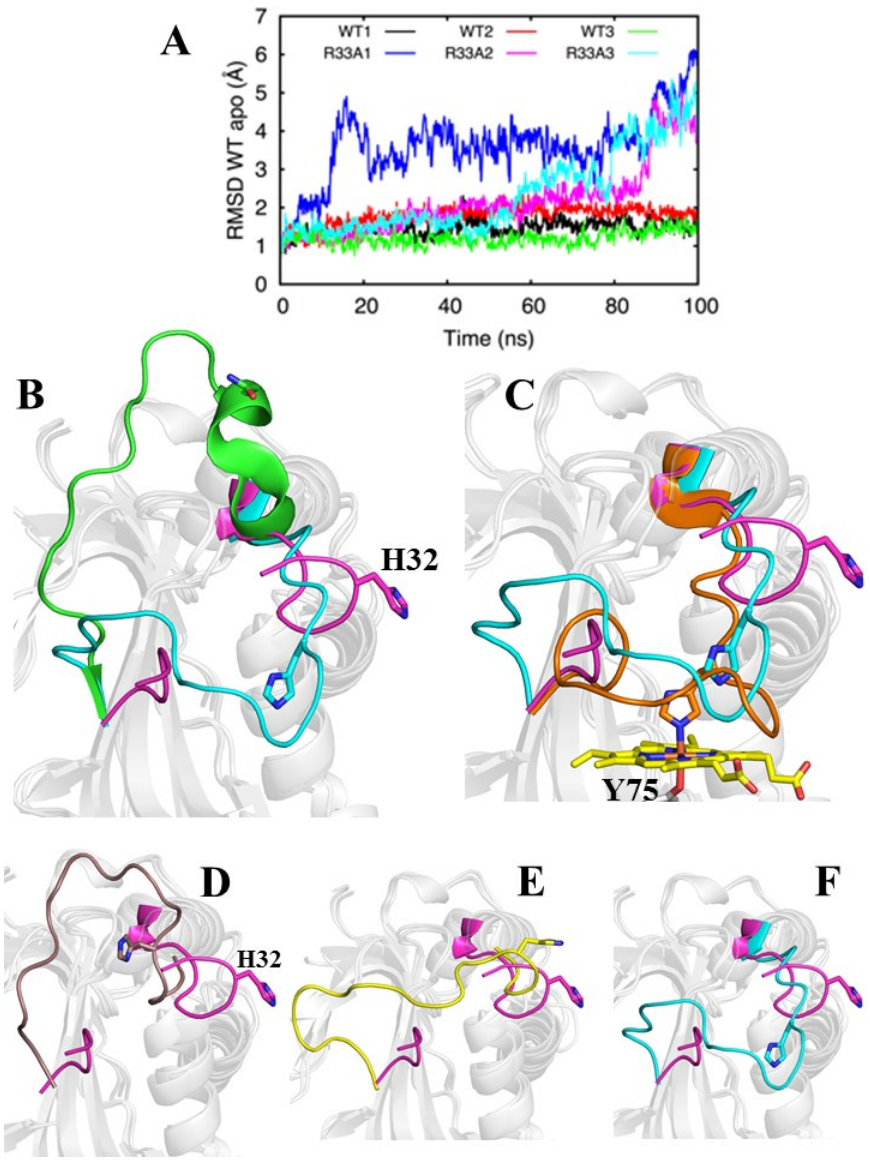


Figure 3.10. (A) RMSD plot showing the comparison of wt and R33A HasAp over the course of simulation, (B) conformation of H32 loop in superimposed crystal structure of wt apo-HasAp (PDB code: 3MOK, green), crystal structure of R33A apo-HasAp (magenta) and R33A1 structure from MD simulation (cyan), (C) crystal structure of wt holo-HasAp (PDB code: 3ELL, orange) with crystal structure of R33A apo-HasAp (magenta) and R33A1 (cyan) and superimposed crystal structure of R33A apo-HasAp with simulation structures (D) R33A3 (brown), (E) R33A2 (yellow), (F) R33A1 (cyan). Molecular dynamics simulation experiments were performed by our collaborators Dr. Wonpil Im and Dr. Qi, Yifei, Center for Bioinformatics and Department of Molecular Biosciences, University of Kansas, Lawrence, KS, USA.

3.4.5 Heme Loading Organizes the H32 Loop in R33A Similar to the Wt holo-HasAp

The X-ray crystal structure of R33A holo-HasAp is almost identical (C_α-RMSD = 0.22 Å) to the wt holo-HasAp and even the heme binding site is very similar, except for few residues in the H32 loop (see Figure 3.9-A). This is also seen in solution where HSQC NMR spectra show complete superimposition of cross peaks from the R33A holo-HasAp spectrum with the cross peaks from wt holo-HasAp (see Figure 3.6-B). It is interesting to note that heme loading onto apo H33A HasAp organizes the H32 loop in a very similar conformation as seen in the wt holo-protein. A significant difference observed between the wt and R33A holo-HasAp is the orientation of the heme. In the wt *Serratia marcescens* hemophore (HasAs), heme binds to the protein in two orientations that differ by a 180 ° rotation about the α-γ-meso axis [5] whereas in the wt *Pseudomonas aeruginosa* hemophore the heme is bound in only one orientation [4, 7, 33]. Hence, this is the first observation in a mutant HasAp protein where the heme binds in two orientations which are flipped by 180° along the α-γ meso axis (see Figure 3.9-B and Figure 3.10-C). At this point we can only speculate that the closed H32 loop conformation in the R33A apo-protein do not sample any particular conformation of the heme and once the heme is loaded, it is quickly coordinated by the His32 side-chain. Additional study may be required to conclude significance of heme binding in one orientation in the wt and other mutants of HasAp.

3.4.6 Extra Amino Acid Residues in the H32 Loop in HasAp

From the amino acid sequence alignment (Figure 3.2), we observe that *P. aeruginosa* has three additional residues in the H32 loop (^{41}SNT^{43}) compared to $HasA_{yp}$ [8] and two residues compared to HasAs [5]. Structural studies in $HasA_{yp}$ [8] suggest that the three residues less in the Q32 loop may decrease the conformational flexibility of the Q32 loop [8]. Additional study may be required to investigate in *P. aeruginosa* that the three extra residues in the H32 loop

might play a role in making specific interactions upon heme uptake and the transfer of the heme to its cognate receptor on the outer membrane of the bacteria.

3.5 Conclusion

In conclusion, this is the first attempt to understand the stabilizing forces that maintain the open H32 loop structure in HasA protein from *P. aeruginosa*. In order to gain insights about the structural mechanisms, we have explored the apo-HasAp open structure and replaced Arg33 with an Alanine. The results are intriguing as the epicenter of this severe structural change was held by a single Arg33 residue whose replacement caused the closure of the H32 loop adopting a conformation which is similar to the H32 loop in wt holo-HasAp.

3.6 Bibliography

1. Nobles CL, Maresso AW. The theft of host heme by Gram-positive pathogenic bacteria. *Metallomics : integrated biometal science*. 2011 Aug;3(8):788-96.
2. Grigg JC, Ukpabi G, Gaudin CF, Murphy ME. Structural biology of heme binding in the Staphylococcus aureus Isd system. *Journal of inorganic biochemistry*. 2010 Mar;104(3):341-8.
3. Bullen JJ, Rogers HJ, Spalding PB, Ward CG. Iron and Infection: The Heart of the Matter. *FEMS Immunol Med Microbiol*. 2005;43:325-30.
4. Jepkorir G, Rodriguez JC, Rui H, Im W, Lovell S, Battaile KP, Alontaga AY, Yukl ET, Moenne-Loccoz P, Rivera M. Structural, NMR spectroscopic, and computational investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa. *Journal of the American Chemical Society*. 2010 Jul 21;132(28):9857-72.
5. Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. The crystal structure of HasA, a hemophore secreted by Serratia marcescens. *Nature structural biology*. 1999 Jun;6(6):516-20.
6. Wolff N, Izadi-Pruneyre N, Couprie J, Habeck M, Linge J, Rieping W, Wandersman C, Nilges M, Delepierre M, Lecroisey A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*. 2008 Feb 15;376(2):517-25.
7. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P, Rivera M. Structural characterization of the hemophore HasAp from Pseudomonas aeruginosa: NMR spectroscopy reveals protein-protein

interactions between Holo-HasAp and hemoglobin. *Biochemistry*. 2009 Jan 13;48(1):96-109.

8. Kumar R, Lovell S, Matsumura H, Battaile KP, Moenne-Loccoz P, Rivera M. The hemophore HasA from Yersinia pestis (HasAyp) coordinates hemin with a single residue, Tyr75, and with minimal conformational change. *Biochemistry*. 2013 Apr 23;52(16):2705-7.
9. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moenne-Loccoz P. Kinetic and spectroscopic studies of hemin acquisition in the hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 Aug 10;49(31):6646-54.
10. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*. 2013 Jul;41(Web Server issue):W597-600.
11. Rodriguez-Maranon MJ, Qiu F, Stark RE, White SP, Zhang X, Foundling SI, Rodriguez V, Schilling CL, 3rd, Bunce RA, Rivera M. 13C NMR spectroscopic and X-ray crystallographic study of the role played by mitochondrial cytochrome b5 heme propionates in the electrostatic binding to cytochrome c. *Biochemistry*. 1996 Dec 17;35(50):16378-90.
12. Kabsch W. Xds. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):125-32.
13. Ahn HC, Le YTH, Nagchowdhuri PS, Derose EF, Putnam-Evans C, London RE, Markley JL. NMR Characterization of an Amyloidogenic Conformational Ensemble of the PI3K SH3 Domain. *Protein Sci*. 2006;15:2552-7.

14. Kabsch W. Automatic-Indexing of Rotation Diffraction Patterns. *J Appl Crystallogr*. 1988 Feb 1;21:67-71.

15. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *Phaser* crystallographic software. *J Appl Cryst*. 2007;40:658-74.

16. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):213-21.

17. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography*. 2004 Dec;60(Pt 12 Pt 1):2126-32.

18. Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. 2003 Feb 15;50(3):437-50.

19. Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr*. 2004 Dec;60(Pt 12 Pt 1):2288-94.

20. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

21. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci*. 2003 Sep;12(9):1865-71.

22. Matthews BW. Solvent Content of Protein Crystals. *J Mol Biol*. 1968;33:491-7.

23. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*. 2008 Aug;29(11):1859-65.

24. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *Journal of computational chemistry*. 2005 Dec;26(16):1781-802.

25. Feller SE, Zhang Y, Pastor PR, Brooks BR. Constant pressure molecular dynamics simulation: The Langevin piston method *J Chem Phys*. 1995;103:4613.

26. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys*. 1994;101.

27. Steinbach PJ, Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of computational chemistry*. 1994;15(7):667-83.

28. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method *J Chem Phys*. 1995;103:8577.

29. Evans PR. An Introduction to Data Reduction: Space-Group Determination, scaling and intentisy statistics. *Acta Cryst*. 2011;D67:282-92.

30. Diederichs K, Karplus PA. Improved R-factors for Diffraction Data Analysis in Macromolecular Crystallography. *Nat Struct Biol*. 1997;4:269-75.

31. Weiss MS. Global indicators of X-ray data quality. *J Appl Cryst*. 2001;34:130-5.

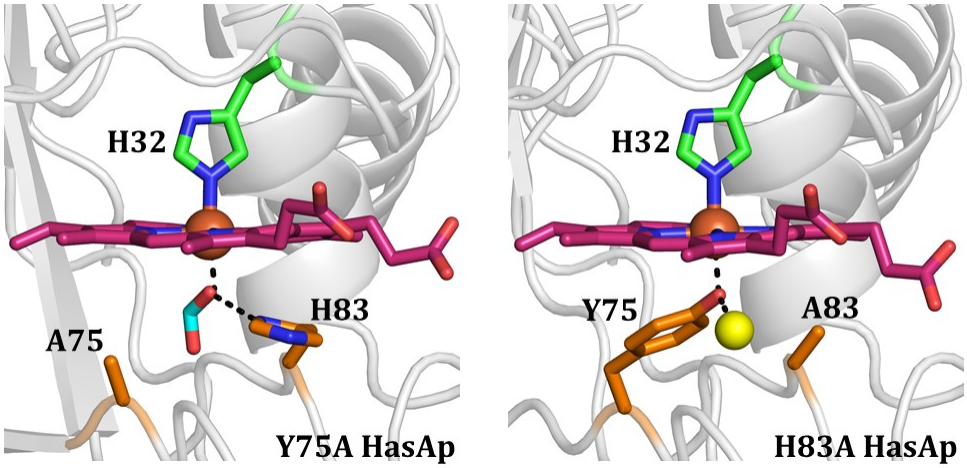
32. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica Section D, Biological crystallography*. 2004 Dec;60(Pt 12 Pt 1):2256-68.

33. Kumar R, Matsumura H, Lovell SW, Yao H, Rodriguez JC, Battaile KP, Moenne-Loccoz P, Rivera M. Replacing the Axial Ligand Tyrosine 75 or its Hydrogen Bond Partner

Histidine 83 Minimally Affects Hemin Acquisition by the Hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2014, 52, 2112-25.

Chapter 4

Replacing the Axial Ligand Tyrosine 75 or its Hydrogen Bond Partner Histidine 83 minimally affects hemin acquisition by the Hemophore HasAp from *Pseudomonas aeruginosa*



Summary

Hemophores from *Pseudomonas aeruginosa* (HasAp), *Serratia marcescens* ($HasA_{sm}$), and *Yersinia pestis* ($HasA_{yp}$) bind hemin between two loops. One of the loops harbors conserved axial ligand Tyr75 (Y75 loop) in all three structures, whereas the second loop (H32 loop) contains axial ligand His32 in HasAp and $HasA_{sm}$, but a non-coordinating Gln32 in $HasA_{yp}$. Hemin binding to the Y75 loop of HasAp or $HasA_{sm}$ causes a large rearrangement of the H32 loop that enables His32 coordination. The Q32 loop in apo-$HasA_{yp}$ is already in the closed conformation, such that hemin binding to the conserved Y75 loop occurs with minimal structural rearrangement and without coordinative interaction with the Q32 loop. In this study, structural and spectroscopic investigations of the hemophore HasAp were conducted to probe (*i*) the role of the conserved Tyr75 loop in hemin binding, and (*ii*) the proposed requirement of the His83-Tyr75 hydrogen bond to enable hemin coordination by Tyr75. High-resolution crystal structures of H83A holo-HasAp obtained at pH 6.5 (0.89 Å) and 5.4 (1.25 Å) show that Tyr75 remains coordinated to the heme iron, and that a water molecule can substitute for His83 N_δ to interact with the O_η atom of Tyr75, likely stabilizing the Tyr75-Fe interaction. NMR spectroscopy revealed that in apo-Y75A and H83A HasAp, the Y75 loop is disordered, and that disorder propagates to nearby elements of secondary structure, suggesting that the His83 N_δ to Tyr75 O_η interaction is important to the organization of the Y75 loop in apo-HasA. Kinetic analysis of hemin loading carried out with stopped flow UV-vis and rapid-freeze-quench resonance Raman show that both mutants load hemin with biphasic kinetic parameters that are not significantly dissimilar from those previously observed with wild type HasAp. When the structural and kinetic data are taken together a tentative model emerges, which suggests that HasA hemophores utilize hydrophobic, π-π stacking, and van der Waals interactions to load hemin efficiently, while

axial ligation likely functions to slow hemin release, thus enabling the hemophore to meet the challenge of capturing hemin under inhospitable conditions and deliver it selectively to its cognate receptor.

4.1 Introduction

Iron-containing proteins are integral components of fundamental physiological processes such as DNA synthesis, gene regulation, respiration, oxygen transport, and degradation of xenobiotics [1, 2]. The relative availability of iron in the host environment is an important determinant of virulence, which is why iron bioavailability is associated with sepsis [3] and an innate response of the immune system is to deploy iron-binding proteins such as lactoferrin and transferrin to lower free-iron to vanishingly small concentrations ($\approx 10^{-18}$ M) in plasma [4]. In spite of the iron-limiting conditions imposed by the host, pathogenic bacteria can establish infections, which is a manifestation of the molecular mechanisms evolved by pathogens to procure iron from their host. Approximately 95% of all iron in mammals is found in hemoglobin heme [5]. Consequently, hemoglobin and other host proteins that scavenge hemoglobin or heme from blood plasma, such as hemopexin, haptoglobin and serum albumin, are a source of iron for pathogenic bacteria [6]. Indeed, pathogenic bacteria are able to overcome the extremely low levels of free iron within their mammalian hosts by deploying several iron and heme-iron acquisition systems [7-9]. Some Gram negative bacteria deploy *has* (heme acquisition system), which is characterized by the secretion of the hemophore HasA to bind and sequester extracellular heme and deliver it to a specific outer membrane receptor, HasR for subsequent internalization [7, 10, 11].

HasA hemophores from *Serratia marcescens* ($HasA_{sm}$) [12, 13], *Pseudomonas aeruginosa* (HasAp) [14, 15] and *Yersinia pestis* ($HasA_{yp}$) [16] have been structurally characterized in their heme-free (apo) and heme-bound (holo) forms. The structures of HasAp, $HasA_{yp}$ and $HasA_{sm}$ are composed of an "α-helix wall" packed against a "β-sheet wall" and two extended loops which bridge the α-helix and β-sheet walls. The heme iron of HasAp and $HasA_{sm}$ is coordinated

by Tyr75 and His32, with each axial ligand harbored in distinct loops, termed the Y75 and H32 loops (Figure 4.1-A). This combination of axial ligands has been observed only in hemoglobin M Saskatoon [17], hemoglobins from invertebrates [18, 19], and heme d_1 in cytochrome cd_1 reductase [20], whereas examples of heme proteins that utilize Tyr as the only protein provided ligand, which include hemoglobins M [17, 21] methemalbumin [22], catalase [23], $HasA_{yp}$ [16], Isd hemophores secreted by Gram positive organisms [2], and heme-binding periplasmic-binding proteins [24, 25], are relatively more abundant. The main difference between apo- and holo-HasAp (or $HasA_{sm}$) is a large reorganization of the H32 loop, which relocates the His32 side chain approximately 30 Å (Figure 4.1-B). The role played by His32 in heme binding has been investigated using the H32A mutant of HasAp [14, 26]. The X-ray crystal structure of the holo form revealed a non-crystallographic dimer linked by cofacial interaction of two heme molecules, which are coordinated by Tyr75; each subunit is nearly identical in structure to the apo-protein, including the open conformation of the H32 loop. These observations suggested that heme loads onto the Y75 loop while the H32 loop is in the open (apo-like) conformation. Subsequent closure of the H32 loop in wild type (wt) HasAp excludes water from the heme environment and provides the sixth ligand to the heme [14]. This idea found additional support in the structure of monomeric holo-H32A in complex with imidazole, which revealed the heme-iron coordinated by Tyr75 and an exogenous imidazole. While the H32 loop is not completely defined in the imidazole complex, segments with clearly discernible electron density indicate a closed loop conformation similar to that seen in wt holo-HasAp. Moreover, the plane of the coordinated imidazole in H32A holo-HasAp is nearly superimposable with the His32 imidazole plane in wt HasAp. Solution NMR studies of H32A holo-HasAp and its imidazole complex

provided corroborating evidence that loading of hemin on the Y75 loop triggers closure of the H32 loop, even in the absence of the His32 side chain [14].

Kinetic studies of heme binding to wt apo-HasAp using rapid-mixing techniques coupled to electronic absorption, EPR and resonance Raman spectroscopic methods identified two distinct kinetic phases [26]. The spectroscopic changes in the first phase, which is complete within 20 ms, are indicative of efficient heme loading onto the Y75 loop, while the slower phase, which lasts ≈1 s at 4 °C, reflects closure of the H32 loop and coordination of the heme-iron by His32; accordingly, the H32A variant exhibits only the first millisecond phase corresponding to the loading of hemin onto the Y75 loop [26]. Taken together, our structural and kinetic investigations clearly demonstrate that hemin loads onto the Y75 loop within a few ms, where it is coordinated by Y75, and triggers closure of the H32 loop and coordination of the heme within the second timescale. A recent study carried out with the H32A mutant of $HasA_{sm}$ arrived to a similar model for heme loading [27].

Amino acid sequence alignments of HasA proteins from Yersinia species show conservation of Tyr75 and His83 in the Y75 loop, but no conservation for His32 [16]. Structural characterization of HasA from *Yersinia pestis*, apo- and holo-$HasA_{yp}$, where residue 32 is a glutamine, led to three important observations: (*i*) the Gln32 bearing loop adopts a closed conformation in apo-$HasA_{yp}$ (Figure 4.1-C), (*ii*) Tyr75 is the only endogenous heme ligand in holo-$HasA_{yp}$, and (*iii*) heme binding to apo-$HasA_{yp}$ occurs on a sub-millisecond timescale with minimal structure reorganization (Figure 4.1-D) [16]. This study of $HasA_{yp}$ highlights variations in the coordination state of hemin in HasA hemophores, but it also underscores a strong structural conservation of the proximal loop, which includes the invariant Tyr75 ligand, and its hydrogen bond partner His81 (Figure 4.1). The function of this hydrogen bond, which is thought

to increase the phenolate character of Tyr75, has been proposed to be essential for proximal hemin coordination by Tyr75 [15, 28, 29]:

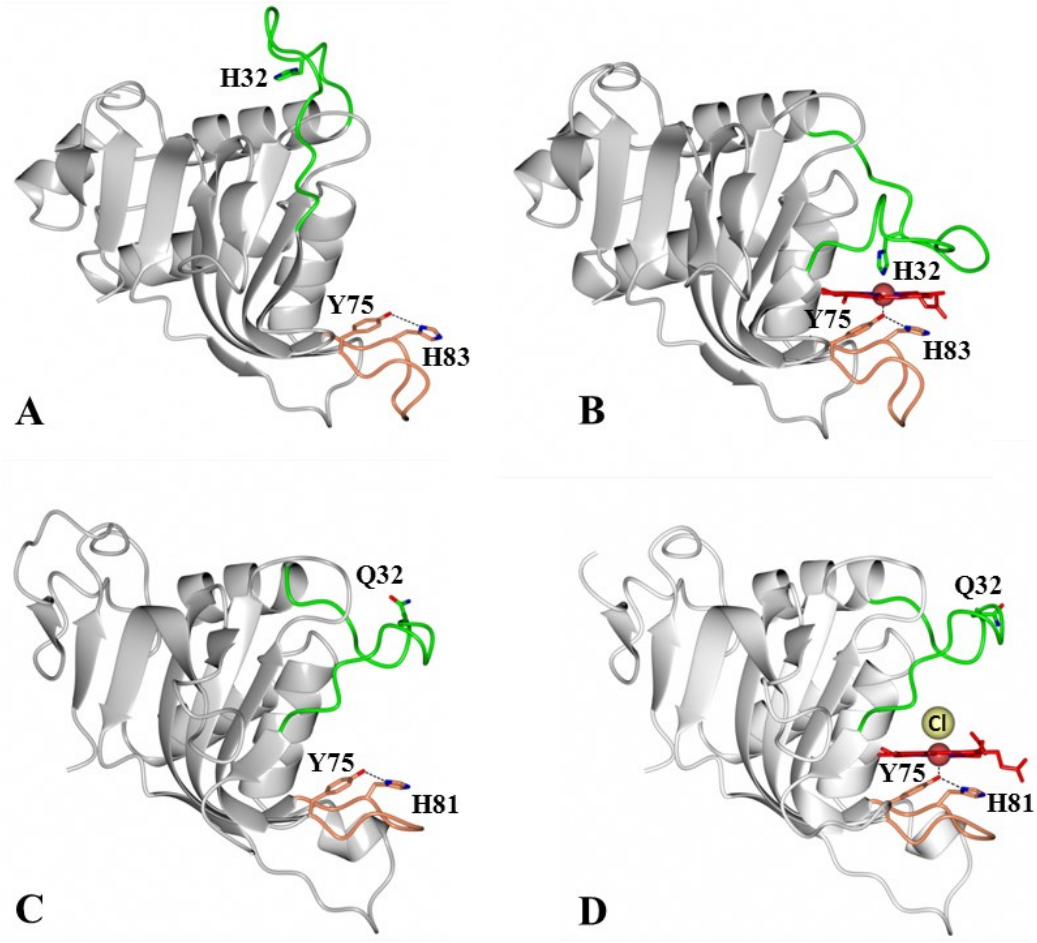


Figure 4.1. HasA hemophores display structural conservation of the proximal loop (coral) but divergence in the structure and function of the distal loop (green). Structures of (A) apo-HasAp (PDB code: 3MOK) and (B) holo-HasAp (PDB code: 3ELL) illustrate the large conformational rearrangement of the distal loop (green) caused by heme binding, which results in heme axial ligation by His32 and Tyr75. In contrast, the structures of (C) apo-$HasA_{yp}$ (PDB code: 4JER), and (D) holo-$HasA_{yp}$ (PDB code: 4JET) illustrate the minimal reorganization of the distal loop upon heme binding and the coordination of heme by only one protein-provided ligand, Tyr75. The Cl^- ion that coordinates to the distal site is shown as a yellow sphere. Figure adapted from reference [30].

A low-resolution crystal structure of H83A holo-$HasA_{sm}$ has been reported, but poor electron density does not establish whether Tyr75 remains coordinated to the heme iron in H83A holo-$HasA_{sm}$ [31]. Resonance Raman and NMR spectroscopic studies, however, led the same authors to suggest that in the absence of His83 the Tyr75-Fe bond is severed. Thus the H83A mutant of $HasA_{sm}$ was described as a 5-coordinate species, with His32 as the only axial ligand at pH ≈ 5.2, His32 and water as axial ligands at neutral pH, and His32 and a hydroxide ligand at basic pH [31]. A different study proposed that in the absence of Tyr75, His83 likely acts as a heme iron ligand [32, 33], but direct experimental support for this notion is lacking.

Given the structural conservation of the Y75 loop in HasA hemophores, and the proposed requirement for the His83-Tyr75 unit for coordination of hemin by Tyr75, we carried out a structural and spectroscopic characterization of the Y75A and H83A HasAp mutants. Our high-resolution crystal structures of H83A-HasAp clearly demonstrate that in the absence of His83, the Tyr75 ligand remains coordinated to the heme-iron at neutral pH and at pH 5.4. These structures also show that in the absence of the His83 side chain, a water molecule can substitute for the missing N_δ of His83 to interact with the Tyr75 O_η atom and aid its coordination to the ferric heme iron. NMR spectroscopic studies conducted with Y75A and H83A apo-HasAp demonstrate that disruption of the Tyr75 to His83 hydrogen bond induces significant conformational disorder in the Y75 loop, which propagates into neighboring elements of secondary structure. Stopped-flow UV-vis absorption and rapid freeze quench resonance Raman (RFQ-RR) analyses of hemin binding to Y75A and H83A apo-HasAp show biphasic kinetics comparable to those observed in wt HasAp. These results suggest that the importance of hemin iron axial ligation in the process of hemin acquisition by HasAp should not be overestimated and

point toward non-covalent interactions as factors dominating heme capture and the formation of holo-HasAp.

4.2 Experimental Procedures

I would like to thank and acknowledge Dr. Grace Jepkorir, Rivera Lab alumni, who started this work, prepared the site-directed mutagenesis and set up for protein crystallization. Also, I would like to thank and acknowledge Dr. Scott Lovell, Director, Protein Structure Lab, Shankel Structure Biology Center, University of Kansas, KS, for data collection and solving the structures of the mutants Y75A and H83A HasAp.

4.2.1 Site-directed Mutagenesis

Y75A- and H83A-HasAp mutants were constructed in the background of the gene coding for the truncated form of HasAp missing the last 21 C'-terminal residues [15]. Primers were synthesized by Integrated DNA Technologies, Inc. and were used without further purification in conjunction with the QuickChange site-directed mutagenesis kit (Stratagene; La Jolla, CA). Primers used to install the Y75A mutation are 5`-GGCGGCGATCTGCATGCGACCCTGTTT-3` and 5`GCTAAACAGGGTCGCATGCAGAT-CGCCGCC-3` and those use for the H83A mutation are 5`-CTGTTTAGCAACCC-GAGCGCGACCCTGTGGGGC-3` and 5`-GCCCCACAGGGTCGCGCTCGGGTTGCTAAA-CAG-3`, where the underlined codons represent target substitutions. Each gene construct was transformed into XL1-Blue Competent cells (Stratagene) for amplification, and DNA sequence verified by SeqWright (Huston, Texas). Recombinant DNA plasmid with the correct sequence was transformed into *E. coli* BL21-GOLD (DE3) competent cells for subsequent protein expression.

4.2.2 Protein Expression and Purification.

Expression, purification and reconstitution of Y75A- and H83A-HasAp with heme were carried out using previously described methods [14]. The oligomeric state of the mutant hemophores was determined with the aid of a gel filtration column (Superdex 75 prep grade, GE Healthcare), calibrated with cytochrome *c* (12.4 kDa), chymotrypsin (25 kDa), ovalbumin (44 kDa), and albumin (67 kDa) standards, in 50 mM Tris-HCl (pH 7.5), 100 mM KCl buffer. The preparation of wt, Y75A- and H83A-HasAp uniformly labeled with nitrogen-15 (U-^{15}N-HasAp, U-^{15}N-Y75A- and U-^{15}N-H83A-HasAp) or labeled with both nitrogen-15 and carbon-13 ([U-^{13}C, U-^{15}N]-HasAp-WT, ([U-^{13}C, U-^{15}N]-HasAp-Y75A and ([U-^{13}C, U-^{15}N]-HasAp-H83A) was performed as previously described for wt and H32A-HasAp [14].

4.2.3 Crystallization and Data Collection.

HasAp mutants in 100 mM sodium phosphate pH 7.8 (Y75A) or 100 mM Tris pH 7.6 (H83A) were concentrated to 15 mg/mL and screened for crystal growth at 20 °C using Compact Jr. (Emerald biosystems) sitting drop vapor diffusion plates. Equal volumes of protein and crystallization solution were equilibrated against 100 µL of reservoir volume. Crystals of Y75A HasAp were obtained within 24 h from Crystal Screen 2 (Hampton Research) (appendix V) condition #32 (1.6 M $(NH_4)_2SO_4$, 100 mM HEPES pH 7.5, 100 mM NaCl). Crystals were transferred to a solution containing 80% crystallization solution and 20% glycerol before freezing in liquid nitrogen for data collection. Crystals of H83A-HasAp were obtained from three distinct conditions: (1) Wizard II (Emerald biosystems) (appendix VIII) condition #11 (10% isopropanol, 100 mM sodium cacodylate pH 6.5, 200 mM zinc acetate) produced non-diffracting small crystals that grew as clusters of small needles in approximately 24 h at 20 °C. Refinement of the crystallization conditions was conducted using the Hampton Additive screen

(Hampton Research) (appendix VII). Plate-shaped clusters that diffracted to high resolution were obtained from 10% isopropanol, 100 mM sodium cacodylate pH 6.5, 200 mM zinc acetate, 3% dioxane. These crystals were transferred to a solution containing 75% crystallization solution and 25% ethylene glycol before freezing in liquid nitrogen for data collection. (2) Crystals were obtained in approximately 3 days at 20 °C from Hampton Crystal Screen 2 condition #23 (1.6 M $(NH_4)_2SO_4$, 100 mM MES pH 6.5, 10% dioxane). Crystals were transferred to a solution containing 80% crystallization solution and 20% ethylene glycol before freezing in liquid nitrogen for data collection. (3) Crystals were obtained in approximately 3 days at 20 °C from 1.6 M $(NH_4)_2SO_4$, 100 mM MES pH 5.4, 10% dioxane. Samples were transferred to a solution containing 80% crystallization solution and 20% ethylene glycol before freezing in liquid nitrogen for data collection. High-resolution diffraction data from all crystals were acquired at the Advanced Photon Source IMCA-CAT beamline 17ID using a Dectris Pilatus 6M pixel array detector.

4.2.4 Structure Solution and Refinement

Diffraction intensities were integrated and scaled using the XDS [34] and Scala [35] packages, respectively and the Laue class was checked using Pointless [36]. Coordinates from the structure of wt HasAp (PDB: 3ELL) were used for molecular replacement with Molrep [37] searching for a single molecule in the asymmetric unit for all data sets. Structure refinement of positional and anisotropic atomic displacement parameters and manual model building for both mutants were performed with Phenix [38] and Coot [39] respectively. Structure validation was carried out using Molprobity [40] and figures were prepared using CCP4mg [41] and PyMol (www.pymol.org). Data collection and refinement statistics are summarized in Table 4.1.

4.2.4.1 Y75A Holo-HasAp

Following initial refinement, prominent difference (F_o-F_c) electron density was observed near the heme on the side that would normally be occupied by Tyr 75. Three peaks were present in this region that were at an angle of approximately 120^o and separated by approximately 1.3 Å. Refinement with an ethanol molecule at this site resulted in positive F_o-F_c electron density at the C2 atom of ethanol indicating an underestimation of electrons, whereas placement of a formate ion did not yield any residual electron density. Similar electron density was observed from four diffraction data sets ranging from 0.95 Å to 1.30 Å resolution obtained from four unique crystallization conditions: Hampton Crystal Screen 2 condition #23 described above, Hampton Crystal Screen 2 condition #32 (1.6 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 0.1 M sodium chloride), Hampton Crystal Screen 2 condition #42 (1.5 M Ammonium sulfate, 0.1 M Tris pH 8.5, 12% v/v Glycerol), and Hampton Crystal Screen 1 condition #38 (1.4 M Sodium citrate tribasic dehydrate, 0.1 M HEPES pH 7.5). Herein we only report the 0.95 Å resolution data obtained with Crystal screen 2, condition #23. Since the electron density was not consistent with any of the crystallization components, protein storage buffer or purification buffers, it is likely that the formate ion was present in the heme used in the reconstitution of the apo-protein.

4.2.4.2 H83A Holo-HasAp

Condition 1 produced a monoclinic crystal form, henceforth referred to as H83A^{mono}. Molecular replacement searches in the space groups $P2_1$ and $P2$ produced correlation coefficients of 0.62 and 0.30 respectively. Therefore, subsequent structure refinement was conducted in the space group $P2_1$. Initial structure solution and refinement were conducted using in-house diffraction data collected with Cu-K_α radiation. Large peaks greater than 10σ were observed in the difference Fourier maps (F_o-F_c) that were tentatively assigned as zinc ions from

the crystallization solution. No anomalous signal was observed at these sites using Cu-K_α data. An X-ray fluorescence scan indicated that zinc was indeed present in the crystal. Consequently, data for structure refinement were collected at a wavelength of 1.0000 Å which still yields an appreciable anomalous signal for Zn ~ 2.6e^-. In addition, diffraction data were collected at a low energy remote wavelength (λ =1.28414 Å, 9,655 eV), which resulted in the disappearance of the anomalous difference density at the Zn sites but an increase at the heme-Fe site. Condition 2 produced an orthorhombic crystal form ($H83A^{ortho}$) and condition 3 produced crystals at pH 5.4 ($H83A^{pH5.4}$), which were obtained from attempts to grow crystals in a pH range of 5.0 to 5.5, in order to explore the ligation state of Tyr75 in acidic conditions. The structure was solved as described above for ($H83A^{ortho}$).

4.2.5 NMR Spectroscopy.

Solution NMR spectroscopy experiments were performed by Dr. Juan C. Rodríguez and Dr. Huili Yao, Rivera Lab, University of Kansas, Lawrence, KS, USA and the experimental details can be found in reference [30]. Results relevant to the analysis are presented in this dissertation with due acknowledgement.

4.2.6 Kinetics Experiments

Kinetic measurements (EPR, Resonance Raman, Stopped-flow experiments and Rapid-Freeze-Quench) were performed by our collaborators, Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU) at Portland, Oregon, USA and experimental details can be found in reference [30] Results relevant to the analysis are presented in this dissertation with due acknowledgement.

Table 4.1. X-ray data and refinement statistics for Y75A and H83A HasAp

	Y75A-HasA$_p$	**H83A^{mono}-HasA$_p$**	**H83A^{ortho}-HasA$_p$**	**H83A$^{pH5.4}$-HasA$_p$**
Data Collection				
Unit-cell parameters (Å, ^{o})	*a*=34.44, *b*=46.73, *c*=101.04	*a*=34.90, *b*=66.23, *c*=41.00, β=97.1	*a*=34.38, *b*=46.99, *c*=100.90	*a*=34.43, *b*=46.78, *c*=101.81
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)[1]	40.00-0.95 (1.00-0.95)	66.23-1.32 (1.39-1.32)	100.90-0.89 (0.94-0.89)	101.81-1.25 (1.32-1.25)
Wavelength (Å)	1.0000	1.0000	0.9184	1.0000
Temperature (K)	100	100	100	100
Observed reflections	901,848	151,394	714,733	296,237
Unique reflections	97,317	43,290	120,441	46,491
$\langle I/s(I)\rangle$[1]	18.7 (2.6)	7.7 (2.2)	18.0 (2.9)	12.1 (2.5)
Completeness (%)[1]	94.1 (80.0)	99.6 (99.8)	95.8 (87.2)	100 (100)
Multiplicity[1]	9.3 (5.6)	3.5 (3.5)	5.9 (3.7)	6.4 (6.3)
R_{merge} (%)[1, 2]	5.5 (60.1)	9.5 (68.1)	5.0 (38.7)	9.2 (80.6)
R_{meas} (%)[1, 4]	5.8 (66.3)	11.2 (80.5)	5.4 (44.7)	10.0 (88.1)
R_{pim} (%)[1, 4]	1.7 (27.3)	5.9 (42.3)	2.1 (21.8)	3.9 (35.1)
Refinement				
Resolution (Å)	34.30 - 0.95	34.63 - 1.32	32.54 – 0.89	42.51-1.25
Reflections	92,362 / 4,861	41,067 / 2,180	114,343 / 6,000	44,065 / 2,350

(working/test)				
R_{factor} / R_{free} (%)[3]	13.9 / 14.6	14.3 / 16.8	12.4 / 13.4	15.3 / 18.2
No. of atoms (Protein/Heme/ Water)	1,357/43/322	1,356 /43/215	1,436/43/235	1,360/43/213
Model Quality				
R.m.s deviations				
Bond lengths (Å)	0.012	0.010	0.015	0.011
Bond angles (^{o})	1.456	1.356	1.688	1.370
Average *B*-factor (Å^2)				
All Atoms	12.4	13.5	8.9	12.7
Protein	10.3	11.8	7.3	10.9
Heme	8.0	9.2	6.5	8.6
Zinc	-	16.3	-	-
Ethylene glycol	-	23.3	20.0	19.9
Formate	7.4	-	-	-
Water	22.2	24.5	18.9	24.4
Coordinate error, maximum likelihood (Å)	0.07	0.16	0.13	0.30
Ramachandran Plot				
Most favored (%)	98.4	98.9	97.1	98.4
Additionally allowed (%)	1.6	1.1	2.4	1.6

1) Values in parenthesis are for the highest resolution shell.

2) R_{merge} = $S_{hkl}S_i$ $|I_i(hkl) - \langle I(hkl)\rangle|$ / $S_{hkl}S_i$ $I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i*th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.
3) R_{factor} = S_{hkl} $||F_{obs}(hkl)| - |F_{calc}(hkl)||$ / S_{hkl} $|F_{obs}(hkl)|$; Rfree is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.
4) R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}. [35, 42] R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}[43, 44].

4.3 Results

4.3.1 Purification of Y75A and H83A Proteins

Similar to the wt HasAp, the bacterial expression of Y75A and H83A HasAp results in accumulation of a much larger proportion of apo-hemophore relative to its holo-counterpart in the *E. coli* host cells [14]. Hence, the two mutants were purified in their apo-form using a previously reported protocol, which utilizes hydrophobic affinity chromatography to separate apo and holo hemophores [14]. Holo-hemophore was prepared from its corresponding apo form by heme reconstitution in vitro, as described in the Experimental section and both the mutants adopt monomeric structures in the solution. SDS PAGE shows that both mutant proteins migrate with apparent mass of ~ 19 kDa. Mass values measured with the aid of ESI-MS are 18,667 Da (Y75A) and 18,693 Da (H83A), which are in good agreement with the molecular weight calculated from amino acid sequences, 18,668 Da (Y75A) and 18,694 Da (H83A).

4.3.2 UV-vis and Spin State of the Heme Iron in Y75A and H83A HasAp Mutants

Figure 4.2-A shows the electronic absorption spectra of wt (black), Y75A (red) and H83A (blue) HasAp. The UV-vis absorption spectrum of Y75A exhibits a Soret peak at 407 nm and a characteristic high-spin marker charge transfer band at 624 nm. Intense Q bands at 535 and 568 nm also support the presence of low-spin contribution. This high-spin/low-spin equilibrium

is consistent with axial coordination that includes a His side chain and an exogenous ligand as revealed by the X-ray crystal structure of Y75A holo-HasAp (see below). The UV-vis spectrum of H83A is primarily indicative of a high-spin state at room temperature with a broader Soret band at 405 nm, Q bands at 489 and 550 nm, and high-spin marker band at 608 nm (Figure 4.2-A). UV-vis spectra obtained at low-temperature confirm that the low-spin state becomes dominant below 100 K (Figure 4.2-B).

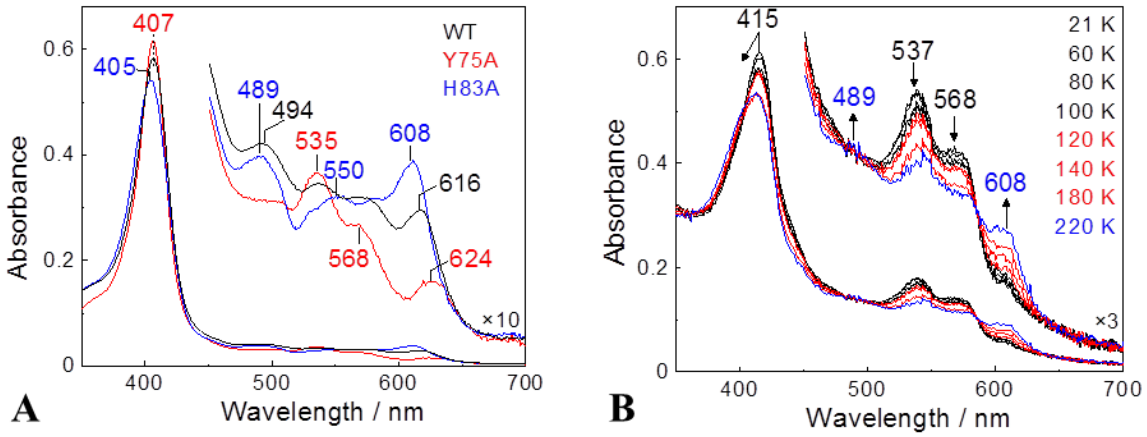


Figure 4.2. (A) Room temperature UV-vis spectra of wt, Y75A and H83A holo-HasAp, (B) Temperature dependence of UV-vis spectra of H83A holo-HasAp. Kinetic experiments were performed by our collaborators, Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU) at Portland, Oregon, USA and details can be found in reference [30].

4.3.3 Homogenous Structural features of Mutants Compared to the Wt Holo-HasAp

The structure of Y75A holo-HasAp was refined using data diffracting to 0.95 Å resolution (Table 4.1). Crystals of H83A holo-HasAp were obtained from three different conditions (see Experimental) and the corresponding structures were refined using data diffracting to 0.89 Å ($H83A^{ortho}$), 1.32 Å ($H83A^{mono}$) and 1.25 Å ($H83A^{pH5.4}$) resolution. The

overall structure of the mutants is nearly identical to that of wt holo-HasAp (Figure 4.3). It is noteworthy that in the H83A mutant, Tyr75 remains coordinated to the heme iron in both structures obtained at pH 6.5 (Figure 4.3-D) as well as in the structure obtained at pH 5.4 (Figure 4.3-E). In the Y75A HasAp structure (Figure 4.3-B), His83 remains in the position it occupies in the wt protein, and therefore does not coordinate the heme iron. In the H83A^{mono} structure the hairpin loop adjoining β6 and β7 is disordered, as suggested by the absence of electron density, which is demarcated by asterisks in Figure 4.3-D. In the H83A^{ortho} structure, strand β8 is shorter than in the wt protein (Figure 4.3-C).

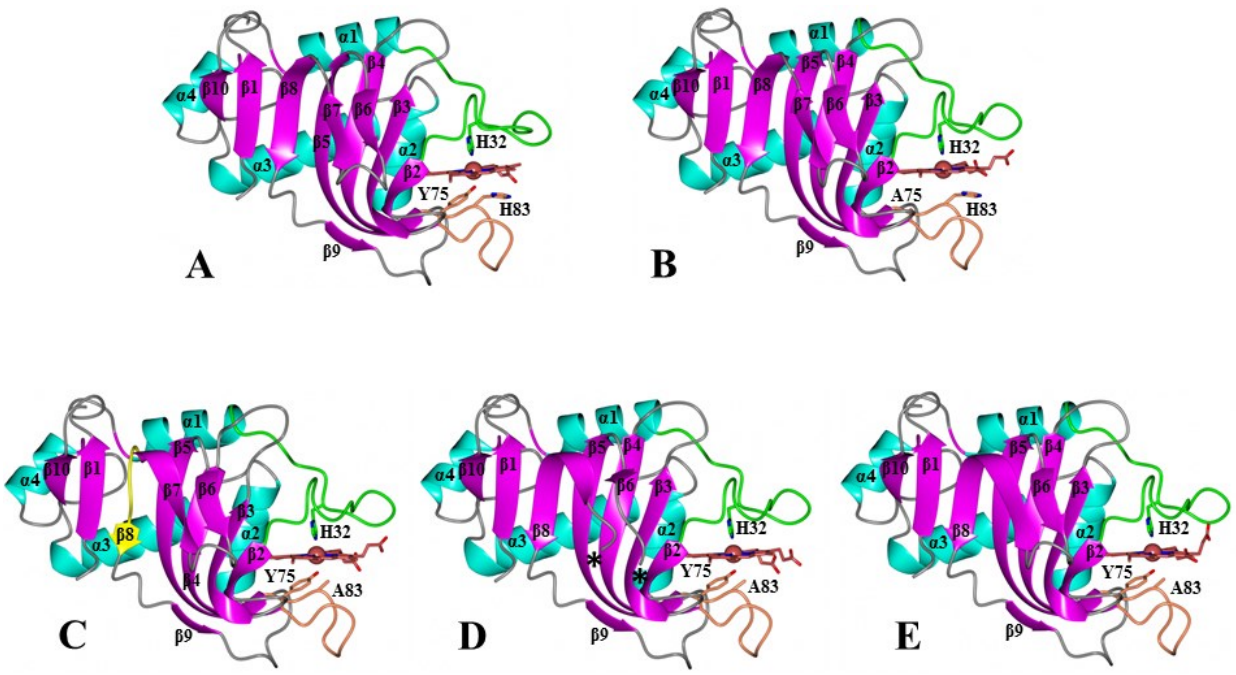


Figure 4.3. The overall structure of the Y75A and H83A mutants is very similar to that of wt HasAp (PDB code: 3ELL). C_{α}-RMSD values from comparing the wild type structure (A) to the structure of Y75A HasAp, (B) is 0.571 Å, H83A^{ortho}-HasAp (C) is 0.564 Å, H83A^{mono}-HasAp (D) is 0.448 Å, and H83ApH5.4 (E) is 0.532 Å. Secondary structures are colored cyan (α-helices), magenta (β-strands) and grey (loops). Loss in secondary structure is colored yellow (C). Missing electron density in (D) is delimited by black asterisks. Residues Ser2 to Ala183 were used to calculate the RMSD values using the secondary structure matching algorithm in the program Superpose [45] via the CCP4 interface [46].

Surprisingly, although both mutations are on the Y75 loop, the conformation of the H32 loop is significantly more affected than the conformation of the Y75 loop. This is illustrated in Figure 4.4, which depicts the conformational changes affecting the backbone and side chains of residues in the H32 loop of H83A-HasAp (cyan) and Y75A-HasAp (green) relative to their conformations in wt HasAp (magenta). In contrast, the structure of the Y75 loop remains nearly unchanged relative to the Y75 loop in wt HasAp. Note, however, that despite the conformational

changes in the H32 loop, the coordinated imidazole rings of His32 in the wt and mutant proteins are identical, as is the seating of the heme.

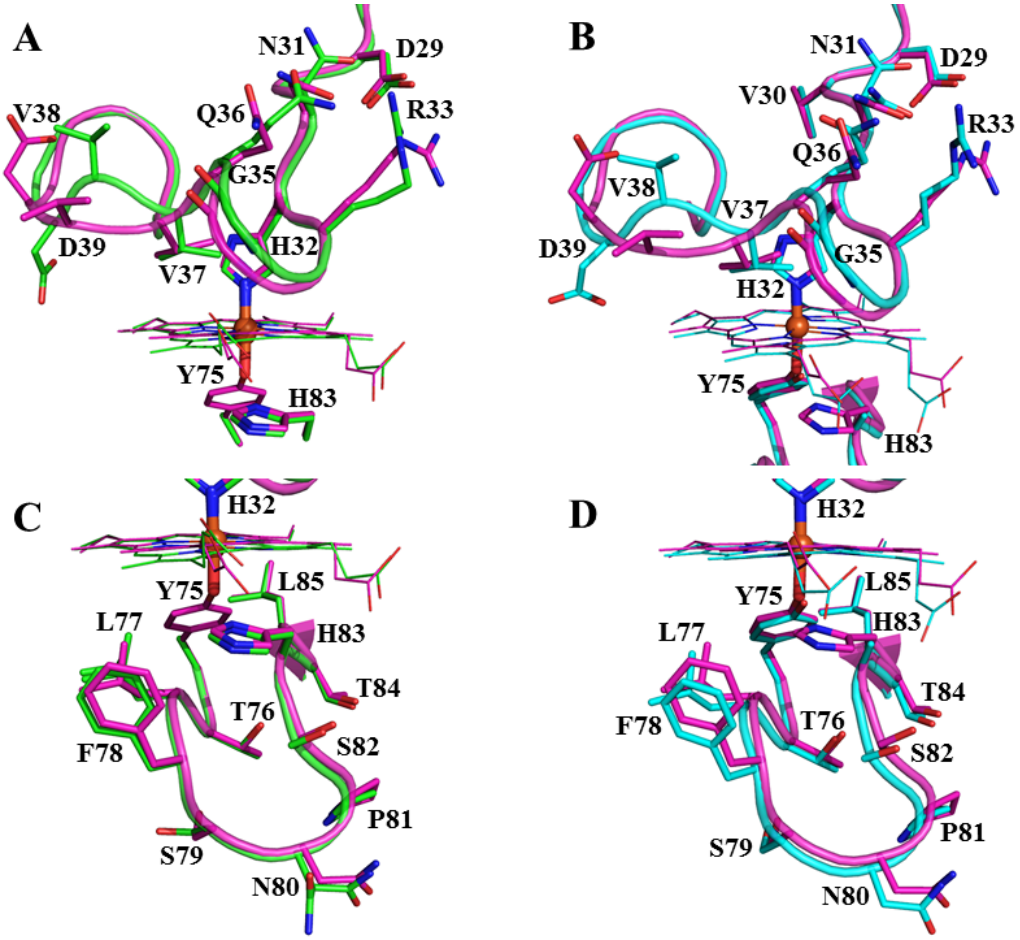


Figure 4.4. Replacing Tyr75 or His83 in the Y75 loop of HasAp by a non-coordinating Alanine causes conformational changes in the His32 loop. Zoomed-in views of the H32 (top) and Y75 (bottom) loops of H83A-HasAp (cyan) and Y75A-HasAp (green) superimposed with the structure of wild type HasAp (magenta). Figure adapted from reference [30].

Heme binds to the Y75A-HasAp mutant in a single orientation, which is evident from the F_o-F_c electron density map in Figure 4.5-A. Given that the heme seating is also identical to that

observed in wt HasAp, it enables similar interactions between the heme propionates and Gly35 and Arg129 (Figure 4.6). Prominent difference electron density ($F_o - F_c$) was observed near the heme on the side that would normally be occupied by Tyr75 in the wt protein. Structure refinement with an ethanol molecule at this site resulted in positive $F_o - F_c$ electron density at the C_2 atom of ethanol, indicating an underestimation of electrons. Placement of a formate ion at this site did not yield residual electron density upon refinement, thus a formate ion was modeled at this site. It is noteworthy that the same electron density was observed in four distinct diffraction data sets ranging in resolution from 0.95 Å to 1.3 Å, which were obtained from four different crystallization conditions (see experimental). Herein we report only the structure obtained from crystals diffracting to 0.95 Å resolution. Given that the protein is initially purified in its apo-form and then reconstituted with heme in vitro, we surmise that the most likely source of formate stems from the heme or the DMSO used to dissolve heme for the heme reconstitution experiments. The plane of the modeled formate ion is aligned with the heme α-γ-meso axis, which enables one of the formate oxygen atoms to coordinate the heme and also form a hydrogen bond with N_δ of His83 (2.6 Å). This hydrogen bonding interaction is similar to the hydrogen bond formed in the wt protein between the coordinated phenolic oxygen of Tyr75 and the N_δ of His83. The second oxygen in the formate molecule interacts with the backbone carbonyl group of Thr76 (3.8 Å). The side chain of His83, which also displays very well defined electron density, is placed in a position identical to that it occupies in the wt protein, where it does not coordinate the heme-iron.

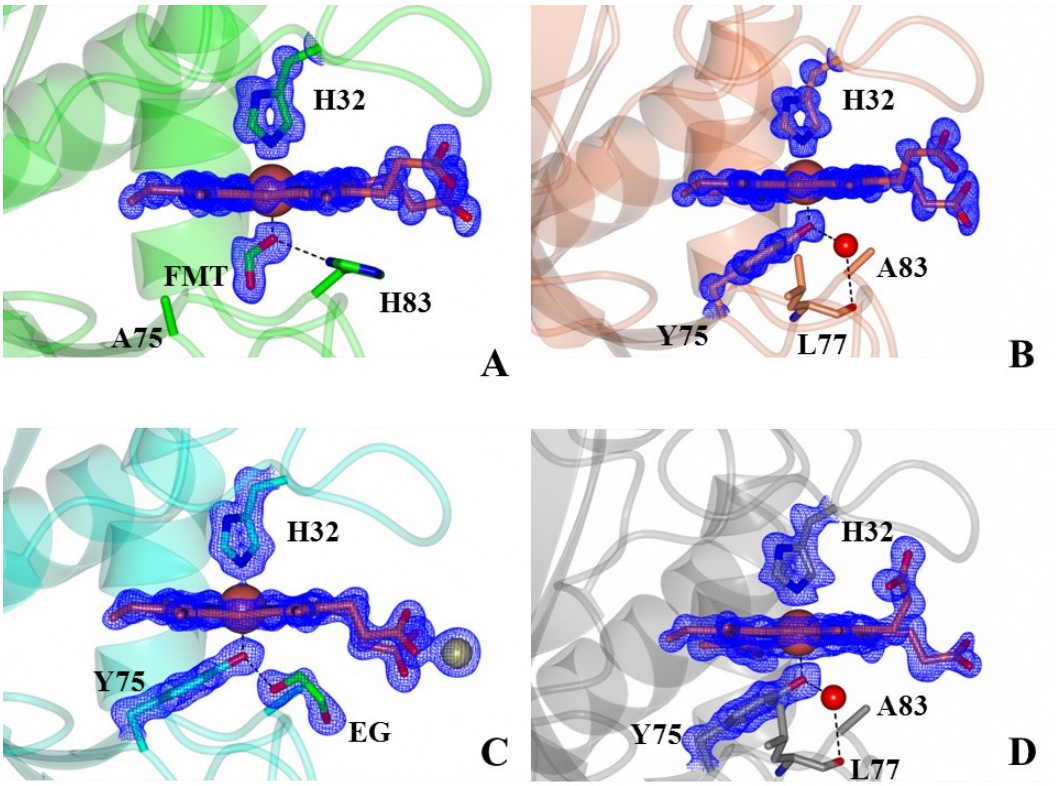


Figure 4.5. Zoomed-in views of the heme-binding sites in (A) Y75A HasAp, (B) H83A^{ortho}-HasAp, (C) H83A^{mono}-HasAp, and (D) H83A$^{pH5.4}$-HasAp. F_o-F_c electron density omit maps contoured at 3σ are shown as blue mesh. FMT is formate ion, EG ethylene glycol, and the Zn^{2+} ion in (C) is shown as a sphere. Figure adapted from reference [30].

Tyr75 and His32 axially coordinate the heme-iron in the H83A^{ortho} (Figure 4.5-B), H83A^{mono} (Figure 4.5-C) and H83A$^{pH5.4}$ (Figure 4.5-D) structures. The F_o-F_c electron density maps show that both axial ligands exhibit very well defined electron density in all structures, as does the heme, which is bound to the protein in a single orientation, thus clearly demonstrating that Tyr75 remains coordinated to the heme iron, even at pH 5.4. The seating of the heme is the same in all three structures and identical to that observed in wt HasAp, although the heme propionates adopt different conformations in the three structures (also see Figure 4.6). In all

structures Tyr75 coordinates the heme iron in a conformation identical to that seen in wt HasAp. In addition to coordinating the heme iron, the phenolic oxygen of Tyr75 forms a hydrogen bond with a water molecule in the $H83A^{ortho}$ and $H83A^{pH5.4}$ structures (Figures 4.5-B and 4.5-D), or with the oxygen atom of an adventitious ethylene glycol (used as cryo-protectant) in the $H83A^{mono}$ structure (Figure 4.5-C). These hydrogen bonding interactions experienced by Tyr75 in the H83A mutant effectively replace the Tyr75 O_η to His83 N_δ hydrogen bond seen in wt HasAp. These observations are in agreement with the significance of the hydrogen bond between Tyr75 and His83 in enabling coordination of Tyr75 to the heme iron [28, 32]. Importantly, however, our finding that Tyr75 coordinates the heme in all available structures is in contrast with conclusions derived from spectroscopic studies suggesting that the Tyr-Fe bond is severed in the H83A mutant of $HasA_{sm}$ [27, 31].

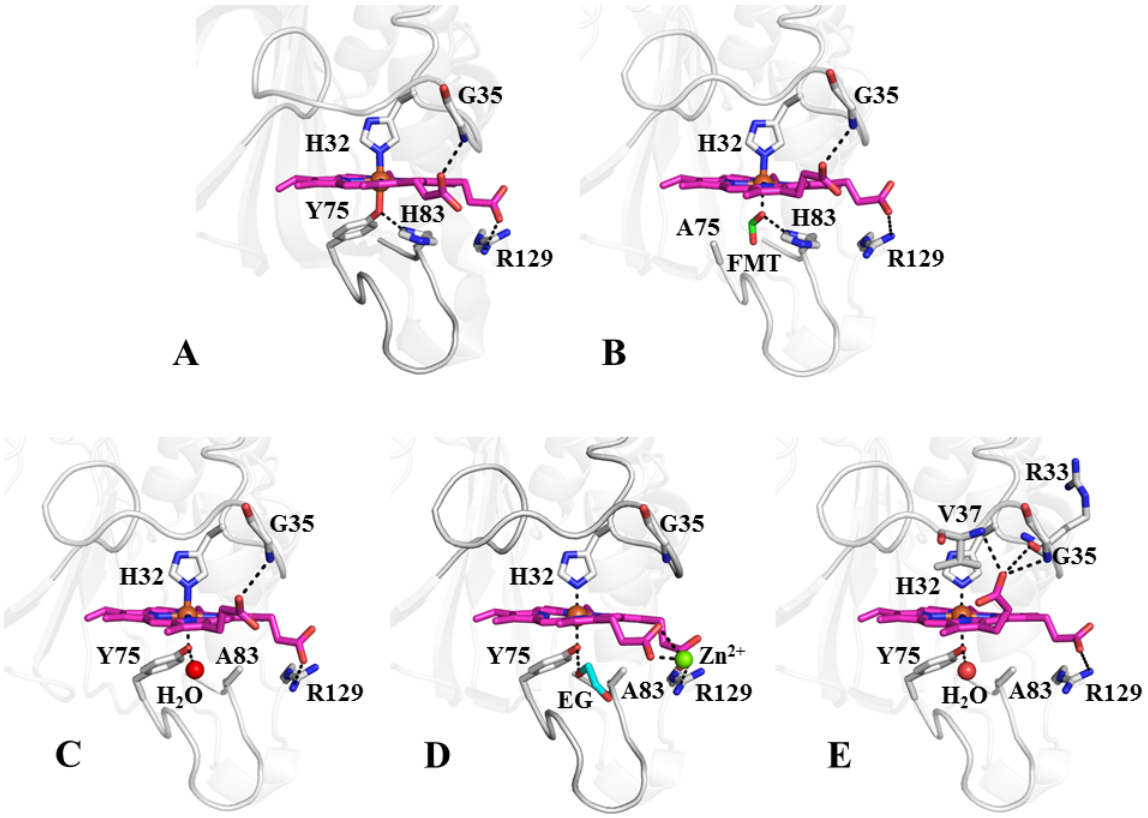


Figure 4.6. Heme propionate interactions in (A) wt holo-HasAp monomer B (PDB code: 3ELL), (B) Y75A-HasAp, (C) H83A^{ortho}-HasAp, (D) H83A^{mono}-HasAp; the zinc ion coordinating the heme propionate is shown as a green sphere and (E) H83A$^{pH5.4}$-HasAp. Figure adapted from reference [30].

During refinement of the H83A^{mono} structure, large peaks greater than 10σ were observed in the F_o–F_c maps, which were tentatively attributed to zinc ions from the crystallization solution. No anomalous signal was observed at these sites using diffraction data from Cu-K_α X-rays. An X-ray fluorescence scan at the synchrotron indicated that zinc was indeed present in the crystal (Figure 4.7-A). Hence, data for structure refinement that were collected at a wavelength of 1.0000 Å yielded appreciable anomalous signal for Zn. In contrast, diffraction data collected at a low energy remote wavelength (1.28414 Å), which resulted in the disappearance of the

anomalous difference density at the Zn sites, and increased anomalous signal at the heme-iron site, confirmed the presence of Zn (Figure 4.7-B). One of the zinc ions interacts with the heme propionates, which induces a conformation distinct from that observed for the heme propionates in wt HasAp. In the H83A$^{pH5.4}$ structure heme propionate 6 interacts with the side chain of Arg129 and thus adopts a conformation similar to that in the wt protein. In contrast, heme propionate 7 adopts a different conformation, which places the carboxylate group within hydrogen bonding distance of the backbone NH and CO groups of Arg33, Gly35 and Val37 (Figure 4.6-E).

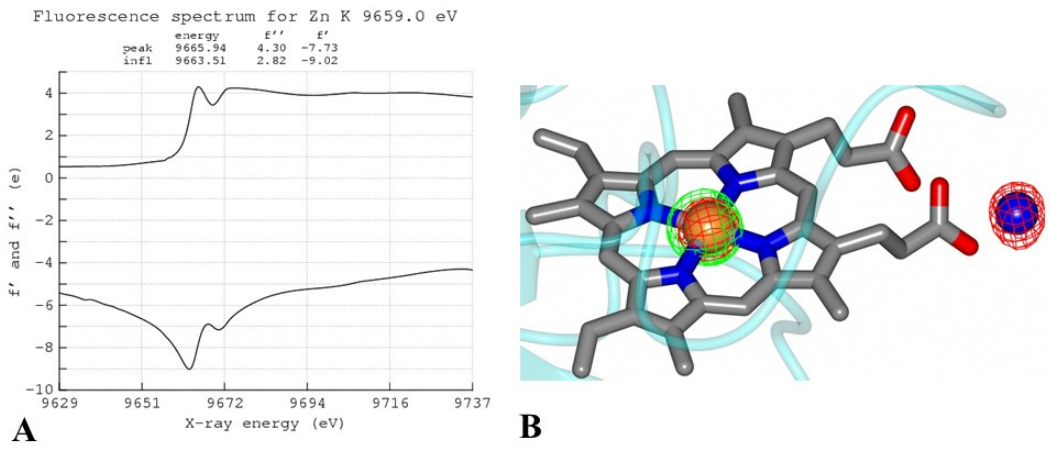


Figure 4.7. (A) X-ray fluorescence scan of the H83A^{mono}-HasAp crystal (B) A zinc ion coordinates the heme propionates in H83A^{mono}-HasAp. Phased anomalous difference map using data collected at λ=1.000 Å (red mesh) and low energy 1.28414 Å (green mesh) contoured at 4σ. Note that the anomalous difference density increases at the heme Fe atom and disappears at the Zn ion site in the data acquired at low energy. Figure adapted from reference [30]

4.3.4 Y75A and H83A Apo-HasAp have Disordered Y75 loops

I would like to thank and acknowledge Dr. Huili Yao, Research associate, Rivera Lab, University of Kansas, Lawrence, Kansas, USA for performing the solution NMR experiments and the details can be found in reference [30]. Results relevant to the analysis are presented in this dissertation.

The X-ray crystal structures of holo-Y75A and holo-H83A-HasAp indicate minimal changes in the structure of the Y75 loop and small changes in the H32 loop relative to the wt protein. Despite considerable effort, however, we were unable to grow crystals of apo-Y75A and apo-H83A HasAp. We hypothesized that conformational disorder in the apo-proteins may prevent crystal growth, thus we conducted NMR spectroscopic studies in solution aimed at probing disorder in the apo-proteins. The ^{1}H-^{15}N HSQC spectrum of each of the mutant apo-proteins, with the exception of a few cross-peaks, is nearly superimposable to the HSQC spectrum of wt apo-HasAp (Figure 4.8), which indicates that overall, the structures of the mutant apo-hemophores are similar to the structure of wt apo-HasAp. Given the large similarity in the HSQC spectra of wt and mutant proteins, the sequential backbone assignments of apo-Y75A and apo-H83A HasAp were obtained relatively readily (for detailed backbone assignments, please refer reference [30]).

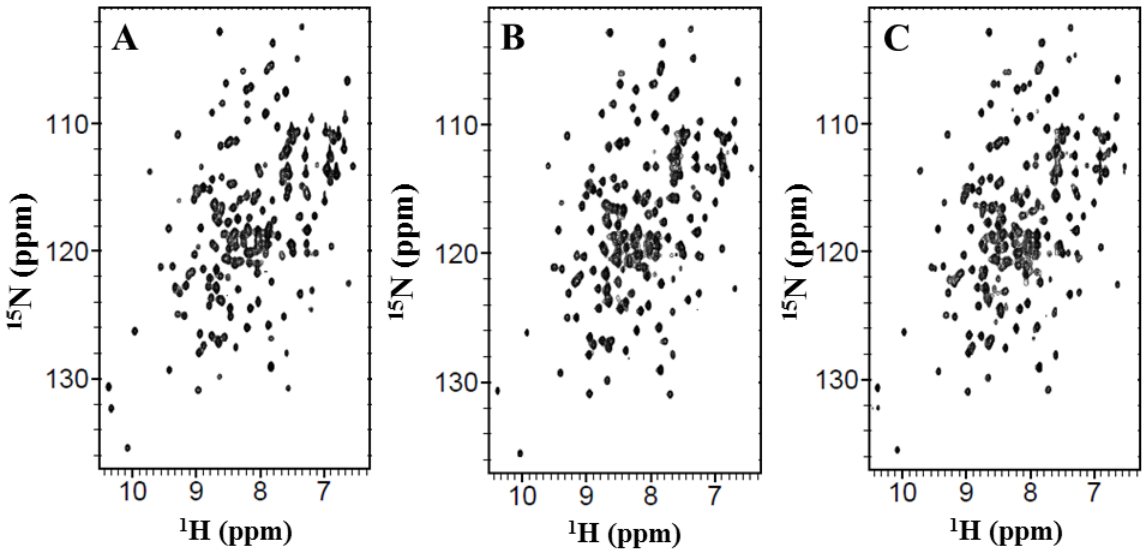


Figure 4.8. ^{1}H,^{15}N-HSQC spectra of wt apo-HasAp (A), Y75A apo-HasAp (B), and H83A apo-HasAp (C). Spectra were taken at 32 °C using a 600MHz Bruker spectrometer. Acquisition parameters are as follows: 2048 (^{1}H) × 256 (^{15}N) complex points; 10.7 kHz (^{1}H) and 2.4 kHz (^{15}N) spectral widths; 32 scans per increment; d1=1sec. Solution NMR experiments were performed by Dr. Huili Yao, Research Associate, Rivera's Lab, The University of Kansas, KS, USA Figure adapted from reference [30].

Closer inspection of the HSQC spectra show mostly small differences in the weighed chemical shifts ($\Delta\delta_{weighed}$) obtained from comparing the chemical shifts of cross peaks in apo-Y75A (Figure 4.9-A), or apo-H83A (Figure 4.9-B) with corresponding cross-peaks in wt apo-HasAp. Importantly, the comparison also reveals that several cross peaks in the spectra of the mutants are (Figure 4.9) (*i*) missing (red triangles), (*ii*) are broadened and have significant lower intensity compared to their counterparts in the spectrum of wt HasAp (blue squares), or (*iii*) are split (green rhombs). Mapping the missing, broadened and split cross-peaks in the spectra of the mutants onto the X-ray crystal structure of wt apo-HasAp using the same color scheme as in the

plots (Figure 4.9) reveals that the Y75 loop, the neighboring α2 helix, some of the β strands near the loop, and a portion of the distal loop emerging from β2 are affected by conformational disorder in both mutants.

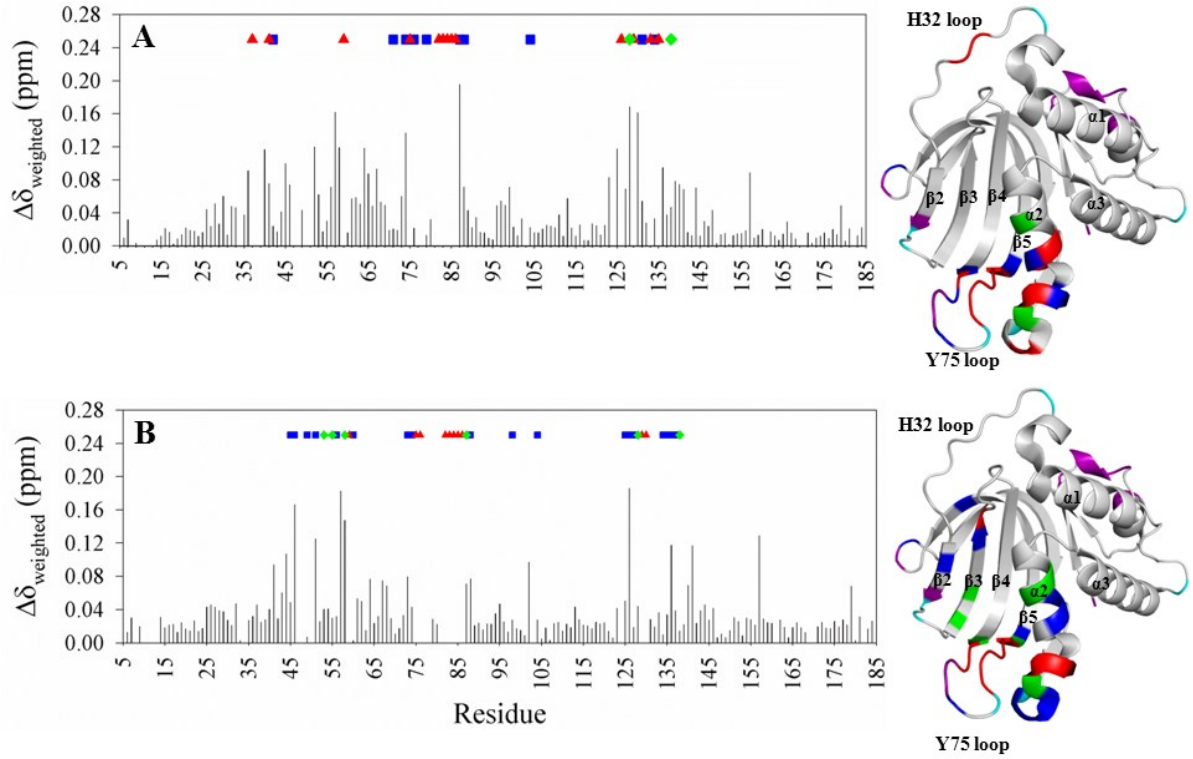


Figure 4.9. Per-residue differences obtained from comparing cross-peaks in the HSQC spectrum of wt apo-HasAp with cross-peaks in the HSQC spectra of (A) Y75A apo-HasAp and (B) H83A apo-HasAp. Residues in the mutants whose cross peaks disappear are highlighted with a red triangle, residues whose cross peaks split are denoted by a green rhomb and residues whose cross-peaks are broadened and have decreased intensity are highlighted by a blue square. These residues have been mapped on the wt apo-HasAp structure (PDB: 3MOK) using the same color code; purple denotes residues with no assignment, and cyan identifies prolines. $\Delta\delta_{weighed} = \sqrt{\left[\left(\frac{\Delta\delta N}{5}\right)^2 + (\Delta\delta H)^2\right]/2}$ Solution NMR experiments were performed by Dr. Huili Yao, Research Associate, Rivera's Lab, The University of Kansas, KS, USA. Figure adapted from reference [30].

4.3.5 Conserved Biphasic Kinetics of Hemin loading in Y75A and H83A apo-HasAp

Spectroscopic experiments were done by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU), Portland, Oregon, USA. Results relevant to the analysis are presented in this dissertation.

In the previous work [47], the analysis of biphasic behavior of hemin loading to wt apo-HasAp was done with a multi-step model, with the formation of reversible hemin-apoprotein complex proceeds quickly before slower iron(III) coordination steps begin to take place [47]. Similar to the wt apo-HasAp [26], the stopped-flow absorption data of hemin acquisition by the Y75A and H83A mutants correspond to a biphasic processes

$$ApoHasAp + hemin \underset{k_1,\ k_{-1}}{\rightleftharpoons} [\ ApoHasAp{\bullet}hemin\ ] \xrightarrow[k_{coord1}]{} HasAp(L1)\text{-}hemin \xrightarrow[k_{coord2}]{} HoloHasAp$$

Assuming that $k_1[ApoHasAp]$ and $k_{-1} >> k_{coord1}$, and $k_{coord1} >> k_{coord2}$ gives eq 1

$$k_{1obs} = k_{coord1}[apoHasAp]/\{(k_{-1}+k_{coord1})/k_1 + [apoHasAp]\} \approx k_{coord1}[apoHasAp]/\{K_d+[apoHasAp]\}$$

eq 1.

This model explains the hyperbolic dependence of k_{1obs} on the apo-HasAp concentration. In contrast, k_{2obs} shows no dependence on the apo-HasAp concentration, as expected for an intramolecular reaction that corresponds to the closure of the H32 loop and coordination of the iron by His32.

This same analysis was used to extract K_d, k_{coord1}, and k_{coord2} values from the stopped-flow data obtained with the variant proteins (Table 4.2). The Y75A substitution results in a 4-fold higher K_d value without significant change in the rate of the first coordination step, k_{coord1}. The second kinetic phase reflects k_{coord2} and is ~3 times faster in Y75A than in wt-HasAp, but it remains low, i.e., in the millisecond timescale expected for a significant protein loop rearrangement (Table 4.2).

Table 4.2. Kinetic parameters for hemin binding to wt and mutant HasAp proteins at 4 °C.

	Hemophore			
Kinetic Parameter	**WT**	**H32A**	**Y75A**	**H83A**
$\boldsymbol{K_d}$ **(μM)**	35	61	129	30
$\boldsymbol{k_{coord1}}$ **(s^{-1})**	552	830	539	118
$\boldsymbol{k_{coord2}}$ **(s^{-1})**	5.2	n.o.	14.4	20.5

The k_{1obs} values for H83A apo-HasAp also follow a hyperbolic dependence on apo protein concentration as seen with the Y75A variant, but the curve for H83A rises quicker and achieves a lower plateau indicating lower K_d and k_{coord1} values (Figure 4.10). Compared to wt HasAp, the H83A substitution does not appear to change the K_d value, but k_{coord1} is lowered 5-fold. In contrast, the k_{coord2} value is increased ~4 fold in the H83A variant compared to wt-HasAp (Table 4.2).

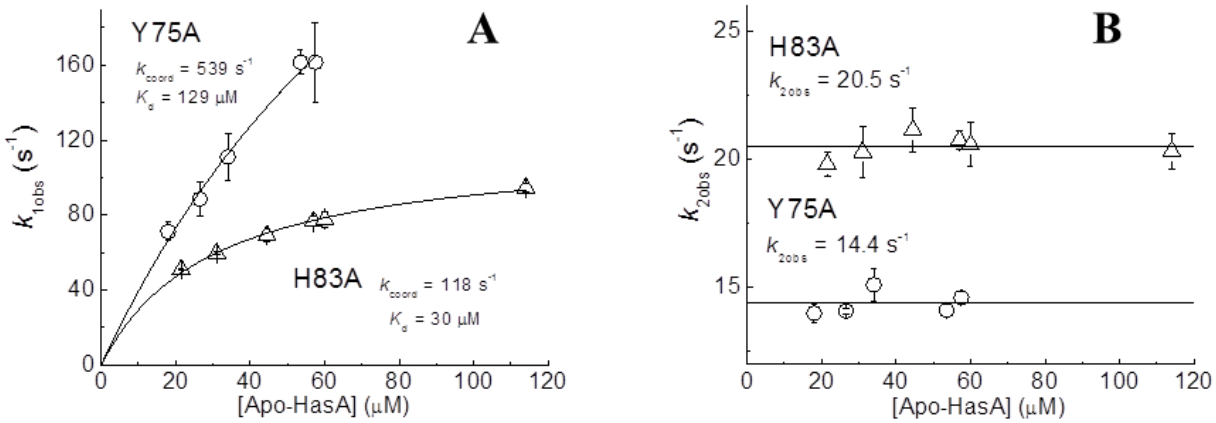


Figure 4.10. Plots of observed rate constants vs apo-HasAp concentration. Spectroscopic experiments were performed by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU), Portland, Oregon, USA. Figure adapted from reference [30].

4.3.6 The millisecond intermediate in Y75A is likely coordinated by His83, while in H83A, Y75 is the likely axial ligand to the hemin iron.

Rapid freeze quench resonance Raman (RFQ-RR) experiments were performed in an effort to characterize the millisecond intermediate heme-protein complexes in Y75A- and H83A-HasAp. The high-frequency RR spectrum of the 6-ms RFQ sample of Y75A shows porphyrin modes indicative of a mixture of 5cHS and 6cHS species with ν_3 modes at 1492 and 1481 cm^{-1}, respectively (Figure 4.11-A). High-spin heme iron (III) species with O-atom axial coordination typically exhibit intense ν_3 modes relative to the ν_4 and ν_2 modes [18, 48-54], and the high ν_3/ν_4, ν_2 intensity ratio observed in the RFQ-RR spectra of the millisecond intermediate formed in wt-HasAp supported the notion that Y75 is the first coordinating ligand to the iron (III) [26]. In contrast, the RR spectra of the 6-ms samples of Y75A show weak ν_3 relative to the ν_2 and ν_4

modes which suggest that the initial coordination to the iron (III) involves an endogenous ligand, presumably His83, rather than a solvent molecule. Over time, these high-spin contributions are replaced by 6cLS modes with ν_3 and ν_2 at 1505 and 1582 cm^{-1}, respectively. On the basis of the crystal structure of Y75A holo-HasAp, we assign this transition to the closure of the H32 loop on the heme pocket and intercalation of a solvent molecule between the heme iron (III) and His83.

The high-frequency RR spectra of the 6-ms RFQ sample of H83A exhibit ν_4, ν_3, and ν_2 modes at 1374, 1493, and 1572 cm^{-1} that are characteristic of a pure 5cHS species with an O-atom axial ligand; they are analogous to but spectrally distinguishable from free hemin (Figure 4.11-B). These 5cHS porphyrin skeletal modes are also different from the 5cHS signals observed in the 6-ms RFQ sample of wt HasAp (Figure 4.11-C). Longer reaction times allow for the conversion of this 5-coordinate complex to a mixture of 6cHS and 6cLS heme conformers. Attempts to detect vibrational modes involving iron (III) axial ligands, such as tyrosinate modes or ν(Fe-OH) modes using off-Soret excitation and $^{18}OH_2/D_2O$ solvent exchange were unsuccessful for both the Y75A and H83A variants.

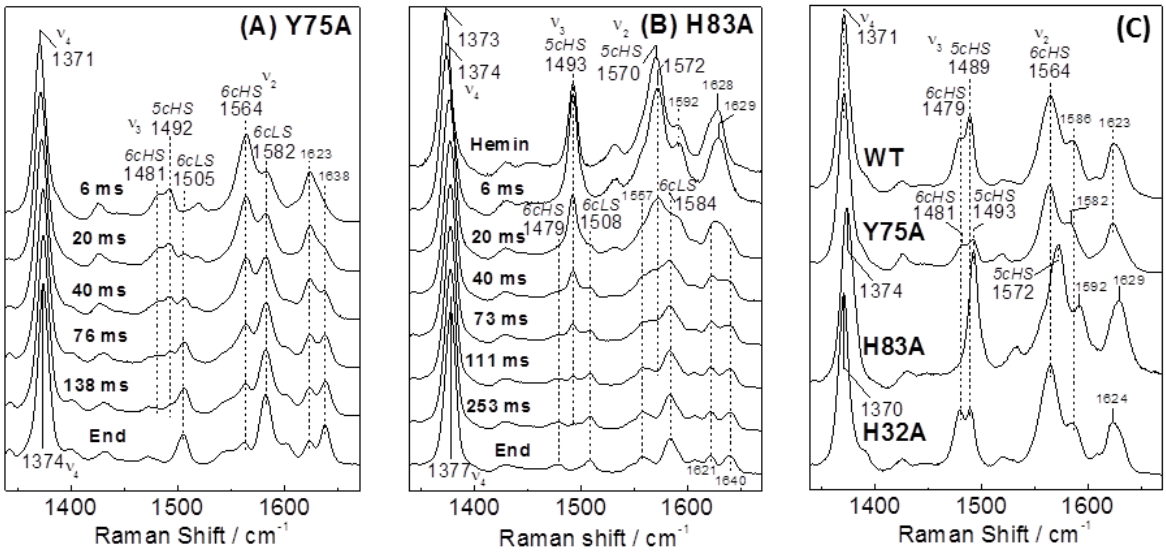


Figure 4.11. High-frequency RR spectra of RFQ samples of the reaction of 1 equivalent of hemin with (A) Y75A and (B) H83A apo-HasAp (protein concentration after mixing is 300 μM; excitation wavelength 406 nm; all spectra are normalized to the intensity of the ν_4 feature observed between 1371 and 1377 cm^{-1}). (C) RR spectra of 6-ms RFQ samples. Spectroscopic experiments were performed by our collaborators Dr. Hirotoshi Matsumura and Dr. Pierre Moënne-Loccoz at Oregon Health and Science University (OHSU), Portland, USA. Figure adapted from reference [30].

4.4 Discussion

4.4.1 Y75A and H83A Substitutions do not affect the Structure of Holo-HasAp.

Previous investigations of the process of hemin loading onto HasAp showed that within 20 ms hemin binds first to the Y75 loop. This rapid process, which is believed to include the coordination of Tyr75 to the heme iron, is followed by a slower reorganization of the H32 loop

and coordination of the heme by His32, which occurs within ~ 1 s at 4 °C [14, 26]. In contrast to the large rearrangement of the His32-bearing loop, the Y75 loop retains its structure in the apo- and holo-forms of the hemophore. The structural conservation includes the Tyr75 O_η to His83 N_δ hydrogen bond, which is thought to play a pivotal role in hemin binding by increasing the phenolate character of Tyr75 axial ligand and enabling it to coordinate the ferric ion in hemin. Yet, our high-resolution crystal structures of the Y75A and H83A holo-HasAp mutants are remarkably similar to that of wt holo-HasAp. In particular, in all three H83A holo-HasAp structures the heme iron is coordinated by both His32 and Tyr75 and the structure of the Y75 loop is unaltered relative to wt holo-HasAp (see Figures 4.3 and 4.5). Interestingly, the void left by the His83-to-Ala substitution is filled by a water molecule in the $H83A^{ortho}$ structure at pH 6.5 and in the structure at pH 5.4, and this crystallographically defined water molecule is hydrogen bonded to the phenol oxygen of the coordinated Tyr75. In the $H83A^{mono}$ structure the same void is filled with an ethylene glycol molecule that acts as a hydrogen bond partner to the coordinated Tyr75. It is also noteworthy that the side chain of Tyr75 remains coordinated to heme, even at pH 5.4, despite the absence of the His83 side chain. Clearly, these observations demonstrate that His83 and the Tyr75 O_η to His83 N_δ hydrogen bond are not essential for maintaining the integrity of the heme-binding pocket in holo-HasAp or the coordination of hemin-iron by Tyr75.

In the Y75A holo-HasAp structure we observed well defined elongated electron density below the hemin iron, which was modeled as a formate ion (Figure 4.5A); as indicated above, the most likely source of this formate ion is our hemin stock. Nevertheless, the structure of the loop, including the position of the His83 side chain, is superimposable with that observed in the Y75 loop of holo wt HasAp, suggesting that in the absence of formate, the heme iron is coordinated to

a water molecule and His32, as has been suggested on the basis of spectroscopic studies with holo-Y75A $HasA_{sm}$ [27].

4.4.2 Partially Disordered Y75 loop in the variant apo proteins.

In contrast to the success we experienced growing single crystals of wt, Y75A and H83A holo-HasAp, we were unable to grow crystals of the apo form of these variant proteins despite extensive screening of conditions. Because we suspected conformational disorder as the reason hindering crystal growth, we turned to solution NMR and obtained the sequential backbone assignments of wt, Y75A and H83A apo-HasAp. As expected, the NMR spectra of wt apo-HasAp are consistent with a well-ordered structure, which is in agreement with the well-defined electron density defining its X-ray crystal structure [14]. The HSQC spectra of apo-Y75A or apo-H83A HasAp are nearly identical to the HSQC spectrum of their wt counterpart. Detailed inspection, however, reveals that several cross peaks corresponding to residues in the Y75 loop, helix α2 and several β strands in Y75A and H83A apo-HasAp are affected by conformational disorder (see Figure 4.9). The crystal structure of wt apo-HasAp (Figure 4.12) shows that Tyr75 and His83 are involved in several hydrogen bonding and hydrophobic interactions within the loop. Figure 4.12-A shows the hydrogen bond between the phenol oxygen of Tyr75 and the N_δ of His83, as well as the hydrogen bonds between the carbonyl oxygen of His83 and the side chains of Arg129 and His134 in α2, and the Leu85 carbonyl to His134 NH hydrogen bond, which mediate interactions between the Y75 loop and helix α2. Also apparent is a group of hydrophobic interactions (Figure 4.12-B) that include the hydrophobic portions of the Tyr75 and His83 side chains, helix α2 (Met141), strands β5 (Leu85), β4 (Leu73), β3 (Tyr56), β2 (Phe46), and the hairpin loop connecting β2 and β3 (Phe51). This extensive and intricate network of H-bonding and hydrophobic interactions which propagates into a significant portion of the structure

stabilizes the structure of the Y75 loop in the apo-protein, but replacing Tyr75 or His83 with Ala results in conformational disorder of the Y75 loop in apo-HasAp.

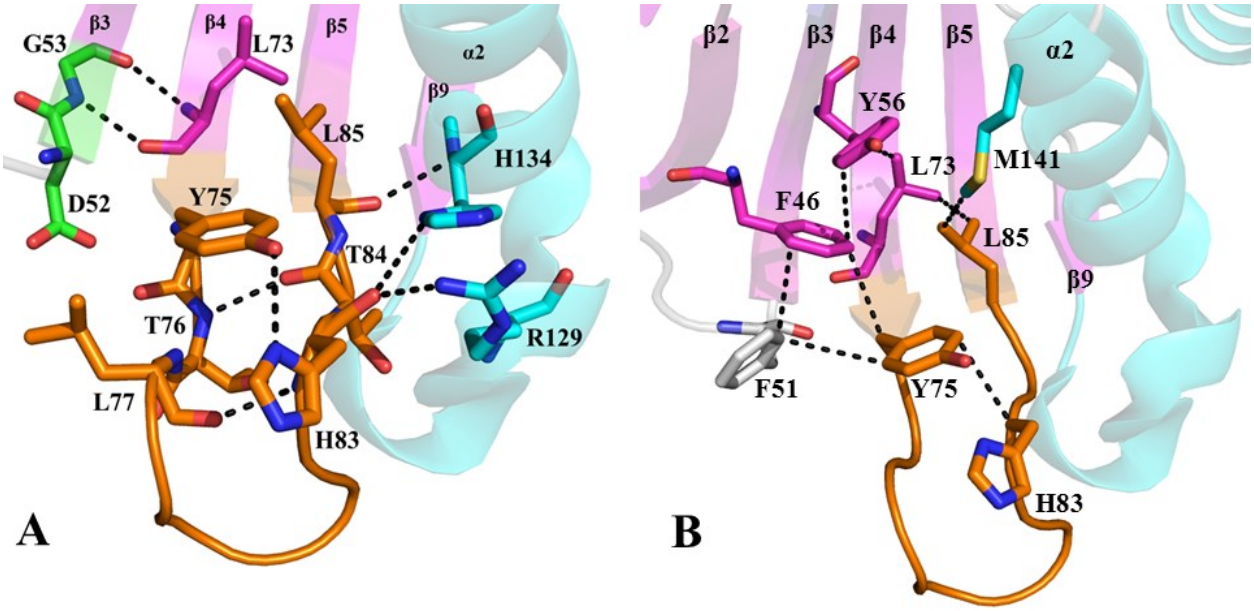


Figure 4.12. Close-up view of wt apo-HasAp (PDB: 3MOK) (A) hydrophilic and (B) hydrophobic interactions emanating from the Tyr75 loop into helix α2 and nearby β strands. β-strands are shown in magenta, α-helices in cyan and the Tyr75 loop in orange, with corresponding residues in stick representation. Figure adapted from reference [30].

4.4.3 Formation of Distinct Millisecond Intermediates in the Mutants

The relatively severe changes introduced to the Y75 loop and concomitant conformational disorder into neighboring structural domains in Y75A and H83A apo-HasAp is likely to result in initial distribution of [apo-HasAp●hemin] sub-states during the process of hemin loading. However, rapid interconversion between these [apo-HasAp●hemin] sub-states is

likely to take place, and the conformational disorder observed in Y75A and H83A apo-HasAp does not dramatically impact the rates of hemin loading in the variants compared to the wt protein (Table 4.2). Both variants show K_d values for the initial [apo-HasAp●hemin] complex in the same micromolar range as the wt protein, and the first coordination event defined by k_{coord1} occurs in similar timeframes for all three proteins. The variants show slightly higher k_{coord2} than the wt, but here again the differences are small.

In contrast, the RFQ-RR analyses of the two variant proteins clearly reveal that they form distinct millisecond intermediates. In Y75A, the intermediate shows intensity ratio between ν4 and ν3 porphyrin skeletal modes consistent with coordination of His83 to the heme iron before closure of the H32 loop, coordination of H32 and displacement of the H83 ligand by a water molecule; coordination of His 32 and rearrangement at the Fe-His83 site are not kinetically resolved and are possibly coupled events. In comparison, the RFQ-RR spectrum of the intermediate in H83A supports iron coordination by an O-atom, either from solvent or from Tyr75, before coordination of the second axial ligand, His32. Importantly, despite these differences in coordination geometry in the [HasAp-hemin] millisecond intermediates of the variant and wt proteins, all three proteins follow equivalent kinetic profiles.

We have previously proposed that π-π stacking interactions between residues in the Y75 loop and hemin contribute significantly to a "sticky" hydrophobic platform that facilitates hemin capture [14, 16]. Within this framework, the Tyr75 O_η to His83 N_δ hydrogen bond contributes to the hemin loading process by positioning the Tyr O_η to coordinate the heme iron (Figure 4.13).

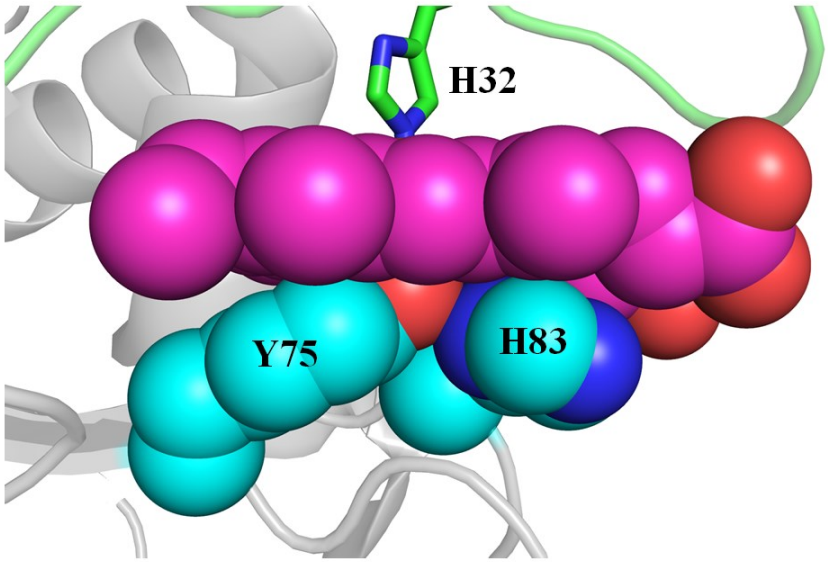


Figure 4.13. Zoomed-in view of the Y75 loop in wild type holo-HasAp (PDB: 3ELL) illustrating the hydrophobic and π-stacking interactions between the Y75 and H83 side chains with the heme. Figure adapted from reference [30].

In absence of a Tyr75 O_η to His83 N_δ hydrogen bond in the apo-Y75A mutant, the conformational disorder of residues in the Y75 loop may permit coordination of the incoming hemin by His83. Subsequent scission of the His83-Fe coordination bond and movement of the His83 side chain away from the iron and toward its position in the structure of wt holo-HasAp, is likely a consequence of loop reorganization driven by the establishment of productive π-π stacking and hydrophobic interactions between hemin and side-chains in the Y75-loop. Isothermal titration calorimetry measurements with $HasA_{sm}$ revealed ≈ 250- and ≈ 400-fold decrease in heme affinity in the H83A and Y75A variants relative to the wt protein [32]. Our pre-steady state measurements show that these mutations have little impact on the formation

rates of holoproteins and therefore suggest that changes in binding affinity primarily reflect increases in dissociation rates. Consequently, the data suggest that axial ligation plays a prominent role in slowing hemin loss from the hemophore, but the on-rate of hemin loading onto the apo hemophore is governed by non-covalent hydrophobic interactions. In this context, the changes in K_d and k_{coord1} observed with the Y75A and H83A variants relative to wt HasAp likely reflect changes in hydrophobic and π-π stacking interactions between the porphyrin macrocycle and the Y75 loop. Hence, the data at hand suggest that HasA hemophores utilize hydrophobic interactions to capture hemin while controlling its release via the heme iron axial coordination. Future studies aimed at dissecting the thermodynamics and kinetics of hemin capture and release with complementary variants, such as Y75F and H83F, Y75F/H83F, and Y75A/H83A will contribute to establish this mechanistic model.

4.5 Bibliography

1. Nobles CL, Maresso AW. The theft of host heme by Gram-positive pathogenic bacteria. *Metallomics : integrated biometal science*. 2011 Aug;3(8):788-96.
2. Grigg JC, Ukpabi G, Gaudin CF, Murphy ME. Structural biology of heme binding in the Staphylococcus aureus Isd system. *Journal of inorganic biochemistry*. 2010 Mar;104(3):341-8.
3. Weinberg ED. Iron Availability and Infection. *Biochim et Biophys Acta*. 2009;1790:600-5.
4. Bullen JJ, Rogers HJ, Spalding PB, Ward CG. Iron and Infection: The Heart of the Matter. *FEMS Immunol Med Microbiol*. 2005;43:325-30.
5. Otto BR, Verweij-van Vught AM, MacLaren DM. Transferrins and Heme-Compounds as Iron Sources of Pathogenic Bacteria. *Crit Rev Microbiol*. 1992;18:217-33.
6. Ascenzi P, Bocedi A, Visca P, Altruda F, Tolosano E, Beringhelli T, Fasano M. Hemoglobin and heme scavenging. *IUBMB life*. 2005 Nov;57(11):749-59.
7. Benson DR, Rivera M. Heme Uptake and Metabolism in Bacteria. Met Ions Life Sci. 122013. p. 279-332.
8. Andrews SC. Iron Storage in Bacteria. *Advan Microbial Physiol*. 1998;40:281-351.
9. Tullius MV, Harmston CA, Owens CP, Chim N, Morse RP, McMath LM, Iniguez A, Kimmey JM, Sawaya MR, Whitelegge JP, Horwitz MA, Goulding CW. Discovery and characterization of a unique mycobacterial heme acquisition system. *Proceedings of the National Academy of Sciences of the United States of America*. 2011 Mar 22;108(12):5051-6.

10. Ghigo JM, Letoffe S, Wandersman C. A new type of hemophore-dependent heme acquisition system of Serratia marcescens reconstituted in Escherichia coli. *Journal of bacteriology*. 1997 Jun;179(11):3572-9.

11. Letoffe S, Ghigo JM, Wandersman C. Iron acquisition from heme and hemoglobin by a Serratia marcescens extracellular protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1994 Oct 11;91(21):9876-80.

12. Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. The crystal structure of HasA, a hemophore secreted by Serratia marcescens. *Nature structural biology*. 1999 Jun;6(6):516-20.

13. Wolff N, Izadi-Pruneyre N, Couprie J, Habeck M, Linge J, Rieping W, Wandersman C, Nilges M, Delepierre M, Lecroisey A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*. 2008 Feb 15;376(2):517-25.

14. Jepkorir G, Rodriguez JC, Rui H, Im W, Lovell S, Battaile KP, Alontaga AY, Yukl ET, Moenne-Loccoz P, Rivera M. Structural, NMR spectroscopic, and computational investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa. *Journal of the American Chemical Society*. 2010 Jul 21;132(28):9857-72.

15. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P, Rivera M. Structural characterization of the hemophore HasAp from Pseudomonas aeruginosa: NMR spectroscopy reveals protein-protein interactions between Holo-HasAp and hemoglobin. *Biochemistry*. 2009 Jan 13;48(1):96-109.

16. Kumar R, Lovell S, Matsumura H, Battaile KP, Moenne-Loccoz P, Rivera M. The hemophore HasA from Yersinia pestis (HasAyp) coordinates hemin with a single residue, Tyr75, and with minimal conformational change. *Biochemistry*. 2013 Apr 23;52(16):2705-7.

17. Aki Y, Nagai M, Nagai Y, Imai K, Aki M, Sato A, Kubo M, Nagatomo S, Kitagawa T. Differences in coordination states of substituted tyrosine residues and quaternary structures among hemoglobin M probed by resonance Raman spectroscopy. *J Biol Inorg Chem*. 2010 Feb;15(2):147-58.

18. Das TK, Couture M, Lee HC, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB, Guertin M. Identification of the ligands to the ferric heme of Chlamydomonas chloroplast hemoglobin: evidence for ligation of tyrosine-63 (B10) to the heme. *Biochemistry*. 1999 Nov 16;38(46):15360-8.

19. Moore CD, Lecomte JTJ. Structural Properties of Apocytochrome b_5: Presence of a Stable Native Core. *Biochemistry*. 1990;29:1984-9.

20. Williams PA, Fulop V, Garman EF, Saunders NF, Ferguson SJ, Hajdu J. Haem-ligand switching during catalysis in crystals of a nitrogen-cycle enzyme. *Nature*. 1997 Sep 25;389(6649):406-12.

21. Jin Y, Nagai M, Nagai Y, Nagatomo S, Kitagawa T. Heme structures of five variants of hemoglobin M probed by resonance Raman spectroscopy. *Biochemistry*. 2004 Jul 6;43(26):8517-27.

22. Wardell M, Wang Z, Ho JX, Robert J, Ruker F, Ruble J, Carter DC. The atomic structure of human methemalbumin at 1.9 A. *Biochem Biophys Res Commun*. 2002 Mar 8;291(4):813-9.

23. Reid TJ, 3rd, Murthy MR, Sicignano A, Tanaka N, Musick WD, Rossmann MG. Structure and heme environment of beef liver catalase at 2.5 A resolution. *Proc Natl Acad Sci U S A*. 1981 Aug;78(8):4767-71.

24. Ho WW, Li H, Eakanunkul S, Tong Y, Wilks A, Guo M, Poulos TL. Holo- and apo-bound structures of bacterial periplasmic heme-binding proteins. *The Journal of biological chemistry*. 2007 Dec 7;282(49):35796-802.

25. Mattle D, Zeltina A, Woo JS, Goetz BA, Locher KP. Two stacked heme molecules in the binding pocket of the periplasmic heme-binding protein HmuT from Yersinia pestis. *Journal of molecular biology*. 2010 Nov 26;404(2):220-31.

26. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moenne-Loccoz P. Kinetic and spectroscopic studies of hemin acquisition in the hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 Aug 10;49(31):6646-54.

27. Caillet-Saguy C, Piccioli M, Turano P, Lukat-Rodgers G, Wolff N, Rodgers KR, Izadi-Pruneyre N, Delepierre M, Lecroisey A. Role of the Iron Axial Ligands of Heme Carrier HasA in Heme Uptake and Release. *Journal of Biological Chemistry*. 2012 Aug 3;287(32):26932-43.

28. Wolff N, Deniau C, Letoffe S, Simenel C, Kumar V, Stojiljkovic I, Wandersman C, Delepierre M, Lecroisey A. Histidine pK(a) shifts and changes of tautomeric states induced by the binding of gallium-protoporphyrin IX in the hemophore HasA(SM). *Protein Sci*. 2002 Apr;11(4):757-65.

29. Caillet-Saguy C, Delepierre M, Lecroisey A, Bertini I, Piccioli M, Turano P. Direct-detected 13C NMR to investigate the iron(III) hemophore HasA. *J Am Chem Soc*. 2006 Jan 11;128(1):150-8.

30. Kumar R, Matsumura H, Lovell SW, Yao H, Rodriguez JC, Battaile KP, Moenne-Loccoz P, Rivera M. Replacing the Axial Ligand Tyrosine 75 or its Hydrogen Bond Partner Histidine 83 Minimally Affects Hemin Acquisition by the Hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2014 Mar 13.

31. Caillet-Saguy C, Turano P, Piccioli M, Lukat-Rodgers GS, Czjzek M, Guigliarelli B, Izadi-Pruneyre N, Rodgers KR, Delepierre M, Lecroisey A. Deciphering the structural role of histidine 83 for heme binding in hemophore HasA. *J Biol Chem*. 2008 Feb 29;283(9):5960-70.

32. Deniau C, Gilli R, Izadi-Pruneyre N, Lettofe S, Delepierre M, Wandersman C, Briand C, Lecroisey A. Thermodynamics of Heme Binding to the $HasA_{SM}$ Hemophore: Effect of Mutations at the Three Key Residues for Heme Uptake. *Biochemistry*. 2003;42:10627-33.

33. Letoffe S, Deniau C, Wolff N, Dassa E, Delepelaire P, Lecroisey A, Wandersman C. Haemophore-mediated bacterial haem transport: evidence for a common or overlapping site for haem-free and haem-loaded haemophore on its specific outer membrane receptor. *Mol Microbiol*. 2001 Jul;41(2):439-50.

34. Kabsch W. Xds. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):125-32.

35. Evans P. Scaling and assessment of data quality. *Acta crystallographica Section D, Biological crystallography*. 2006 Jan;62(Pt 1):72-82.

36. Kabsch W. Automatic-Indexing of Rotation Diffraction Patterns. *J Appl Crystallogr*. 1988 Feb 1;21:67-71.

37. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr*. 1997 Dec 1;30:1022-5.

38. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb;66(Pt 2):213-21.

39. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography*. 2004 Dec;60(Pt 12 Pt 1):2126-32.

40. Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. 2003 Feb 15;50(3):437-50.

41. Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr*. 2004 Dec;60(Pt 12 Pt 1):2288-94.

42. Evans PR. An Introduction to Data Reduction: Space-Group Determination, scaling and intentisy statistics. *Acta Cryst*. 2011;D67:282-92.

43. Diederichs K, Karplus PA. Improved R-factors for Diffraction Data Analysis in Macromolecular Crystallography. *Nat Struct Biol*. 1997;4:269-75.

44. Weiss MS. Global indicators of X-ray data quality. *J Appl Cryst*. 2001;34:130-5.

45. Krissinel E, Henrick K. Secondary Structure Matching (SSM), A New tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Cryst*. 2004;D60:2256-68.

46. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS,

Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography*. 2011 Apr;67(Pt 4):235-42.

47. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moënne-Loccoz P. Kinetic and Spectroscopic Studies of Hemin Acquisition in the Hemophore HasAp from Pseudomonas aeruginosa. *Biochemistry*. 2010 2010/08/10;49(31):6646-54.

48. Chuang WJ, Johnson S, Van Wart HE. Resonance Raman spectra of bovine liver catalase: enhancement of proximal tyrosinate vibrations. *Journal of inorganic biochemistry*. 1988 Nov;34(3):201-19.

49. Sharma KD, Andersson LA, Loehr TM, Terner J, Goff HM. Comparative spectral analysis of mammalian, fungal, and bacterial catalases. Resonance Raman evidence for iron-tyrosinate coordination. *J Biol Chem*. 1989 Aug 5;264(22):12772-9.

50. Rohlfs RJ, Mathews AJ, Carver TE, Olson JS, Springer BA, Egeberg KD, Sligar SG. The Effects of Amino Acid Substitution at Position E7 (Residue 64) on the Kinetics of Ligand Binding to Sperm Whale Myoglobin. *J Biol Chem*. 1990;256:3168-76.

51. Adachi S, Nagano S, Ishimori K, Watanabe Y, Morishima I, Egawa T, Kitagawa T, Makino R. Roles of proximal ligand in heme proteins: replacement of proximal histidine of human myoglobin with cysteine and tyrosine by site-directed mutagenesis as models for P-450, chloroperoxidase, and catalase. *Biochemistry*. 1993 Jan 12;32(1):241-52.

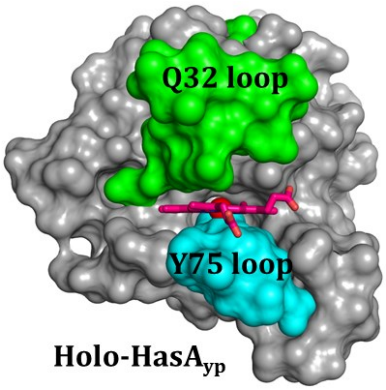
52. Liu D, Williamson DA, Kennedy ML, Williams TD, Morton MM, Benson DR. Aromatic Side Chain-Porphyrin Interactions in Designed Hemoproteins. *J Am Chem Soc*. 1999;121:11798-812.

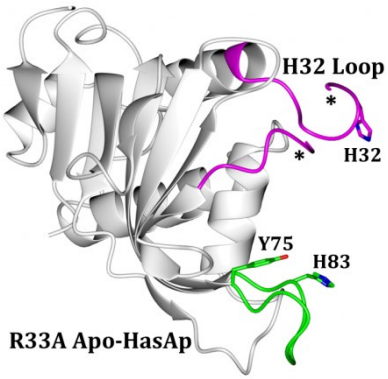
53. Uchida T, Stevens JM, Daltrop O, Harvat EM, Hong L, Ferguson SJ, Kitagawa T. The interaction of covalently bound heme with the cytochrome c maturation protein CcmE. *J Biol Chem*. 2004 Dec 10;279(50):51981-8.

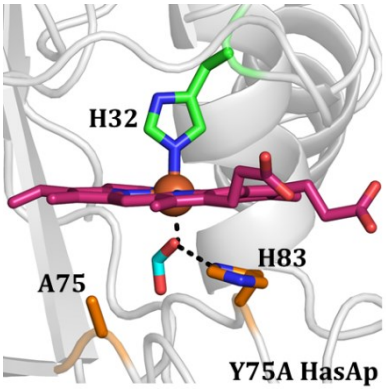
54. Eakanunkul S, Lukat-Rodgers GS, Sumithran S, Ghosh A, Rodgers KR, Dawson JH, Wilks A. Characterization of the periplasmic heme-binding protein shut from the heme uptake system of Shigella dysenteriae. *Biochemistry*. 2005 Oct 4;44(39):13179-91.

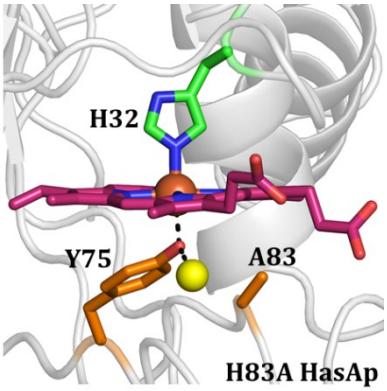
Chapter 5

Summary









5.1 Epilogue

The nutritional requirement of iron in opportunistic pathogens has been established for growth and successful colonization in vertebrate hosts [1]. However, the extremely low availability of free iron in the host cells poses a significant challenge for bacteria to fulfill their metabolic needs. To combat the challenging environment, bacteria have evolved sophisticated mechanisms to scavenge iron and circumvent host iron sequestration mechanisms.

One such sophisticated mechanism is by deploying the *has* (heme acquisition system) system comprised of a hemophore (HasA, heme binding protein) and an outer membrane receptor HasR [1]. Hemophores have been structurally and biochemically characterized from *Pseudomonas aeruginosa* (HasAp), *Yersinia pestis* ($HasA_{yp}$) and *Serratia marcescens* (HasAs). Under iron limiting conditions, hemophores are secreted in the host which binds heme and shuttles it to the outer membrane receptor for subsequent internalization. Once in the cytoplasm, the heme macrocycle is broken down by the heme oxygenase to release iron which is then either stored or used as an essential nutrient [2].

Because of the fact that iron is required for growth and pathogenicity [3], infections caused by bacteria could be minimized or even stopped at an early stage by limiting or depriving the organism of iron. To limit the progression of infection, it is essential to understand the mechanism of iron/heme uptake and the structural-functional relationship of various proteins that are activated during the process. The insights from the structural aspects will help to better understand and design inhibitors/drugs which would specifically target and inhibit the heme acquisition from the host cells, thus curbing the progression of disease.

In order to better understand the structural and functional relationship, we carried out our investigations by characterizing a novel hemophore structure from *Yersinia pestis* using analytical tools and had the opportunity to discover new findings about this protein. Further, new structural features are identified by replacing a single residue Arg33 to Alanine in HasA from *Pseudomonas aeruginosa*. The findings and discoveries from experimental data are summarized below.

5.1.1 Heme Binding Protein from Yersinia pestis ($HasA_{yp}$) was Over-expressed, Purified and Crystallized

A synthetic gene encoding for a full-length $HasA_{yp}$ protein was cloned in *E. coli*. The protein was overexpressed and purified using a previously developed protocol [4] with modifications. Purification of full-length protein resulted in mixture of apo- and holo-protein truncated at C-terminus. Following this, a gene encoding for the truncated $HasA_{yp}$ (full-length – 12) residues was constructed by placing a stop codon after Val193. The truncated $HasA_{yp}$ was purified to homogeneity was analyzed by electrospray ionization mass spectrometry (Chapter II).

Crystals were obtained for apo- and holo-$HasA_{yp}$ (full length - 12) from various conditions. The crystals for apo-form grew from seventy six unique conditions and for holo-form from eight unique conditions. The initial diffraction data was collected in-house and the model was built using various crystallographic software packages.

5.1.2 Novel Structure of the Apo- and Holo-$HasA_{yp}$ with Fast Kinetics of Heme Loading

The crystal structure of $HasA_{yp}$, (apo- and holo-form) shows a unique structure which is different from previously characterized hemophores from *P. aeruginosa* [4, 5] and *S. marcescens* [6, 7]. The heme ion ion is only coordinated by one protein provided ligand, Tyr75 from the

Y75 loop. The distal Q32 loop in the apo-$HasA_{yp}$ is already in closed conformation and does not coordinate heme iron ion upon heme loading. Kinetic analysis suggests that heme loading is fast compared to the HasAp and is completed within millisecond time scale with minimum structural reorganization.

5.2.1 Arg33 in Apo-HasAp Anchors the H32 Loop in Open Conformation

Differences in the distal loop (H32 in HasAp and Q32 in $HasA_{yp}$) organization lead to investigate the open H32 loop in HasAp. Detailed structural analysis of wt apo-HasAp protein points to a unique position of Arginine at position 33 networked with salt bridges and H-bond interactions. Arg33 was replaced with an Alanine at the same position with a hypothesis that it anchors the H32 loop in the open position.

R33A apo- and holo proteins were expressed, purified and screened for crystal growth. Apo-protein crystals were obtained after 43 days and holo-protein crystals were obtained after 72 hours (Chapter III).

5.2.2 Replacing Arg33 to Alanine in H32 Loop in Apo-HasAp Leads to Closure of the H32 Loop

The superimposed HSQC spectrum of wt apo- and R33A apo-protein shows completely different set of cross-peaks suggesting that H32 loop might be organized differently compared to the wt apo-protein, whereas upon heme reconstitution, the HSQC spectrum shows the cross peaks matched for the wt holo- and R33A holo-proteins. This observation was structurally supported by the crystal structure of R33A apo-protein which showed that the H32 loop adopts a

conformation which is very similar to the wt holo-HasAp. Moreover, the hinge regions in the H32 loop are already in a conformation which is superimposable with the conformation in the wt holo-HasAp. Further, results from molecular dynamics simulations provide additional insights and shows the loop is indeed adopts a closed form similar to the wt holo-HasAp. An interesting feature of the mutant holo-protein is the heme binding to the polypeptide. This is the first observation of dual binding of the heme molecule to mutant HasAp protein. From the available structural data, we conclude that a single Arginine at position 33 in the HasAp is the anchor residue holding the H32 loop in the open position and replacing it with Alanine shifts the loop in a conformation which is similar to the wt holo-protein.

5.3.1 Severe Changes in the Axial Ligands Minimally Affect Heme Acquisition in HasAp from P. aeruginosa

Hydrogen bond between Tyr75 and His83 in the Y75 loop is proposed to modulate the Tyr75-Fe^{3+} coordination [7]. To investigate the role played by the hydrogen bond in coordinating heme iron, Tyr75 and His83 were replaced by Alanine in two separate sets of experiments. Mutated proteins were purified, heme reconstituted and set up for crystal screening. For the apo-proteins solution NMR experiments were carried out to probe disorder in the Y75 loop.

5.3.2 Overall the Structure of Mutant Holo-Proteins is Similar to the Wt Holo-HasAp

Despite the severe changes introduced in the Y75 loop by replacing Tyr75 and His83 to alanine, the mutant holo-proteins show a high degree of similarity to the wt holo-protein. High resolution crystal structures obtained for Y75A do not show His83 coordinating the hemin-iron

from the Y75 loop, instead a formate ion was modeled which occupies a similar position as of Tyr75 in the wt protein. The plane of formate ion is aligned with the α-γ-meso axis of heme which allows the oxygen atom from formate ion to coordinate the heme iron ion. Similarly, all the three high resolution crystal structures obtained for the H83A mutants shows that the heme iron is coordinated by both His32 and Tyr75 and the structure of Y75 loop is almost identical to that of wt holo-HasAp. In the absence of His83, in two of the structures ($H83A^{ortho}$ and $H83A^{pH5.4}$), a water molecule occupies the site which would be normally be occupied by the side chain of His83 in the wt holo-protein. This crystallographic water molecule is hydrogen bonded to the phenol oxygen of the coordinated Tyr75. In the $H83A^{mono}$ structure, the same position is occupied by an ethylene glycol molecule that acts as a hydrogen bond partner to the coordinated Tyr75.

5.3.3 Y75 Loop is Disorder in the Mutant Apo-Proteins

Sequential backbone assignments of the wt, Y75A and H83A HasAp were obtained. The NMR spectra of wt apo-protein were consistent with a well-ordered structure which was also complemented with the high resolution X-ray crystal structure. The HSQC spectra of Y75A and H83A apo-proteins are nearly identical to the wt apo-protein, however, a closer inspection reveals several cross peaks, corresponding to residues in Y75 loop, are conformationally disordered in Y75A and H83A apo-protein. Moreover, the conformational disorder propagates to the nearby secondary structural elements like helix α2, and several β-strands. Replacing Tyr75 and His83 to alanine results in the loss of several stabilizing network of H-bonds and hydrophobic interaction, which otherwise kept the Y75 loop in a stable conformation in the wt apo-HasAp.

5.3.4 Heme Capture is Driven by Hydrophobic Interactions

Combining the structural and kinetic data suggests that heme capture is mainly driven by hydrophobic and π–π stacking interactions which are favored by the "sticky" hydrophobic platform provided by the Y75 loop while controlling its release via the heme-iron axial coordination.

5.2 Future Directions

In the current study, we were able to show our novel findings and important results understanding the hemophore structures and heme loading mechanisms in *Pseudomonas aeruginosa* and *Yersinia pestis*. Differences in the distal loop structure in $HasA_{yp}$ and HasAp would suggest differences in interactions and mechanisms of heme delivery to the cognate receptor on the bacterial outer membrane. One of the future directions would be the study of hemophores with their cognate receptor. Further, studies could be aimed at understanding the kinetics and thermodynamics of heme capture and release in variants like R33A and double mutants of Y75A/H83A, Y75F/H83A HasAp.

5.3 Bibliography

1. Vasil ML, Ochsner UA. The Response of *Pseudomonas aeruginosa* to Iron: Genetics, Biochemistry and Virulence. *Mol Microbiol*. 1999;34:399-413.
2. Benson DR, Rivera M. Heme Uptake and Metabolism in Bacteria. Met Ions Life Sci. 122013. p. 279-332.
3. Weinberg ED. Iron Availability and Infection. *Biochim et Biophys Acta*. 2009;1790:600-5.
4. Jepkorir G, Rodriguez JC, Rui H, Im W, Lovell S, Battaile KP, Alontaga AY, Yukl ET, Moenne-Loccoz P, Rivera M. Structural, NMR spectroscopic, and computational investigation of hemin loading in the hemophore HasAp from Pseudomonas aeruginosa. *Journal of the American Chemical Society*. 2010 Jul 21;132(28):9857-72.
5. Alontaga AY, Rodriguez JC, Schonbrunn E, Becker A, Funke T, Yukl ET, Hayashi T, Stobaugh J, Moenne-Loccoz P, Rivera M. Structural characterization of the hemophore HasAp from Pseudomonas aeruginosa: NMR spectroscopy reveals protein-protein interactions between Holo-HasAp and hemoglobin. *Biochemistry*. 2009 Jan 13;48(1):96-109.
6. Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. The crystal structure of HasA, a hemophore secreted by Serratia marcescens. *Nature structural biology*. 1999 Jun;6(6):516-20.
7. Wolff N, Izadi-Pruneyre N, Couprie J, Habeck M, Linge J, Rieping W, Wandersman C, Nilges M, Delepierre M, Lecroisey A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*. 2008 Feb 15;376(2):517-25.

List of Publications

1. **Ritesh Kumar** and Mario Rivera (**2014**), Replacing Arginine 33 for Alanine in the Hemophore HasA from *Pseudomonas aeruginosa* Causes Closure of the H32 Loop in the Apo-Protein. *Manuscript in preparation*.

2. **Ritesh Kumar**, Hirotoshi Matsumura, Scott Lovell, Huili Yao, Juan C. Rodriguez, Kevin P. Battaile, Pierre Moënne-Loccoz, and Mario Rivera (**2014**), Replacing the Axial Ligand Tyrosine 75 or its Hydrogen Bond Partner Histidine 83 Minimally Affects Hemin Acquisition by the Hemophore HasAp from *Pseudomonas aeruginosa, Biochemistry*, 53(13), 2112-25.

3. **Ritesh Kumar**, Scott Lovell, Hirotoshi Matsumura, Kevin P. Battaile, Pierre Moënne-Loccoz, and Mario Rivera (**2013**), The hemophore HasA from *Yersinia pestis* ($HasA_{yp}$) coordinates hemin with a single residue, Tyr75, and with minimal conformational change, *Biochemistry,* Apr 23; 52(16), 2705-7.

4. Huili Yao, Yan Wang, Scott Lovell, **Ritesh Kumar**, Anatoly M. Ruvinsky, Kevin P. Battaile, Ilya A. Vakser, and Mario Rivera (**2012**), The structure of the BfrB-Bfd complex reveals protein-protein interactions enabling iron release from bacterioferritin, *J Am Chem Soc.,* Aug 15;134(32), 13470-81.

5. Huan Rui, **Ritesh Kumar** and Wonpil Im (**2011**), Membrane tension, lipid adaptation, conformational changes, and energetics in MscL gating, *Biophysical Journal*, Aug 3; 101(3), 671-9.

Appendices

Appendix I: Index HT protein crystallization screen (Hampton Research)

A1. 0.1 M Citric acid pH 3.5, 2.0 M Ammonium sulfate
A2. 0.1 M Sodium acetate trihydrate pH 4.5, 2.0 M Ammonium sulfate
A3. 0.1 M BIS-TRIS pH 5.5, 2.0 M Ammonium sulfate
A4. 0.1 M BIS-TRIS pH 6.5, 2.0 M Ammonium sulfate
A5. 0.1 M HEPES pH 7.5, 2.0 M Ammonium sulfate
A6. 0.1 M Tris pH 8.5, 2.0 M Ammonium sulfate
A7. 0.1 M Citric acid pH 3.5, 3.0 M Sodium chloride
A8. 0.1 M Sodium acetate trihydrate pH 4.5, 3.0 M Sodium chloride
A9. 0.1 M BIS-TRIS pH 5.5, 3.0 M Sodium chloride
A10. 0.1 M BIS-TRIS pH 6.5, 3.0 M Sodium chloride
A11. 0.1 M HEPES pH 7.5, 3.0 M Sodium chloride
A12. 0.1 M Tris pH 8.5, 3.0 M Sodium chloride
B1. 0.1 M BIS-TRIS pH 5.5, 0.3 M Magnesium formate dihydrate
B2. 0.1 M BIS-TRIS pH 6.5, 0.5 M Magnesium formate dihydrate
B3. 0.1 M HEPES pH 7.5, 0.5 M Magnesium formate dihydrate
B4. 0.1 M Tris pH 8.5, 0.3 M Magnesium formate dihydrate
B5. 1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 5.6
B6. 1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 6.9
B7. 1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 8.2
B8. 0.1 M HEPES pH 7.5, 1.4 M Sodium citrate tribasic dihydrate
B9. 1.8 M Ammonium citrate tribasic pH 7.0
B10. 0.8 M Succinic acid pH 7.0
B11. 2.1 M DL-Malic acid pH 7.0
B12. 2.8 M Sodium acetate trihydrate pH 7.0
C1. 3.5 M Sodium formate pH 7.0
C2. 1.1 M Ammonium tartrate dibasic pH 7.0
C3. 2.4 M Sodium malonate pH 7.0
C4. 35% v/v Tacsimate pH 7.0
C5. 60% v/v Tacsimate pH 7.0
C6. 0.1 M Sodium chloride, 0.1 M BIS-TRIS pH 6.5, 1.5 M Ammonium sulfate
C7. 0.8 M Potassium sodium tartrate tetrahydrate, 0.1 M Tris pH 8.5, 0.5% w/v Polyethylene glycol monomethyl ether 5,000
C8. 1.0 M Ammonium sulfate, 0.1 M BIS-TRIS pH 5.5, 1% w/v Polyethylene glycol 3,350
C9. 1.1 M Sodium malonate pH 7.0, 0.1 M HEPES pH 7.0, 0.5% v/v Jeffamine ED-2001 pH 7.0
C10. 1.0 M Succinic acid pH 7.0, 0.1 M HEPES pH 7.0, 1% w/v Polyethylene glycol monomethyl ether 2,000
C11. 1.0 M Ammonium sulfate , 0.1 M HEPES pH 7.0, 0.5% w/v Polyethylene glycol 8,000
C12. 15% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 2% w/v Polyethylene glycol 3,350
D1. 25% w/v Polyethylene glycol 1,500
D2. 0.1 M HEPES pH 7.0, 30% v/v Jeffamine M-600 pH 7.0
D3. 0.1 M HEPES pH 7.0, 30% v/v Jeffamine ED-2001 pH 7.0
D4. 0.1 M Citric acid pH 3.5, 25% w/v Polyethylene glycol 3,350
D5. 0.1 M Sodium acetate trihydrate pH 4.5, 25% w/v Polyethylene glycol 3,350
D6. 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
D7. 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
D8. 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350

D9. 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
D10. 0.1 M BIS-TRIS pH 6.5, 20% w/v Polyethylene glycol monomethyl ether 5,000
D11. 0.1 M BIS-TRIS pH 6.5, 28% w/v Polyethylene glycol monomethyl ether 2,000
D12. 0.2 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 5.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E1. 0.2 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 6.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E2. 0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E3. 0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 6.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E4. 0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E5. 0.2 M Ammonium acetate, 0.1 M Tris pH 8.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
E6. 0.05 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 6.5, 30% v/v Polyethylene glycol monomethyl ether 550
E7. 0.05 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% v/v Polyethylene glycol monomethyl ether 550
E8. 0.2 M Potassium chloride, 0.05 M HEPES pH 7.5, 35% v/v Pentaerythritol propoxylate (5/4 PO/OH)
E9. 0.05 M Ammonium sulfate , 0.05 M BIS-TRIS pH 6.5, 30% v/v Pentaerythritol ethoxylate (15/4 EO/OH)
E10. 0.1 M BIS-TRIS pH 6.5, 45% v/v Polypropylene glycol P 400
E11. 0.02 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 22% w/v Polyacrylic acid sodium salt 5,100
E12. 0.01 M Cobalt(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 20% w/v Polyvinylpyrrolidone K 15
F1. 0.2 M L-Proline, 0.1 M HEPES pH 7.5, 10% w/v Polyethylene glycol 3,350
F2. 0.2 M Trimethylamine N-oxide dihydrate, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000
F3. 5% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 10% w/v Polyethylene glycol monomethyl ether 5,000
F4. 0.005 M Cobalt(II) chloride hexahydrate, 0.005 M Nickel(II) chloride hexahydrate, 0.005 M Cadmium chloride hydrate, 0.005 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 12% w/v Polyethylene glycol 3,350
F5. 0.1 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 17% w/v Polyethylene glycol 10,000
F6. 0.2 M Ammonium sulfate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
F7. 0.2 M Ammonium sulfate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
F8. 0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
F9. 0.2 M Ammonium sulfate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
F10. 0.2 M Sodium chloride, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
F11. 0.2 M Sodium chloride, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
F12. 0.2 M Sodium chloride, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
G1. 0.2 M Sodium chloride, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
G2. 0.2 M Lithium sulfate monohydrate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
G3. 0.2 M Lithium sulfate monohydrate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
G4. 0.2 M Lithium sulfate monohydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
G5. 0.2 M Lithium sulfate monohydrate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
G6. 0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
G7. 0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
G8. 0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
G9. 0.2 M Ammonium acetate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
G10. 0.2 M Magnesium chloride hexahydrate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
G11. 0.2 M Magnesium chloride hexahydrate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
G12. 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
H1. 0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
H2. 0.2 M Potassium sodium tartrate tetrahydrate, 20% w/v Polyethylene glycol 3,350
H3. 0.2 M Sodium malonate pH 7.0, 20% w/v Polyethylene glycol 3,350
H4. 0.2 M Ammonium citrate tribasic pH 7.0, 20% w/v Polyethylene glycol 3,350
H5. 0.1 M Succinic acid pH 7.0, 15% w/v Polyethylene glycol 3,350
H6. 0.2 M Sodium formate, 20% w/v Polyethylene glycol 3,350
H7. 0.15 M DL-Malic acid pH 7.0, 20% w/v Polyethylene glycol 3,350
H8. 0.1 M Magnesium formate dihydrate, 15% w/v Polyethylene glycol 3,350
H9. 0.05 M Zinc acetate dihydrate, 20% w/v Polyethylene glycol 3,350

H10. 0.2 M Sodium citrate tribasic dihydrate, 20% w/v Polyethylene glycol 3,350
H11. 0.1 M Potassium thiocyanate, 30% w/v Polyethylene glycol monomethyl ether 2,000
H12. 0.15 M Potassium bromide, 30% w/v Polyethylene glycol monomethyl ether 2,000

Appendix II: pH buffer protein crystallization screen (Rigaku)

1M Sodium Lactate / HCl	A1	2.4	B1	2.8	C1	3.2	D1	3.6	E1	4.0	F1	4.4	G1	4.8	H1	5.2
1M Sodium Acetate / Acetic Acid	A2	3.4	B2	3.8	C2	4.2	D2	4.6	E2	5.0	F2	5.4	G2	5.8	H2	6.2
1M MES / NaOH	A3	4.6	B3	5.0	C3	5.4	D3	5.8	E3	6.2	F3	6.6	G3	7.0	H3	7.4
1M Bis-Tris / HCl	A4	5.2	B4	5.6	C4	6.0	D4	6.4	E4	6.8	F4	7.2	G4	7.6	H4	8.0
1M Imidazole / HCl	A5	5.4	B5	5.8	C5	6.2	D5	6.6	E5	7.0	F5	7.4	G5	7.8	H5	8.2
1M K2HPO4/ NaH2PO4	A6	5.8	B6	6.2	C6	6.6	D6	7.0	E6	7.4	F6	7.8	G6	8.2	H6	8.6
1M HEPES / NaOH	A7	6.0	B7	6.4	C7	6.8	D7	7.2	E7	7.6	F7	8.0	G7	8.4	H7	8.8
1M Tris / HCl	A8	6.6	B8	7.0	C8	7.4	D8	7.8	E8	8.2	F8	8.6	G8	9.0	H8	9.4
1M Bicine / NaOH	A9	7.0	B9	7.4	C9	7.8	D9	8.2	E9	8.6	F9	9.0	G9	9.4	H9	9.8
0.5M CHES / NaOH	A10	8.0	B10	8.4	C10	8.8	D10	9.2	E10	9.6	F10	10.0	G10	10.4	H10	10.8
1M Glycine / NaOH	A11	8.2	B11	8.6	C11	9.0	D11	9.4	E11	9.8	F11	10.2	G11	10.6	H11	11.0
0.5M CAPS / NaOH	A12	8.8	B12	9.2	C12	9.6	D12	10.0	E12	10.4	F12	10.8	G12	11.2	H12	11.6

All formulations are made with ultrapure ASTM Type I water and sterile-filtered stock solutions. Store at 4-25 °C

Appendix III: Crystal screen (Hampton Research)

1. 0.02 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
2. 0.4 M Potassium sodium tartrate tetrahydrate
3. 0.4 M Ammonium phosphate monobasic
4. 0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium sulfate
5. 0.2 M Sodium citrate tribasic dihydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
6. 0.2 M Magnesium chloride hexahydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000
7. 0.1 M Sodium cacodylate trihydrate pH 6.5, 1.4 M Sodium acetate trihydrate
8. 0.2 M Sodium citrate tribasic dihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% v/v 2-Propanol
9. 0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 30% w/v Polyethylene glycol 4,000
10. 0.2 M Ammonium acetate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol 4,000

11. 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Ammonium phosphate monobasic
12. 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v 2-Propanol
13. 0.2 M Sodium citrate tribasic dihydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% v/v Polyethylene glycol 400
14. 0.2 M Calcium chloride dihydrate, 0.1 M HEPES sodium pH 7.5, 28% v/v Polyethylene glycol 400
15. 0.2 M Ammonium sulfate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% w/v Polyethylene glycol 8,000
16. 0.1 M HEPES sodium pH 7.5, 1.5 M Lithium sulfate monohydrate
17. 0.2 M Lithium sulfate monohydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000
18. 0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 20% w/v Polyethylene glycol 8,000
19. 0.2 M Ammonium acetate, 0.1 M TRIS hydrochloride pH 8.5, 30% v/v 2-Propanol
20. 0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 25% w/v Polyethylene glycol 4,000
21. 0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
22. 0.2 M Sodium acetate trihydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000
23. 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v Polyethylene glycol 400
24. 0.2 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 20% v/v 2-Propanol
25. 0.1 M Imidazole pH 6.5, 1.0 M Sodium acetate trihydrate
26. 0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
27. 0.2 M Sodium citrate tribasic dihydrate, 0.1 M HEPES sodium pH 7.5, 20% v/v 2-Propanol
28. 0.2 M Sodium acetate trihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% w/v Polyethylene glycol 8,000
29. 0.1 M HEPES sodium pH 7.5, 0.8 M Potassium sodium tartrate tetrahydrate
30. 0.2 M Ammonium sulfate, 30% w/v Polyethylene glycol 8,000
31. 0.2 M Ammonium sulfate, 30% w/v Polyethylene glycol 4,000
32. 2.0 M Ammonium sulfate
33. 4.0 M Sodium formate
34. 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium formate
35. 0.1 M HEPES sodium pH 7.5, 0.8 M Sodium phosphate monobasic monohydrate, 0.8 M Potassium phosphate monobasic
36. 0.1 M TRIS hydrochloride pH 8.5, 8% w/v Polyethylene glycol 8,000
37. 0.1 M Sodium acetate trihydrate pH 4.6, 8% w/v Polyethylene glycol 4,000
38. 0.1 M HEPES sodium pH 7.5, 1.4 M Sodium citrate tribasic dehydrate
39. 0.1 M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400, 2.0 M Ammonium sulfate
40 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 20% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000
41. 0.1 M HEPES sodium pH 7.5, 10% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000
42. 0.05 M Potassium phosphate monobasic, 20% w/v Polyethylene glycol 8,000
43. 30% w/v Polyethylene glycol 1,500
44. 0.2 M Magnesium formate dehydrate
45. 0.2 M Zinc acetate dihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 18% w/v Polyethylene glycol 8,000
46. 0.2 M Calcium acetate hydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5 , 18% w/v Polyethylene glycol 8,000
47. 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Ammonium sulfate
48. 0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium phosphate monobasic
49. 1.0 M Lithium sulfate monohydrate, 2% w/v Polyethylene glycol 8,000
50. 0.5 M Lithium sulfate monohydrate, 15% w/v Polyethylene glycol 8,000

Appendix IV: Wizard IV protein crystallization screen (Rigaku)

Tube	Precipitation Reagent	Buffer	Salt	Detergent
1	16% (w/v) PEG 8000		40 mM Potassium phosphate dibasic	20% (v/v) Glycerol
2	5% (v/v) MPD	100 mM Tris base/ Hydrochloric acid pH 8.0	100 mM Sodium chloride	15% (v/v) Reagent alcohol
3	5% (w/v) PEG 1000	100 mM Sodium phosphate dibasic / Citric acid pH 4.2		40% (v/v) Reagent alcohol
4		100 mM Bis Tris/ Hydrochloric acid pH 5.5	200 mM Ammonium sulfate	
5	2% (v/v) PEG 400	100 mM Sodium acetate/ Acetic acid pH 5.5	2000 mM Ammonium sulfate	
6		100 mM Sodium citrate/ Citric acid pH 4.0	800 mM Ammonium sulfate	
7	2000 mM Lithium sulfate	100 mM Sodium acetate/ Acetic acid pH 4.5	100 mM Magnesium sulfate	5% (v/v) 2-Propanol
8	2% (v/v) PEG 400	100 mM Tris base/ Hydrochloric acid pH 8.5	2000 mM Lithium sulfate	
9	5% (v/v) PEG 400	100 mM Sodium acetate/ Acetic acid pH 5.5	2000 mM Lithium sulfate	100 mM Magnesium sulfate
10	50% (v/v) PEG 200	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Magnesium chloride	
11	40% (v/v) PEG 300	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Calcium acetate	
12	30% (v/v) Jeffamine M-600 pH 7.0	100 mM HEPES/ Sodium hydroxide pH 7.0		
13	800 mM Succinic acid pH 7.0			
14	40% (v/v) PEG 400	100 mM Tris base/ Hydrochloric acid pH 8.5	200 mM Lithium sulfate	
15	50% (v/v) PEG 400	100 mM Sodium acetate/ Acetic acid pH 4.5	200 mM Lithium sulfate	
16	15% (v/v) PEG 550 MME	100 mM MES/ Sodium hydroxide pH 6.5		
17	25% (w/v) PEG 1500	100 mM SPG buffer pH 5.5		
18	25% (w/v) PEG 1500	100 mM SPG buffer pH 8.5		
19	25% (w/v) PEG 1500	100 mM MMT buffer pH 6.5		
20	25% (w/v) PEG 1500	100 mM MMT buffer pH 9.0		
21	25% (w/v) PEG 1500	100 mM MIB buffer pH 5.0		
22	25% (w/v) PEG 1500	100 mM PCB buffer pH 7.0		

23	12% (w/v) PEG 1500	100 mM Sodium acetate/ Acetic acid pH 5.5	2500 mM Sodium chloride	1.5% (v/v) MPD
24	2400 mM Sodium malonate dibasic			
25	30% (w/v) PEG 2000 MME		150 mM Potassium bromide	
26	10% (w/v) PEG 2000 MME	100 mM Sodium acetate/ Acetic acid pH 5.5	200 mM Ammonium sulfate	
27	20% (w/v) PEG 2000 MME	100 mM Tris base/ Hydrochloric acid pH 8.5	200 mM Trimethylamine n-oxide	
28	20% (w/v) PEG 3350	100 mM Bis tris propane/ Hydrochloric acid pH 6.5	200 mM Sodium fluoride	
29	20% (w/v) PEG 3350	100 mM Sodium citrate/ Citric acid pH 4.0	200 mM Sodium citrate tribasic	
30	20% (w/v) PEG 3350	100 mM Bis tris propane/ Hydrochloric acid pH 8.5	200 mM Sodium malonate dibasic	
31	20% (w/v) Polyacrylic acid 5100	100 mM HEPES/ Sodium hydroxide pH 7.0	20 mM Magnesium chloride	
32	2100 mM DL Malic acid pH 7.0			
33	800 mM Potassium phosphate dibasic	100 mM HEPES/ Sodium hydroxide pH 7.5	800 mM Sodium phosphate monobasic	
34	20% (w/v) PEG 6000	100 mM MES/ Sodium hydroxide pH 6.0	200 mM Ammonium chloride	
35	20% (w/v) PEG 6000	100 mM HEPES/ Sodium hydroxide pH 7.0	200 mM Sodium chloride	
36	20% (w/v) PEG 6000	100 mM Tris base/ Hydrochloric acid pH 8.0	200 mM Lithium chloride	
37	20% (w/v) Polyvinylpyrrolidone K15	100 mM Tris base/ Hydrochloric acid pH 8.5	100 mM Cobalt chloride	
38	50% (v/v) Ethylene glycol	100 mM Tris base/ Hydrochloric acid pH 8.5	200 mM Magnesium chloride	
39	20% (w/v) PEG 8000	100 mM Imidazole/ Hydrochloric acid pH 6.5		3% (v/v) MPD
40	20% (w/v) PEG 8000	100 mM Tris base/ Hydrochloric acid pH 8.5	100 mM Magnesium chloride	20% (v/v) PEG 400
41	20% (w/v) PEG 8000	100 mM HEPES/ Sodium hydroxide pH 7.5	200 mM Ammonium sulfate	10% (v/v) 2-propanol
42	30% (v/v) MPD	100 mM Sodium acetate/ Acetic acid pH 4.5		25% (w/v) PEG 1500
43	30% (v/v) MPD	100 mM Imidazole/ Hydrochloric acid pH 6.5	200 mM Ammonium sulfate	10% (w/v) PEG 3350
44	30% (v/v) MPD	100 mM Tris base/ Hydrochloric acid pH 8.5	500 mM Sodium chloride	8% (w/v) PEG 8000
45	40% (v/v) 2-propanol	100 mM Imidazole/ Hydrochloric acid pH 6.5		15% (w/v) PEG 8000

46	30% (v/v) 2-propanol	100 mM Tris base/ Hydrochloric acid pH 8.5		30% (w/v) PEG 3350
47	17% (w/v) PEG 10,000	100 mM Bis tris/ Hydrochloric acid pH 5.5	100 mM Ammonium acetate	
48	15% (w/v) PEG 20,000	100 mM HEPES/ Sodium hydroxide pH 7.0		

Appendix V: Crystal screen 2 crystallization screen (Hampton Research)

1. 2.0 M Sodium chloride, 10% w/v Polyethylene glycol 6,000
2. 0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate, 0.01 M Hexadecyltrimethylammonium bromide
3. 25% v/v Ethylene glycol
4. 35% v/v 1,4-Dioxane
5. 2.0 M Ammonium sulfate, 5% v/v 2-Propanol
6. 1.0 M Imidazole pH 7.0
7. 10% w/v Polyethylene glycol 1,000, 10% w/v Polyethylene glycol 8,000
8. 1.5 M Sodium chloride, 10% v/v Ethanol
9. 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium chloride
10. 0.2 M Sodium chloride, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
11. 0.01 M Cobalt(II) chloride hexahydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 1.0 M 1,6-Hexanediol
12. 0.1 M Cadmium chloride hydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v Polyethylene glycol 400
13. 0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol monomethyl ether 2,000
14. 0.2 M Potassium sodium tartrate tetrahydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2.0 M Ammonium sulfate
15. 0.5 M Ammonium sulfate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Lithium sulfate monohydrate
16. 0.5 M Sodium chloride, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2% v/v Ethylene imine polymer
17. 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 35% v/v tert-Butanol
18. 0.01 M Iron(III) chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 10% v/v Jeffamine M-600
19. 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2.5 M 1,6-Hexanediol
20. 0.1 M MES monohydrate pH 6.5, 1.6 M Magnesium sulfate heptahydrate
21. 0.1 M Sodium phosphate monobasic monohydrate, 0.1 M Potassium phosphate monobasic, 0.1 M MES monohydrate pH 6.5, 2.0 M Sodium chloride
22. 0.1 M MES monohydrate pH 6.5, 12% w/v Polyethylene glycol 20,000
23. 1.6 M Ammonium sulfate, 0.1 M MES monohydrate pH 6.5, 10% v/v 1,4-Dioxane
24. 0.05 M Cesium chloride, 0.1 M MES monohydrate pH 6.5, 30% v/v Jeffamine M-600
25. 0.01 M Cobalt(II) chloride hexahydrate, 0.1 M MES monohydrate pH 6.5, 1.8 M Ammonium sulfate
26. 0.2 M Ammonium sulfate, 0.1 M MES monohydrate pH 6.5, 30% w/v Polyethylene glycol monomethyl ether 5,000
27. 0.01 M Zinc sulfate heptahydrate, 0.1 M MES monohydrate pH 6.5, 25% v/v Polyethylene glycol monomethyl ether 550
28. 1.6 M Sodium citrate tribasic dihydrate pH 6.5
29. 0.5 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
30. 0.1 M HEPES pH 7.5, 10% w/v Polyethylene glycol 6,000, 5% v/v (+/-)-2-Methyl-2,4-pentanediol

31. 0.1 M HEPES pH 7.5, 20% v/v Jeffamine M-600
32. 0.1 M Sodium chloride, 0.1 M HEPES pH 7.5, 1.6 M Ammonium sulfate
33. 0.1 M HEPES pH 7.5, 2.0 M Ammonium formate
34. 0.05 M Cadmium sulfate hydrate, 0.1 M HEPES pH 7.5, 1.0 M Sodium acetate trihydrate
35. 0.1 M HEPES pH 7.5, 70% v/v (+/-)-2-Methyl-2,4-pentanediol
36. 0.1 M HEPES pH 7.5, 4.3 M Sodium chloride
37. 0.1 M HEPES pH 7.5, 8% v/v Ethylene glycol, 10% w/v Polyethylene glycol 8,000
38. 0.1 M HEPES pH 7.5, 20% w/v Polyethylene glycol 10,000
39. 0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 3.4 M 1,6-Hexanediol
40 0.1 M Tris pH 8.5, 25% v/v tert-Butanol
41. 0.01 M Nickel(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 1.0 M Lithium sulfate monohydrate
42. 1.5 M Ammonium sulfate, 0.1 M Tris pH 8.5, 12% v/v Glycerol
43. 0.2 M Ammonium phosphate monobasic, 0.1 M Tris pH 8.5, 50% v/v (+/-)-2-Methyl-2,4-pentanediol
44. 0.1 M Tris pH 8.5, 20% v/v Ethanol
45. 0.01 M Nickel(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000
46. 0.1 M Sodium chloride, 0.1 M BICINE pH 9.0, 20% v/v Polyethylene glycol monomethyl ether 550
47. 0.1 M BICINE pH 9.0, 2.0 M Magnesium chloride hexahydrate
48. 0.1 M BICINE pH 9.0, 2% v/v 1,4-Dioxane, 10% w/v Polyethylene glycol 20,000

Appendix VI: Crystal screen 1 crystallization screen (Hampton Research)

A1 0.02 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
A2 0.4 M Potassium sodium tartrate tetrahydrate
A3 0.4 M Ammonium phosphate monobasic
A4 0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium sulfate
A5 0.2 M Sodium citrate tribasic dihydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
A6 0.2 M Magnesium chloride hexahydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000
A7 0.1 M Sodium cacodylate trihydrate pH 6.5, 1.4 M Sodium acetate trihydrate
A8 0.2 M Sodium citrate tribasic dihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% v/v 2-Propanol
A9 0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 30% w/v Polyethylene glycol 4,000
A10 0.2 M Ammonium acetate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol 4,000
A11 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Ammonium phosphate monobasic
A12 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v 2-Propanol
B1 0.2 M Sodium citrate tribasic dihydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% v/v Polyethylene glycol 400
B2 0.2 M Calcium chloride dihydrate, 0.1 M HEPES sodium pH 7.5, 28% v/v Polyethylene glycol 400
B3 0.2 M Ammonium sulfate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% w/v Polyethylene glycol 8,000
B4 0.1 M HEPES sodium pH 7.5, 1.5 M Lithium sulfate monohydrate
B5 0.2 M Lithium sulfate monohydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000
B6 0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 20% w/v Polyethylene glycol 8,000
B7 0.2 M Ammonium acetate, 0.1 M TRIS hydrochloride pH 8.5, 30% v/v 2-Propanol
B8 0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 25% w/v Polyethylene glycol 4,000

B9 0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol

B10 0.2 M Sodium acetate trihydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4,000

B11 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v Polyethylene glycol 400

B12 0.2 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 20% v/v 2-Propanol

C1 0.1 M Imidazole pH 6.5, 1.0 M Sodium acetate trihydrate

C2 0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol

C3 0.2 M Sodium citrate tribasic dihydrate, 0.1 M HEPES sodium pH 7.5, 20% v/v 2-Propanol

C4 0.2 M Sodium acetate trihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% w/v Polyethylene glycol 8,000

C5 0.1 M HEPES sodium pH 7.5, 0.8 M Potassium sodium tartrate tetrahydrate

C6 0.2 M Ammonium sulfate, 30% w/v Polyethylene glycol 8,000

C7 0.2 M Ammonium sulfate, 30% w/v Polyethylene glycol 4,000

C8 2.0 M Ammonium sulfate

C9 4.0 M Sodium formate

C10 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium formate

C11 0.1 M HEPES sodium pH 7.5, 0.8 M Sodium phosphate monobasic monohydrate, 0.8 M Potassium phosphate monobasic

C12 0.1 M TRIS hydrochloride pH 8.5, 8% w/v Polyethylene glycol 8,000

D1 0.1 M Sodium acetate trihydrate pH 4.6, 8% w/v Polyethylene glycol 4,000

D2 0.1 M HEPES sodium pH 7.5, 1.4 M Sodium citrate tribasic dihydrate

D3 0.1 M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400, 2.0 M Ammonium sulfate

D4 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 20% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000

D5 0.1 M HEPES sodium pH 7.5, 10% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000

D6 0.05 M Potassium phosphate monobasic, 20% w/v Polyethylene glycol 8,000

D7 30% w/v Polyethylene glycol 1,500

D8 0.2 M Magnesium formate dihydrate

D9 0.2 M Zinc acetate dihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 18% w/v Polyethylene glycol 8,000

D10 0.2 M Calcium acetate hydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5 , 18% w/v Polyethylene glycol 8,000

D11 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Ammonium sulfate

D12 0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium phosphate monobasic

E1 2.0 M Sodium chloride, 10% w/v Polyethylene glycol 6,000

E2 0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate, 0.01 M Hexadecyltrimethylammonium bromide

E3 25% v/v Ethylene glycol

E4 35% v/v 1,4-Dioxane

E5 2.0 M Ammonium sulfate, 5% v/v 2-Propanol

E6 1.0 M Imidazole pH 7.0

E7 10% w/v Polyethylene glycol 1,000, 10% w/v Polyethylene glycol 8,000

E8 1.5 M Sodium chloride, 10% v/v Ethanol

E9 0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium chloride

E10 0.2 M Sodium chloride, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol

E11 0.01 M Cobalt(II) chloride hexahydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 1.0 M 1,6-Hexanediol

E12 0.1 M Cadmium chloride hydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v Polyethylene glycol 400

F1 0.2 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol monomethyl ether 2,000

F2 0.2 M Potassium sodium tartrate tetrahydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2.0 M

Ammonium sulfate

F3 0.5 M Ammonium sulfate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Lithium sulfate monohydrate

F4 0.5 M Sodium chloride, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2% v/v Ethylene imine polymer

F5 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 35% v/v tert-Butanol

F6 0.01 M Iron(III) chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 10% v/v Jeffamine M-600

F7 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 2.5 M 1,6-Hexanediol

F8 0.1 M MES monohydrate pH 6.5, 1.6 M Magnesium sulfate heptahydrate

F9 0.1 M Sodium phosphate monobasic monohydrate, 0.1 M Potassium phosphate monobasic, 0.1 M MES monohydrate pH 6.5, 2.0 M Sodium chloride

F10 0.1 M MES monohydrate pH 6.5, 12% w/v Polyethylene glycol 20,000

F11 1.6 M Ammonium sulfate, 0.1 M MES monohydrate pH 6.5, 10% v/v 1,4-Dioxane

F12 0.05 M Cesium chloride, 0.1 M MES monohydrate pH 6.5, 30% v/v Jeffamine M-600

G1 0.01 M Cobalt(II) chloride hexahydrate, 0.1 M MES monohydrate pH 6.5, 1.8 M Ammonium sulfate

G2 0.2 M Ammonium sulfate, 0.1 M MES monohydrate pH 6.5, 30% w/v Polyethylene glycol monomethyl ether 5,000

G3 0.01 M Zinc sulfate heptahydrate, 0.1 M MES monohydrate pH 6.5, 25% v/v Polyethylene glycol monomethyl ether 550

G4 1.6 M Sodium citrate tribasic dihydrate pH 6.5

G5 0.5 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol

G6 0.1 M HEPES pH 7.5, 10% w/v Polyethylene glycol 6,000, 5% v/v (+/-)-2-Methyl-2,4-pentanediol

G7 0.1 M HEPES pH 7.5, 20% v/v Jeffamine M-600

G8 0.1 M Sodium chloride, 0.1 M HEPES pH 7.5, 1.6 M Ammonium sulfate

G9 0.1 M HEPES pH 7.5, 2.0 M Ammonium formate

G10 0.05 M Cadmium sulfate hydrate, 0.1 M HEPES pH 7.5, 1.0 M Sodium acetate trihydrate

G11 0.1 M HEPES pH 7.5, 70% v/v (+/-)-2-Methyl-2,4-pentanediol

G12 0.1 M HEPES pH 7.5, 4.3 M Sodium chloride

H1 0.1 M HEPES pH 7.5, 8% v/v Ethylene glycol, 10% w/v Polyethylene glycol 8,000

H2 0.1 M HEPES pH 7.5, 20% w/v Polyethylene glycol 10,000

H3 0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 3.4 M 1,6-Hexanediol

H4 0.1 M Tris pH 8.5, 25% v/v tert-Butanol

H5 0.01 M Nickel(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 1.0 M Lithium sulfate monohydrate

H6 1.5 M Ammonium sulfate, 0.1 M Tris pH 8.5, 12% v/v Glycerol

H7 0.2 M Ammonium phosphate monobasic, 0.1 M Tris pH 8.5, 50% v/v (+/-)-2-Methyl-2,4-pentanediol

H8 0.1 M Tris pH 8.5, 20% v/v Ethanol

H9 0.01 M Nickel(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000

H10 0.1 M Sodium chloride, 0.1 M BICINE pH 9.0, 20% v/v Polyethylene glycol monomethyl ether 550

H11 0.1 M BICINE pH 9.0, 2.0 M Magnesium chloride hexahydrate

H12 0.1 M BICINE pH 9.0, 2% v/v 1,4-Dioxane, 10% w/v Polyethylene glycol 20,000

Appendix VII: Additive protein crystallization screen (Hampton Research)

1.	(A1) 0.1 M Barium chloride dihydrate	Multivalent
2.	(A2) 0.1 M Cadmium chloride hydrate	Multivalent

3. (A3) 0.1 M Calcium chloride dihydrate	Multivalent
4. (A4) 0.1 M Cobalt(II) chloride hexahydrate	Multivalent
5. (A5) 0.1 M Copper(II) chloride dihydrate	Multivalent
6. (A6) 0.1 M Magnesium chloride hexahydrate	Multivalent
7. (A7) 0.1 M Manganese(II) chloride tetrahydrate	Multivalent
8. (A8) 0.1 M Strontium chloride hexahydrate	Multivalent
9. (A9) 0.1 M Yttrium(III) chloride hexahydrate	Multivalent
10. (A10) 0.1 M Zinc chloride	Multivalent
11. (A11) 0.1 M Iron(III) chloride hexahydrate	Multivalent
12. (A12) 0.1 M Nickel(II) chloride hexahydrate	Multivalent
13. (B1) 0.1 M Chromium(III) chloride hexahydrate	Multivalent
14. (B2) 0.1 M Praseodymium(III) acetate hydrate	Multivalent
15. (B3) 1.0 M Ammonium sulfate	Salt
16. (B4) 1.0 M Potassium chloride	Salt
17. (B5) 1.0 M Lithium chloride	Salt
18. (B6) 2.0 M Sodium chloride	Salt
19. (B7) 0.5 M Sodium fluoride	Salt
20. (B8) 1.0 M Sodium iodide	Salt
21. (B9) 2.0 M Sodium thiocyanate	Salt
22. (B10) 1.0 M Potassium sodium tartrate tetrahydrate	Salt
23. (B11) 1.0 M Sodium citrate tribasic dihydrate	Salt
24. (B12) 1.0 M Cesium chloride	Salt
25. (C1) 1.0 M Sodium malonate pH 7.0	Salt
26. (C2) 0.1 M L-Proline	Amino Acid
27. (C3) 0.1 M Phenol	Dissociating Agent
28. (C4) 30% v/v Dimethyl sulfoxide	Dissociating Agent
29. (C5) 0.1 M Sodium bromide	Dissociating Agent
30. (C6) 30% w/v 6-Aminohexanoic acid	Linker
31. (C7) 30% w/v 1,5-Diaminopentane dihydrochloride	Linker
32. (C8) 30% w/v 1,6-Diaminohexane	Linker
33. (C9) 30% w/v 1,8-Diaminooctane	Linker
34. (C10) 1.0 M Glycine	Linker
35. (C11) 0.3 M Glycyl-glycyl-glycine	Linker
36. (C12) 0.1 M Taurine	Linker
37. (D1) 0.1 M Betaine hydrochloride	Linker
38. (D2) 0.1 M Spermidine	Polyamine
39. (D3) 0.1 M Spermine tetrahydrochloride	Polyamine
40. (D4) 0.1 M Hexammine cobalt(III) chloride	Polyamine
41. (D5) 0.1 M Sarcosine	Polyamine
42. (D6) 0.1 M Trimethylamine hydrochloride	Chaotrope
43. (D7) 1.0 M Guanidine hydrochloride	Chaotrope
44. (D8) 0.1 M Urea	Chaotrope
45. (D9) 0.1 M β-Nicotinamide adenine dinucleotide hydrate	Co-factor
46. (D10) 0.1 M Adenosine-5’-triphosphate disodium salt hydrate	Co-factor
47. (D11) 0.1 M TCEP hydrochloride	Reducing Agent
48. (D12) 0.01 M GSH (L-Glutathione reduced), 0.01 M GSSG (L-Glutathione oxidized)	Reducing Agent

49. (E1) 0.1 M Ethylenediaminetetraacetic acid disodium salt dihydrate	Chelating Agent
50. (E2) 5% w/v Polyvinylpyrrolidone K15	Polymer
51. (E3) 30% w/v Dextran sulfate sodium salt	Polymer
52. (E4) 40% v/v Pentaerythritol ethoxylate (3/4 EO/OH)	Polymer
53. (E5) 10% w/v Polyethylene glycol 3,350	Polymer
54. (E6) 30% w/v D-(+)-Glucose monohydrate	Carbohydrate
55. (E7) 30% w/v Sucrose	Carbohydrate
56. (E8) 30% w/v Xylitol	Carbohydrate
57. (E9) 30% w/v D-Sorbitol	Carbohydrate
58. (E10) 12% w/v myo-Inositol	Carbohydrate
59. (E11) 30% w/v D-(+)-Trehalose dihydrate	Carbohydrate
60. (E12) 30% w/v D-(+)-Galactose	Carbohydrate
61. (F1) 30% v/v Ethylene glycol	Polyol
62. (F2) 30% v/v Glycerol	Polyol
63. (F3) 3.0 M NDSB-195	Non-detergent
64. (F4) 2.0 M NDSB-201	Non-detergent
65. (F5) 2.0 M NDSB-211	Non-detergent
66. (F6) 2.0 M NDSB-221	Non-detergent
67. (F7) 1.0 M NDSB-256	Non-detergent
68. (F8) 0.15 mM CYMAL®-7	Amphiphile
69. (F9) 20% w/v Benzamidine hydrochloride	Amphiphile
70. (F10) 5% w/v n-Dodecyl-N,N-dimethylamine-N-oxide	Detergent
71. (F11) 5% w/v n-Octyl-β-D-glucoside	Detergent
72. (F12) 5% w/v n-Dodecyl-β-D-maltoside	Osmolyte
73. (G1) 30% w/v Trimethylamine N-oxide dihydrate	Organic, Non-volatile
74. (G2) 30% w/v 1,6-Hexanediol	Organic, Non-volatile
75. (G3) 30% v/v (+/-)-2-Methyl-2,4-pentanediol	Organic, Non-volatile
76. (G4) 50% v/v Polyethylene glycol 400	Organic, Non-volatile
77. (G5) 50% v/v Jeffamine M-600 pH 7.0	Organic, Non-volatile
78. (G6) 40% v/v 2,5-Hexanediol	Organic, Non-volatile
79. (G7) 40% v/v (±)-1,3-Butanediol	Organic, Non-volatile
80. (G8) 40% v/v Polypropylene glycol P 400	Organic, Non-volatile
81. (G9) 30% v/v 1,4-Dioxane	Organic, Volatile
82. (G10) 30% v/v Ethanol	Organic, Volatile
83. (G11) 30% v/v 2-Propanol	Organic, Volatile
84. (G12) 30% v/v Methanol	Organic, Volatile
85. (H1) 10% v/v 1,2-Butanediol	Organic, Volatile
86. (H2) 40% v/v tert-Butanol	Organic, Volatile
87. (H3) 40% v/v 1,3-Propanediol	Organic, Volatile
88. (H4) 40% v/v Acetonitrile	Organic, Volatile
89. (H5) 40% v/v Formamide	Organic, Volatile
90. (H6) 40% v/v 1-Propanol	Organic, Volatile
91. (H7) 5% v/v Ethyl acetate	Organic, Volatile
92. (H8) 40% v/v Acetone	Organic, Volatile
93. (H9) 0.25% v/v Dichloromethane	Organic, Volatile
94. (H10) 7% v/v 1-Butanol	Organic, Volatile

95. (H11) 40% v/v 2,2,2-Trifluoroethanol	Organic, Volatile
96. (H12) 40% v/v 1,1,1,3,3,3-Hexafluoro-2-propanol	Organic, Volatile

Appendix VIII: Wizard II protein crystallization formulation (Rigaku)

Tube	Precipitation Reagent	Buffer	Salt
1	10% (w/v) PEG 3000	100 mM Sodium acetate/ Acetic acid pH 4.5	200 mM Zinc acetate
2	35% (v/v) MPD	100 mM MES/ Sodium hydroxide pH 6.0	200 mM Lithium sulfate
3	20% (w/v) PEG 8000	100 mM Tris base/ Hydrochloric acid pH 8.5	200 mM Magnesium chloride
4	2000 mM Ammonium sulfate	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Sodium chloride
5	20% (v/v) 1,4-butanediol	100 mM HEPES/ Sodium hydroxide pH 7.5	200 mM Sodium chloride
6	10% (v/v) 2-propanol	100 mM Sodium phosphate dibasic/ Citric acid pH 4.2	200 mM Lithium sulfate
7	30% (w/v) PEG 3000	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Sodium chloride
8	10% (w/v) PEG 8000	100 mM Potassium phosphate monobasic/ Sodium phosphate dibasic pH 6.2	200 mM Sodium chloride
9	2000 mM Ammonium sulfate	100 mM Sodium phosphate dibasic/ Citric acid pH 4.2	
10	1000 mM Ammonium phosphate dibasic	100 mM Tris base/ Hydrochloric acid pH 8.5	
11	10% (v/v) 2-propanol	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Zinc acetate
12	30% (v/v) PEG 400	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Lithium sulfate
13	15% (v/v) Reagent alcohol	100 mM Sodium citrate/ Citric acid pH 5.5	200 mM Lithium sulfate
14	20% (w/v) PEG 1000	100 mM Potassium phosphate monobasic/ Sodium phosphate dibasic pH 6.2	200 mM Sodium chloride
15	1260 mM Ammonium sulfate	100 mM HEPES/ Sodium hydroxide pH 7.5	
16	1000 mM Sodium citrate tribasic	100 mM CHES/ Sodium hydroxide pH 9.5	
17	2500 mM Sodium chloride	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Magnesium chloride
18	20% (w/v) PEG 3000	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Calcium acetate
19	1600 mM Sodium phosphate monobasic/ 400 mM Potassium phosphate dibasic	100 mM Sodium phosphate dibasic/ Citric acid pH 4.2	
20	15% (v/v) Reagent alcohol	100 mM MES/ Sodium hydroxide pH 6.0	200 mM Zinc acetate
21	35% (v/v) MPD	100 mM Sodium acetate/ Acetic acid pH 4.5	

22	10% (v/v) 2-propanol	100 mM Imidazole/ Hydrochloric acid pH 8.0	
23	15% (v/v) Reagent alcohol	100 mM HEPES/ Sodium hydroxide pH 7.5	200 mM Magnesium chloride
24	30% (w/v) PEG 8000	100 mM Imidazole/ Hydrochloric acid pH 8.0	200 mM Sodium chloride
25	35% (v/v) MPD	100 mM HEPES/ Sodium hydroxide pH 7.5	200 mM Sodium chloride
26	30% (v/v) PEG 400	100 mM CHES/ Sodium hydroxide pH 9.5	
27	10% (w/v) PEG 3000	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Magnesium chloride
28	20% (w/v) PEG 8000	100 mM MES/ Sodium hydroxide pH 6.0	200 mM Calcium acetate
29	1260 mM Ammonium sulfate	100 mM CHES/ Sodium hydroxide pH 9.5	200 mM Sodium chloride
30	20% (v/v) 1,4-butanediol	100 mM Imidazole/ Hydrochloric acid pH 8.0	200 mM Zinc acetate
31	1000 mM Sodium citrate tribasic	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Sodium chloride
32	20% (w/v) PEG 1000	100 mM Tris base/ Hydrochloric acid pH 8.5	
33	1000 mM Ammonium phosphate dibasic	100 mM Sodium citrate/ Citric acid pH 5.5	200 mM Sodium chloride
34	10% (w/v) PEG 8000	100 mM Imidazole/ Hydrochloric acid pH 8.0	
35	800 mM Sodium phosphate monobasic/ 1200 mM Potassium phosphate dibasic	100 mM Sodium acetate/ Acetic acid pH 4.5	
36	10% (w/v) PEG 3000	100 mM Sodium phosphate dibasic/ Citric acid pH 4.2	200 mM Sodium chloride
37	1000 mM Potassium sodium tartrate	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Lithium sulfate
38	2500 mM Sodium chloride	100 mM Sodium acetate/ Acetic acid pH 4.5	200 mM Lithium sulfate
39	20% (w/v) PEG 8000	100 mM CAPS/ Sodium hydroxide pH 10.5	200 mM Sodium chloride
40	20% (w/v) PEG 3000	100 mM Imidazole/ Hydrochloric acid pH 8.0	200 mM Zinc acetate
41	2000 mM Ammonium sulfate	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Lithium sulfate
42	30% (v/v) PEG 400	100 mM HEPES/ Sodium hydroxide pH 7.5	200 mM Sodium chloride
43	10% (w/v) PEG 8000	100 mM Tris base/ Hydrochloric acid pH 7.0	200 mM Magnesium chloride
44	20% (w/v) PEG 1000	100 mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200 mM Magnesium chloride
45	1260 mM Ammonium sulfate	100 mM MES/ Sodium hydroxide pH 6.0	
46	1000 mM Ammonium phosphate dibasic	100 mM Imidazole/ Hydrochloric acid pH 8.0	200 mM Sodium chloride
47	2500 mM Sodium chloride	100 mM Imidazole/ Hydrochloric acid pH 8.0	200 mM Zinc acetate
48	1000 mM Potassium sodium tartrate	100 mM MES/ Sodium hydroxide pH 6.0	