Structural characterization of ligand binding in hexacoordinate hemoglobins

Julie Anne Hoy

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Biophysics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/3098

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
# TABLE OF CONTENTS

LIST OF TABLES ............................................................................................................................... v

LIST OF FIGURES............................................................................................................................. vi

ACKNOWLEDGEMENTS .................................................................................................................. viii

CHAPTER 1. Introduction .................................................................................................................. 1
  1.1. Hexacoordinate Hemoglobins ............................................................................................. 1
  1.2. Nonsymbiotic Hemoglobins ............................................................................................... 5
  1.3. Truncated Hemoglobin .................................................................................................... 8
  1.4. Dissertation Organization ................................................................................................. 11
  1.5. References ....................................................................................................................... 11

CHAPTER 2. The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. ............................................................................................................ 17
  2.1. Abstract ............................................................................................................................. 17
  2.2. Introduction ....................................................................................................................... 18
  2.3. Experimental Procedures ................................................................................................. 20
    2.3.1. Protein Production and Crystallization ................................................................... 20
    2.3.2. Structure Determination and Refinement ............................................................ 21
    2.3.3. Calculation of Structural Alignments and Solvent Accessibilities ....................... 22
  2.4. Results ............................................................................................................................... 23
    2.4.1. The Crystal Structure of *SynHb* in Comparison to the Solution NMR Structure ........23
    2.4.2. Solvent Accessibility in *SynHb* ........................................................................... 29
    2.4.3. Comparison of *SynHb* to the Structures of Other trHbs ........................................ 31
    2.4.4. Comparison of *SynHb* with Other Hexacoordinate Hemoglobins ......................... 33
  2.5. Discussion ......................................................................................................................... 37
    2.5.1. Implications for Ligand Binding in trHbs ............................................................... 37
    2.5.2. Implications for Ligand Binding in hxHbs ............................................................... 39
  2.6. Conclusions ....................................................................................................................... 39
  2.7. Abbreviations ................................................................................................................... 40
  2.8. Acknowledgments ............................................................................................................ 41
  2.9. References ....................................................................................................................... 42

CHAPTER 3. Conformational changes upon ligand binding in *Synechocystis* hemoglobin. ...................................................................................................................... 46
  3.1. Introduction ....................................................................................................................... 46
  3.2. Methods ............................................................................................................................. 47
  3.3. Results and Discussion ..................................................................................................... 50
  3.4. Conclusions ....................................................................................................................... 56
  3.5. References ....................................................................................................................... 57

CHAPTER 4. Covalent heme attachment in *Synechocystis* hemoglobin is required to prevent ferrous heme dissociation ......................................................... 59
  4.1. Abstract ............................................................................................................................. 59
4.2. Introduction........................................................................................................ 60
4.3. Results............................................................................................................ 64
  4.3.1. Crystal structures of ferric and ferrous SynHb and SynH117A ............... 64
  4.3.2. Ligand binding in the ferrous oxidation state ........................................... 66
  4.3.3. Spectroelectrochemistry ........................................................................... 68
  4.3.4. Ferric ligand binding ................................................................................ 72
  4.3.5. GdmCl unfolding ...................................................................................... 73
  4.3.6. Heme dissociation ................................................................................. 77
4.4. Discussion ....................................................................................................... 79
  4.4.1. The role of the His$^{117}\text{-vinyl bond in protein stability}$ ..................... 80
  4.4.2. Effects of the His$^{117}\text{-vinyl bond on hexacoordination and exogenous}$
          ligand binding ......................................................................................... 82
4.5. Materials and Methods.................................................................................. 84
  4.5.1. Protein expression and purification ......................................................... 84
  4.5.2. Structural determination ........................................................................ 85
  4.5.3. Carbon monoxide binding ...................................................................... 86
  4.5.4. Spectroelectrochemistry ....................................................................... 86
  4.5.5. Cyanide binding kinetics ........................................................................ 87
  4.5.6. Equilibrium ferric ligand binding ............................................................. 88
  4.5.7. Guanidinium chloride unfolding .............................................................. 88
  4.5.8. Heme Loss Assay .................................................................................. 89
4.6. Supplemental Materials.................................................................................. 90
  4.6.1. Solvent Electron Density ....................................................................... 90
  4.6.2. Crystallographic Data .......................................................................... 91
  4.6.3. Heme Pocket Stereochemistry ............................................................... 91
  4.6.4. Flash Photolysis Data ........................................................................... 92
4.7. Abbreviations.................................................................................................. 93
4.8. References....................................................................................................... 94

CHAPTER 5. Crystal Structures of Phenylalanine B10 mutants in a nonsymbiotic plant
  hemoglobin ......................................................................................................... 99
  5.1. Introduction ................................................................................................... 99
  5.2. Methods ....................................................................................................... 100
  5.3. Results ....................................................................................................... 103
  5.4. Conclusions ............................................................................................... 105

CHAPTER 6. The crystal structure of a liganded non-symbiotic plant hemoglobin
  reveals structural changes in the evolution of oxygen transport .................... 110
  6.1. Summary ................................................................................................... 110
  6.2. Introduction ............................................................................................... 111
  6.3. Results and Discussion.............................................................................. 113
    6.3.1. The structure of barHb:CN and comparison to riceHb1 ....................... 113
    6.3.2. Structural changes upon ligand binding in nsHbs ................................ 117
    6.3.3. A mechanism to achieve reversible ligand binding ........................... 119
    6.3.4. The evolution of oxygen transport in plants ..................................... 125
  6.4. Experimental Procedures ......................................................................... 129
    6.4.1. Protein expression and purification .................................................... 129
    6.4.2. Crystallization and data collection .................................................... 129
    6.4.3. Structure determination and refinement ......................................... 130
    6.4.4. Kinetic experiments ..................................................................... 130
LIST OF TABLES

Table 2-1. Data collection and refinement statistics for SynHb. ........................................... 22
Table 2-2. Helix position designations and residue numbers in trHbs. ...................................... 28
Table 2-3. Heme pocket stereochemistry. .............................................................................. 36
Table 3-1. Data collection and refinement statistics for SynHb:CN ...................................... 48
Table 4-1. Ligand binding constants and redox potentials. .................................................. 72
Table 4-2. Guanidinium chloride denaturation midpoints and energy values ...................... 74
Table 4-3. Crystallographic Data. ......................................................................................... 91
Table 4-4. Heme stereochemistry. ........................................................................................ 92
Table 5-1. Data collection and refinement statistics for riceHb1 mutant proteins .............. 102
Table 6-1. Data collection and refinement statistics for barHb:CN. .................................... 116
Table 6-2. Kinetic Data. ........................................................................................................ 138
# LIST OF FIGURES

**Figure 1-1.** Hemoglobin ................................................................. 2

**Figure 1-2.** Distal histidine configurations ........................................... 4

**Figure 1-3.** Optical spectra ................................................................. 5

**Figure 2-1.** The crystal structure of *Syn*Hb ........................................ 24

**Figure 2-2.** Temperature factors for the crystal structure and r.m.s.d. from the NMR structure ................................................................. 24

**Figure 2-3.** The crystal structure of *Syn*Hb compared with the NMR structure ......................................................... 26

**Figure 2-4.** Solvent accessibility .......................................................... 30

**Figure 2-5.** Structural comparisons .................................................... 32

**Figure 2-6.** Proximal histidine azimuthal angles of selected hemoglobins ........................................................................ 35

**Figure 3-1.** Ribbon diagram and electron density of the *Syn*Hb:CN structure ................................................................. 49

**Figure 3-2.** Structural comparison of *Syn*Hb and *Syn*Hb:CN ................ 50

**Figure 3-3.** Heme pocket overlay .......................................................... 51

**Figure 3-4.** Hydrogen bonding networks in *Syn*Hb and *Syn*Hb:CN ...... 52

**Figure 3-5.** Solvent tunnels in trHbs ...................................................... 54

**Figure 3-6.** Solvent accessibility near the heme propionates .................. 55

**Figure 4-1.** *Syn*Hb versus *Syn*H117A .................................................. 65

**Figure 4-2.** Rapid mixing kinetic measurement of CO binding ............... 67

**Figure 4-3.** Electrochemistry and ferric binding data for *Syn*Hb and *Syn*H117A ................................................................. 70

**Figure 4-4.** Guanidinium chloride denaturation ..................................... 75

**Figure 4-5.** Spectral analysis of GdmCl denaturation ............................. 76

**Figure 4-6.** Heme loss from holoprotein to mutant horse heart apomyoglobin ......................................................... 78

**Figure 4-S1.** Comparative solvent electron density ................................. 90

**Figure 4-S2.** Flash photolysis data for *Syn*H117A ................................. 93

**Figure 5-1.** Nonsymbiotic hemoglobin crystals ..................................... 101

**Figure 5-2.** X-ray crystal structures of ferric, hexacoordinate F40L and F40W riceHb1 compared to the wild type protein ......................................................... 104

**Figure 6-1.** BarHb:CN structure and electron density ........................... 115
Figure 6-2. Comparison of plant Hb structures. ................................................................. 118
Figure 6-3. Hydrogen bonds and hydrophobic networks in the CD loop region. .......... 120
Figure 6-4. Hydrogen bonds in the EF loop region. .......................................................... 121
Figure 6-5. Heme pocket comparison between barHb and riceHb1. ............................... 124
Figure 6-6. Primary sequence mutations in Lba. .............................................................. 127
Figure 6-S1. Ferrous ligand binding kinetics in cyanide bound barley Hb. .................... 139
Figure A-1. The structure of nsHb corn Hbm1. ............................................................... 150
I would like to take this opportunity to express my gratitude to my advisor, Mark Hargrove, my lab mates past and present, my POS committee members, my BBMB and BCB department colleagues, the University, and the federal government for the many ways in which they have all guided and supported me in my scientific development over the course of my Ph.D. research. I would like to thank my partner, David Doty, my family, and my friends for constant inspiration, affirmation, and encouragement. Finally, I would like to thank the taxpayer for recognizing with financial support the importance of science to the progress and survival of humanity.
CHAPTER 1. INTRODUCTION

The goal of biophysics is to study the structures of the components of living organisms and to understand the mechanics of the processes of life. Hemoglobin is a well suited model for this study. As an essential component of the life blood of mammals, and easy to obtain in large quantities, hemoglobin and its monomeric partner myoglobin are two of the most well studied and characterized components of life. Yet hemoglobin studies continue to reveal new forms of hemoglobin, raising new questions, functional possibilities, and research opportunities.

This research focuses on hemoglobins characterized as “hexacoordinate.” The introduction below outlines the relationship of this family of hemoglobins within the hemoglobin superfamily. Two specific types of hexacoordinate hemoglobins, the nonsymbiotic and truncated hemoglobins, are then described in detail, to provide a background for the specific protein studies described in the following chapters.

1.1. Hexacoordinate Hemoglobins

Hemoglobin proteins (Hbs) are composed of $\alpha$-helices, loops, and an heme b, iron protoporphyrin IX prosthetic group (see Figure 1). They are proficient at binding small, polar, diatomic, gaseous ligands such as oxygen, carbon monoxide, and nitric oxide. The general compact globin fold is universal among hemoglobins, yet individual primary sequences and ligand binding kinetics can vary drastically. This versatility has allowed for
the evolution of specialized hemoglobins to meet varied functional needs in practically all living organisms.

**Figure 1. Hemoglobin.** A) The globin fold as illustrated by sperm whale myoglobin (2MBW.pdb). Helices are labeled alphabetically from the N- to C-termini. The soluble protein harbors the hydrophobic heme. Note the universally conserved proximal histidine below the heme plane. Myoglobin has a histidine in the distal position above the heme plane. The proximal and distal residues are referred to as His(F8) and His(E7) respectively based on their helical position. The residues of other hemoglobins are similarly notated based on their homology to myoglobin. B) The heme b chemical structure and notation. Four of the six iron coordination sites bind nitrogens from the heme pyrroles.

The only residue strictly universally conserved in all hemoglobins is the proximal histidine, which coordinates the heme iron, holding the heme within the protein. The vertebrate hemoglobins first studied can be termed “pentacoordinate” hemoglobins because five of the
six iron coordination sites are filled by the four heme pyrroles and the proximal histidine. The sixth site binds exogenous ligands or solvent.

The first hemoglobins found in plants (the so called leghemoglobins because they were first found in legumes) are also pentacoordinate. These globins are only found in plants involved in a symbiotic relationship with a bacterium, and were thus also termed “symbiotic” hemoglobins (sHbs). When hemoglobins were discovered in plants not involved in symbiosis, these hemoglobins were termed “nonsymbiotic” (nsHbs).

The nonsymbiotic hemoglobins were one of the first instances of naturally occurring “hexacoordinate” hemoglobins (hxHbs), so named because the remaining sixth coordination site of the iron atom is endogenously coordinated by an amino acid within the protein, frequently the “distal” histidine. Hexacoordinate hemoglobins have since been found among the “truncated” hemoglobins (trHbs), a group that maintains the globin fold in a shortened form with fewer helices. They have also been found in vertebrates, including humans, and present a new area of research in the hemoglobin field.

In hexacoordinate hemoglobins, the distal histidine is different from the proximal histidine in that its coordination to the heme iron is reversible. Thus, the hxHbs are capable of exogenous ligand binding despite the endogenous coordination of all six binding sites. This feature also distinguishes the hexacoordinate hemoglobins from other bis-histidyl heme proteins like cytochrome b$_5$ which is not capable of ligand binding due to nonreversible endogenous coordination (Figure 2).
Hexacoordinate hemoglobins are easily distinguished by the differences in their characteristic optical spectra, as depicted in Figure 3. Because exogenous ligand binding has to compete with endogenous ligand binding, the kinetics of hxHbs frequently differ as well. They tend to have high ligand affinities, despite distal histidine binding, due to low dissociation constants. Although the function of hxHbs has not yet been precisely determined, it is likely that the distal histidine plays a mechanistic regulatory role in ligand binding that affects the physiological capabilities of this type of hemoglobin.

Figure 2. Distal histidine configurations. In A) the pentacoordinate symbiotic plant leghemoglobin from soybean, Lba (1BIN.pdb), the distal histidine does not coordinate the heme, while it does in B) an hexacoordinate plant nonsymbiotic hemoglobin from rice (1D8U.pdb). Both are capable of exogenous ligand binding. C) The heme in cytochrome b5 is also hexacoordinate, but incapable of ligand binding (1CYO.pdb).
Figure 3. Optical spectra. Ferric (oxidized or Fe$^{3+}$, black) and ferrous (reduced or Fe$^{2+}$, grey) spectra of A) pentacoordinate horse heart myoglobin (hhmb), and B) hexacoordinate rice hemoglobin (riceHb1). The peak near four hundred is called the soret peak. Note the double peak around 550 in ferrous riceHb1 that differs from the single peak in hhmb.

1.2. Nonsymbiotic Hemoglobins

It was first suggested in 1988 that all plants contained hemoglobin in the roots, if not elsewhere [1]. Before then, only symbiotic plant hemoglobins had been discovered. These included the leghemoglobins and a hemoglobin from the non-legume *Parasponia* [2], both of which form a symbiotic relationship with *Rhizobium*, and also a hemoglobin from *Casuarina* which is nodulated by *Frankia* [3]. However, subsequently hemoglobins were found in *Trema*, *Celtis*, and *Ulmus*, all of which are plants uninvolved in symbiosis with nitrogen fixing bacteria [4]. A second, nonsymbiotic, hemoglobin was also found in the roots of *Casuarina*, suggesting different genes coding for separate hemoglobins with tissue specific
expression patterns [5]. It was this that led to the speculation that hemoglobin might be found universally in plant roots or other tissues.

The discovery of hemoglobin in barley expanded the types of plants known to contain nsHbs from dicots to include monocots, and strengthened the supposition that hemoglobins could be found in all plants [6]. It also confirmed a change in thinking regarding the evolutionary origin of plant hemoglobin from one of horizontal transfer to symbiotic plants from animals to concurrent evolution with animal hemoglobin. It became clear that the leghemoglobins were a case of evolutionary specialization from the nsHbs, likely from gene duplication.

A nsHb was then found in soybeans [7] that expressed in a variety of tissues, and had greater similarity to other nsHbs than to symbiotic soybean leghemoglobin. Two nsHbs were found in *Arabidopsis thaliana* with such differing expression patterns and oxygen affinities, that two classes of nsHbs were subsequently designated [8]. Class 1 showed strong sequence similarity to the previous nsHbs, with similar low level root expression, while Class 2 was more similar to the sHbs.

The hexacoordinate nature of Class 1 nsHbs was first characterized in barley [9], and then confirmed in rice studies [10], including the x-ray crystal structure of hexacoordinate riceHb1 [11], the only nsHb structure prior to this research. Both barley and rice nsHbs were found to have low expression in root tissues, and high oxygen affinities due to low dissociation rates, suggesting they might have functions other than oxygen storage and

Since then, a flurry of studies on a number of primarily Class 1 nsHbs have attempted to determine the function or functions of these proteins. In rice, Class 1 nsHb synthesis is elevated in etiolated and flooded plants, but not in plants subjected to oxidative, nitrosative, or hormonal stress [14]. This protein was also found in developing rice seedlings with highest levels during the early stages of the germinating seed [15]. In addition, induction is strongly increased by nitrate, nitrite, and nitric oxide donors in association with induction of genes for NADH-nitrate reductase [16].

Alfalfa root cultures overexpressing barley hemoglobin were found to maintain growth, have lower ATP/ADP ratios, less accumulation of nitric oxide, and no breakdown of cells under hypoxia [17]. They also exhibit nitric oxide dioxygenase-like activity with the nsHb acting with a flavohemoglobin to metabolize nitric oxide to nitrate using NADH as an electron donor [18]. Maize cells expressing barley hemoglobin also experienced lower levels of nitric oxide under anoxic conditions [19], as did Arabidopsis plants overexpressing a Class 1 nsHb [20]. This hemoglobin plays a role in nitric oxide detoxification, but does not interrupt nitric oxide accumulation involved in defense signaling.

A nsHb from cotton has been found to play a role in defense responses by protecting against infection from the fungus pathogen Vericillium dahliae [21], while tobacco plants overproducing a nsHb from alfalfa had reduced necrotic symptoms linked to increased
accumulation of a NO-affected pathogenesis related transcript [22]. That study also indicated that both monocot and dicot nsHbs affect cellular NO levels by scavenging nitric oxide. It was found that in barley, cytosolic monodehydroascorbate reductase supports NADH-dependent NO scavenging and that ascorbate facilitated the rate of reduction of oxidized Hb [23]. An additional study on barley indicates that an organelle calcium channel blocker inhibits anoxia-induced nsHb expression that can be restored by addition of calcium, implicating calcium in Hb regulation [24].

These studies of nsHbs reveal that this type of hemoglobin plays a complicated and important scavenging or signal role in many critical plant pathways. The functional roles of nsHbs are certainly affected by their characteristic hexacoordination, and perhaps also by more subtle aspects of their molecular structure. The process of deciphering the function and evolutionary development of hxHbs can be aided by examining the specific affects of hexacoordination on nonsymbiotic hemoglobin structure and ligand binding behavior.

### 1.3. Truncated Hemoglobin

Short hemoglobins highly related to each other, but without notable similarity to other known hemoglobins were first found in various species of *Paramecium* [25, 26] and *Tetrahymena* [27, 28], in *Nostoc commune* [29], and in *Chlamydomonas eugametos* [30, 31]. Additional studies identified these Hbs in many other species, including *Mycobacterium tuberculosis* [32, 33] and *Synechocystis* [34, 35]. This phylogenetically distinct family of hemoglobins were eventually termed “truncated hemoglobins,” due to the fact that they can be 20-40
residues shorter than other non-vertebrate hemoglobins and exhibit a shortened globin fold [36, 37].

Over one hundred trHbs have currently been identified in eubacteria, cyanobacteria, unicellular eukaryotes, and plants [37, 38]. They have been divided into three subgroups based on sequence identity, and some organisms contain more than one type, notably *Mycobacterium avium* which contains one trHb of each type [36, 37, 39-41]. The first structures of trHbs revealed a compact globin fold with fewer helices, leading to a 2-on-2 α-helical “sandwich” fold, rather than the typical 3-on-3 α-helical fold of longer hemoglobins [36]. TrHb structures indicate these proteins are capable of forming a tunnel within the protein matrix that possibly assists in ligand diffusion to the heme, ligand storage, or multiligand reactions [37, 40-43].

There is less amino acid conservation in the Group I trHbs than in non-vertebrate Hbs and Groups II and III trHbs [38]. While the proximal His(F8) is conserved, the residue at position CD1 can be Phe, Tyr, or His. The residue at position B10 is most frequently a Tyr, but can be a histidine or a hydrophobic residue, while the distal E7 position contains over half a dozen residue types [37, 38]. Prevalent in the trHbs, a number of residues lining the tunnel are conserved within the trHb family, as is Phe(E14) which plays a solvent shielding role near the heme [37, 38, 41].

Only some of the trHbs are capable of hexacoordination. There is evidence that the trHb from *Nostoc commune* is endogenously hexacoordinate at neutral pH [44, 45]. The trHb
from *Chlamydomonas eugametos* is endogenous hexacoordinate at alkaline pH, with the coordinating residue likely Tyr(B10) [31, 46, 47]. In *Arabidopsis thaliana*, a trHb is capable of forming an unstable hexacoordinate complex [48]. The trHb from *Synechocystis* is coordinated by a distal histidine in both the ferric and ferrous forms over a wide pH range, however this distal histidine is at position E10 rather than the typical E7 position [35, 49-51].

In addition to stable hexacoordination, the trHb from *Synechocystis* is interesting on an evolutionary level. Due to the antiquity of the cyanobacteria, the trHb from this organism may provide insights into primitive hemoglobin form and function. *Synechocystis* is also a model organism for the study of photosynthesis, and it is possible that the hemoglobin found in this organism may play some role in this process. In addition, this trHb has a unique third bond to the heme, this one a covalent link between His(H17) and the heme 2-vinyl [52].

The diversity of organisms in which trHbs are found, along with the diversity in expression patterns and cellular localization, indicate that trHbs could play a number of functional roles. Possible functions include a role in photosynthesis, oxygen sequestering, nitrogen fixation, and nitric oxide detoxification [30, 37, 44, 45, 53, 54]. Further characterization is needed to determine how the structural modifications of truncation impact the functional roles possible for these proteins and whether or not hexacoordination plays the same regulatory roles in trHbs as in nsHbs.
1.4. Dissertation Organization

The goal of this research is to increase the body of knowledge about ligand binding in hexacoordinate hemoglobins by examining both their global structural accommodation of ligands and the role of specific residues in regulating ligand affinity. This research has focused specifically on the trHb from *Synechocystis* and on the nsHbs from rice and barley. *Syn*Hb was chosen because it combines several unique characteristics that set it apart from other truncated and hexacoordinate proteins, while riceHb1 and barley Hb are the common model proteins for study within the plant non-symbiotic hemoglobin group.

Following this introduction, Chapter 2 describes structural characterization of endogenously hexacoordinate *Syn*Hb by x-ray crystallography. Chapter 3 examines the structural conformation changes that take place in *Syn*Hb upon exogenous ligand binding. Chapter 4 explores the function of the unique covalent bond found in *Syn*Hb between the heme and a non-axial histidine within the protein. Changing the focus from trHbs to nsHbs, Chapter 5 centers on the effect of the mutation of a key residue in the hexacoordinate nsHb riceHb1, while Chapter 6 describes the effect of exogenous ligand binding on barley Hb. Concluding remarks are made in Chapter 7.

1.5. References


CHAPTER 2. THE CRYSTAL STRUCTURE OF SYNECHOCYSTIS HEMOGLOBIN WITH A COVALENT HEME LINKAGE.

A paper published in The Journal of Biological Chemistry

Julie A. Hoy\textsuperscript{2,3}, Suman Kundu\textsuperscript{2}, James T. Trent III\textsuperscript{2}, S. Ramaswamy\textsuperscript{4}, and Mark S. Hargrove\textsuperscript{2,5}

2.1. Abstract

The x-ray crystal structure of \textit{Synechocystis} hemoglobin has been solved to a resolution of 1.8 Å. The conformation of this structure is surprisingly different from that of the previously reported solution structure, probably due in part to a covalent linkage between the heme 2-vinyl and His\textsuperscript{117} that is present in the crystal structure but not in the structure solved by NMR. \textit{Synechocystis} hemoglobin is a hexacoordinate hemoglobin in which the heme iron is coordinated by both the proximal and distal histidines. It is also a member of the "truncated hemoglobin" family that is much shorter in primary structure than vertebrate and plant

\textsuperscript{2}Primary Author and Researcher
\textsuperscript{3}Department of Biochemistry, Biophysics, and Molecular Biology; Iowa State University, Ames, IA 50011
\textsuperscript{4}Department of Biochemistry; University of Iowa, Iowa City, IA 52242
\textsuperscript{5}To whom correspondence should be addressed
hemoglobins. In contrast to other truncated hemoglobins, the crystal structure of *Synechocystis* hemoglobin displays no "ligand tunnel" and shows that several important amino acid side chains extrude into the solvent instead of residing inside the heme pocket. The stereochemistry of hexacoordination is compared with other hexacoordinate hemoglobins and cytochromes in an effort to illuminate factors contributing to ligand affinity in hexacoordinate hemoglobins.

### 2.2. Introduction

A wide diversity in both form and function has been discovered in the study of hemoglobins (Hbs) from many species, including bacteria (1, 2), plants (3), and humans (4–7). Although the physiological functions of many hemoglobins are still unknown, a number of new functions have recently been described, including the scavenging of nitric oxide and oxygen (2, 8), aerotaxis (9), and phototaxis (10). Although functions vary, the tertiary structures of these hemoglobins conserve many of the general features of the globin fold even when truncated to very short primary structures (11–13). However, like many proteins sharing the same fold but carrying out different functions, the specific amino acid residues surrounding the heme prosthetic group exhibit a large degree of variability across hemoglobins from different species.

An extreme example of this diversity is found in the family of hexacoordinate hemoglobins (hxHbs) in which an endogenous amino acid coordinates the ligand binding site of the heme iron in the absence of exogenous ligands. The fact that hxHbs are capable of reversible
exogenous ligand binding distinguishes them from cytochrome b$_5$ and denatured Hbs that are not capable of ligand binding in the hexacoordinate state (12, 14–16). This unusual characteristic could be an alternative form of regulating ligand affinity (17, 18) or indicate a different functional role for hxHbs compared with the traditional role of oxygen storage and transport (19–22). Although hxHbs are found in many species, few structures have been reported in the hexacoordinate state. Thus, further investigation of structures across a wide variety of species will aid in understanding reversible hexacoordination and discovering its physiological significance.

The hemoglobin found in the single-celled cyanobacterium *Synechocystis* sp. PCC 6803 (SynHb) belongs to the truncated hemoglobin (trHb) family found extensively in eubacteria, bacteria, single celled eukaryotes, and plants (2, 23–26). The trHbs are 20–40 residues shorter than non-vertebrate hemoglobins and have a 2-on-2 α-helical globin fold rather than the typical 3-on-3 fold. They are further characterized by a hydrophobic tunnel connecting the protein surface to the distal heme pocket that could serve as a direct route for ligand entry and/or exit from the heme pocket (2). In addition to being a hexacoordinate member of the trHb family, SynHb is distinguished by the ability to form a covalent bond between the heme 2-vinyl group and the His117 side chain (27); this is a novel example of covalent attachment in a Hb via the porphyrin ring, and the purpose of this modification is not yet understood.

In the present study, the crystal structure of SynHb has been solved to 1.8-Å resolution. The crystal structure reveals the covalent linkage between the heme 2-vinyl and the Ne$_2$ atom of His117 that is not present in the solution NMR structure of this protein (28). The crystal
structure is compared with the solution NMR structure and to other trHb and hxHb structures in an effort to understand the role of this unusual covalent modification as well as ligand entry, stabilization, and exit from \( \text{SynHb} \).

2.3. Experimental Procedures

2.3.1. Protein Production and Crystallization

\( \text{SynHb} \) was produced as described previously (18, 24) with the following exceptions. Expression was performed without induction by isopropyl-1-thio-\( \beta \)-D-galactopyranoside, and free hemin was not added during fermentation. Following inoculation, cells were grown for 16 h and then harvested by centrifugation. The resulting supernatant was bright red due to soluble heme-bound \( \text{SynHb} \) existing predominately in the ferrous state. Protein purification proceeded as described previously (18), resulting in protein with a Soret/\( A_{280} \) absorbance ratio greater than 5.0. The purified protein was oxidized with a slight molar excess of potassium ferricyanide followed by desalting on a Sephadex G-25 column in 0.01 M potassium phosphate, pH 7. The resulting ferric \( \text{SynHb} \) was concentrated to \( \sim3 \) mM and stored at -80 °C until use.

The ferric protein used for crystallization was produced under conditions similar to those in which the His117-heme covalent link was observed by NMR experiments in \( \text{SynHb} \) and the homologue in Synechococcus (28, 29). However, the reported method of dithionite treatment leading to the covalent link (27, 28) was not used in our treatment of the protein. Crystal
growth was achieved by hanging-drop vapor diffusion. Drops were produced by mixing 2 µl of 3 mM protein with 2 µl of well buffer containing 30–35% polyethylene glycol monomethylether 5000, 0.2 M ammonium sulfate, 0.01 M cadmium chloride, and 0.1 M MES at pH 6.5. Single crystals grew in 1–3 days at 4 °C.

2.3.2. Structure Determination and Refinement

Initially, solution of the crystal structure of SynHb was attempted using molecular replacement starting with the NMR structure of this protein (28). The fact that this method failed is not unusual because the use of NMR structures for molecular replacement starting models is often unsuccessful, even with 100% sequence identity (30, 31). Therefore, multiple wavelength anomalous diffraction was employed using anomalous scattering from the heme iron atom. These data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL, Stanford, CA) at 100 °K. Diffraction data were processed using the program d*TREK (32) from Rigaku/MSC (peak and inflection data sets required two batches for proper processing of anomalous data). Phases were calculated to 2.0-Å resolution using SOLVE (33), with a starting figure of merit of 0.58. Amino acids were built into the electron density using RESOLVE (34) followed by manual rebuilding in O (35) and refinement using REFMAC5 (36), a program from the CCP4 suite (37). The resolution was extended to 1.8 Å using a native data set collected on a Rigaku/MSC home source generator and further refinement using REFMAC5. The final model contains 123 amino acids and 77 water molecules with R = 21.8%, Rfree = 22.4%, and acceptably small variations from ideal stereochemistry (Table I). A plot of B factor versus residue number is shown in Fig. 2A. As
expected, B factors are low except in loop and termini regions. The atomic coordinates of SynHb have been deposited in the Protein Data Bank (www.rcsb.org, PDB ID 1RTX.pdb).

<table>
<thead>
<tr>
<th>Table 2-1. Data collection and refinement statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) Data collection statistics</strong>^a^</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Peak</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>R-merge (%)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Unique Reflections</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Remote</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>R-merge (%)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Unique Reflections</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Inflection</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>R-merge (%)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Unique Reflections</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Native</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>R-merge (%)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Unique Reflections</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Refinement statistics and model quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>SynHb</td>
</tr>
<tr>
<td>Space Group</td>
</tr>
<tr>
<td>Unit Cell (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Refined residues</td>
</tr>
<tr>
<td>Refined waters</td>
</tr>
<tr>
<td>R&lt;sub&gt;cryst&lt;/sub&gt; (%)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)</td>
</tr>
<tr>
<td>Ave. B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
</tr>
</tbody>
</table>

^a^Peak and inflection data sets were processed in two batches to properly process anomalous data.

^b^Outer shell statistics are shown in parentheses.

^c^Calculated using 5.8% of reflections.

**2.3.3. Calculation of Structural Alignments and Solvent Accessibilities**

Structural alignments and r.m.s.d. (root mean square deviation) values were calculated using the Deep View Swiss PDB Viewer (38) and checked with LSQMAN from the Uppsala
Software Factory (39). The most representative NMR structure was determined by NMRCLUST (40), and LSQMAN was used to find the average r.m.s.d. among the 20 NMR models relative to the most representative structure. Solvent accessibility was calculated with the program SURFNET (41) using a 1.4-Å probe. All values in Table III and Fig. 6 were calculated with MATLAB (The Math Works, Inc). The least squares plane (42) of the heme was calculated using the 24 atoms of the porphyrin macrocycle, whereas that of the histidine was determined using the 5 atoms of the imidazole ring. The figures were created using the Deep View Swiss PDB Viewer and Igor Pro (Wavemetrics).

2.4. Results

2.4.1. The Crystal Structure of SynHb in Comparison to the Solution NMR Structure

The backbone structure and electron density for SynHb are shown in Fig. 1. The overall structure is typical of other trHbs and the NMR structure of SynHb in that the B and E, and G and H helices form the characteristic 2-on-2 helical core of the globin fold (2) (Fig. 1A). The covalent linkage between the Ne2 atom of His117 and the heme-2-vinyl is observed as continuous electron density between these atoms (Fig. 1B). There is no evidence for a mixture of states for this side chain, suggesting that the protein is exclusively in the covalently cross-linked conformation. The Ne2 atom of His117 is 2.1 Å from the CAB atom of the heme in the crystal structure, whereas these atoms are 5.4 Å apart in the most representative model from NMR. This movement is accompanied by a 136° change in the
Figure 1. The crystal structure of SynHb. A, backbone ribbon diagram of the crystal structure of SynHb, including His$^{46}$ and His$^{70}$ (which coordinate the heme iron), and His$^{117}$, which is covalently bound to the heme 2-vinyl. B, cross-eyed stereo view representation of electron density for the heme pocket of the crystal structure. Density is contoured to 1.5 $\sigma$ and rotated $\sim 180^\circ$ from the view in A.

Figure 2. Temperature factors for the crystal structure and r.m.s.d. from the NMR structure. A, a plot of $B$ factors (black) and r.m.s.d. between the backbone alignment of the crystal and NMR structures of SynHb (gray). B, the SynHb crystal structure, color-coded for r.m.s.d. from the NMR structure. r.m.s.d. values <1.5 Å are in blue, 1.5–3 Å in orange, and >3 Å in red.
angle of CH12 and a $>30^\circ$ increase in the acute angle between the least squares planes of the heme porphyrin and the His117 imidazole ring.

Another distinct feature of the SynHb crystal structure is the relatively large r.m.s.d. from the NMR structure, as illustrated in Figs. 2 and 3. Two methods were used to align the structures for quantitative comparison. The first method aligns the structures based on the C$\alpha$ atoms of 109 residues (89% of the residues in the SynHb structure) with clear backbone density in the crystal structure (residues 5–55, 60–98, 105, and 107–124). The second method aligns the structures based on the C$\alpha$ atoms of 86 residues (70%) involved in clear secondary structure (residues 17–55, 64–98, and 101–112) and was chosen because it was the method used for aligning the 20 NMR models of SynHb (28).

The r.m.s.d. values for the density-based and secondary structure-based alignments were relatively high at 2.18 and 1.6 Å, respectively. The difference in r.m.s.d. values between the two methods arises from the number of residues used for each alignment. Both alignments give an r.m.s.d. of 2.6 Å when all 123 C$\alpha$ atoms of the identical sequence are taken into account, rather than only the subset of C$\alpha$ atoms used in the alignment. As a frame of reference, the average r.m.s.d. among the 20 NMR models relative to the most representative structure is 1.37 Å for all C$\alpha$ atoms. In addition, the comparison between the crystal and NMR structures of SynHb gives a higher r.m.s.d. value than the comparison of the SynHb crystal structure and that of the trHb from Chlamydomonas eugametos (CtrHb). The latter comparison aligns 94 C$\alpha$ (76%) with an r.m.s.d. of 1.53 Å, even though CtrHb is in a different liganded state than SynHb and does not share the same primary sequence.
Figure 3. The crystal structure of SynHb compared with the NMR structure. A, a cross-eyed stereo overlay of the crystal structure of SynHb in red and the most representative NMR structure in blue, with helices labeled; alignment is based on the Cα atoms of 109 residues with clear backbone density. B, a crossed-eyed stereo overlay of the heme pocket, aligned with respect to the heme group.
The source of deviation between the crystal and NMR structures is examined on a per residue basis in Fig. 2. Fig. 2A includes a plot of the r.m.s.d. between the SynHb crystal and NMR structures. Fig. 2B is a ribbon representation of the crystal structure backbone onto which r.m.s.d. values have been color-coded for each amino acid (blue is $<1.5$ Å, orange 1.5–3 Å, and red $>3$ Å). Several regions of increased r.m.s.d. are expected because of increased mobility. These regions include the termini and extended loops, both of which correlate with increased B factors in Fig. 2A. There are also regions described by relatively low B factors that have increased r.m.s.d., including the entire F helix, the first half of the E helix, and a portion of the lower H helix containing the covalently linked His117. These differences between the crystal and NMR structures of SynHb are seen clearly in the backbone alignment shown in Fig. 3A.

In the case of hemoglobins, it can be argued that structural alignment based on the heme prosthetic group is the most important functionally, because it gives a direct comparison of the orientation of the amino acid side chains influencing the steric and electrostatic environment immediately surrounding the ligand binding site. Alignment based on heme position rather than the protein backbone increases the deviation between the crystal and NMR structures of SynHb from 2.6 Å to 3.76 Å for all Cα. This indicates that the heme orientations within the crystal and NMR structures of SynHb are different (as can also be seen in Fig. 3A). Both the Cα and heme-based alignment comparisons indicate significant deviations between the crystal and NMR structures of SynHb, particularly at high resolution.
Inclusion of side chain positions in the backbone comparison of these structures also generates r.m.s.d. values > 3 Å. Differences occur throughout the backbone but are most pronounced in the heme pocket overlay shown in Fig. 3B, which provides a direct comparison of the crystal structure and the most representative NMR structure. Several side chains, including Tyr22 and Phe35, are shifted upwards, away from the heme in the NMR structure. The H helix is closer to the heme in the crystal structure, which brings His117 into position to form a covalent bond to the heme 2-vinyl group. Additional high resolution structural differences include the position of the side chains of the proximal and distal histidines, His70 and His46, respectively (Table II), which have different orientations with respect to the heme porphyrins in the two structures. The proximal histidine in the crystal structure is tilted toward the propionates with an acute angle of 79° between the least squares planes of the heme porphyrin and the histidine imidazole ring, whereas in the NMR structure the proximal histidine is tilted away from the propionates with an acute angle of 77°. The

<table>
<thead>
<tr>
<th>Protein</th>
<th>B10</th>
<th>CD1</th>
<th>E7</th>
<th>E10</th>
<th>E11</th>
<th>E14</th>
<th>F8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synechocystis</td>
<td>Tyr22</td>
<td>Phe35</td>
<td>Gln43</td>
<td>His46</td>
<td>Gln47</td>
<td>Phe50</td>
<td>His70</td>
</tr>
<tr>
<td>C. eugametos</td>
<td>Tyr20</td>
<td>Phe33</td>
<td>Gln41</td>
<td>Lys44</td>
<td>Gln45</td>
<td>Phe48</td>
<td>His68</td>
</tr>
<tr>
<td>P. caudatum</td>
<td>Tyr20</td>
<td>Phe33</td>
<td>Gln41</td>
<td>Lys44</td>
<td>Thr45</td>
<td>Phe48</td>
<td>His68</td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td>Tyr33</td>
<td>Phe46</td>
<td>Leu54</td>
<td>Lys57</td>
<td>Gln58</td>
<td>Phe61</td>
<td>His81</td>
</tr>
</tbody>
</table>

Helical designations are determined by alignment with sperm whale myoglobin
effect of this tilt on the azimuthal angle of the proximal histidine is discussed below. Similarly, the tilts of the distal histidines are nearly opposite; that in the crystal structure tilts toward the 1- and 8-methyl groups with an acute angle of 73°, whereas the one in the NMR structure tilts toward the 4-vinyl and 5-methyl groups with an acute angle of 76°.

### 2.4.2. Solvent Accessibility in SynHb

The crystal structures of three other group I trHbs have been solved to date, CtrHb, Paramecium caudatum (PtrHb), and Mycobacterium tuberculosis (trHbN) (13, 43), all with bound ligands. Each of these structures display a continuous or nearly continuous "ligand tunnel" that connects the distal heme pocket with solvent. The tunnel contains two branches, with openings between the AB and GH inter-helical corners and between the E, G, and H helices (2, 13). Utilizing the same parameters for tunnel calculations in the other trHbs (see "Experimental Procedures"), SynHb does not contain this tunnel. In the SynHb crystal structure, crowding from residues such as Leu51, Phe55, and Met98 narrows the space between the AB and GH corners and the distal heme pocket, whereas Phe50 blocks access to the distal pocket via this path completely. Furthermore, Tyr53 in SynHb prevents access to the heme from the branch of the tunnel found between the G and H helices. For the NMR structure, it was reported that this tunnel path was detected (28). However, when the same method was again used to check for solvent accessibility in the NMR structure, it was found that, despite the porous surface of the protein, no continuous tunnels extend through the protein matrix to the distal heme pocket.
Figure 4. Solvent accessibility. A cross-eyed stereo view of the solvent accessible area around the propionates in the crystal structure of SynHb (A) and the NMR structure (B).
Both the crystal and NMR structures of \textit{Syn}Hb are solvent-accessible around the heme propionates. The solvent-accessible cavity of the crystal structure (Fig. 4A) is smaller than that of the NMR structure (Fig. 4B), which is more extensive and reaches into the heme pocket to surround the distal His46. However, the rest of the heme group is blocked from solvent as is the case in other trHbs. For example, in the crystal structures of CtrHb and PtrHb, the side chains of Phe48 and Trp59 block the heme from solvent accessibility at the CHD methinic bridge (13). In the crystal structure of \textit{Syn}Hb, Tyr53 and Tyr61 guard this side of the heme from solvent. In the NMR structure of \textit{Syn}Hb, Tyr53 also serves this function but Tyr61 does not due to the wide swing of the pre-F loop in this conformation (28).

\textbf{2.4.3. Comparison of \textit{Syn}Hb to the Structures of Other trHbs}

\textit{Syn}Hb is the first crystal structure of a trHb in the unliganded state. Of the three crystal structures previously solved for other group I trHbs, \textit{Syn}Hb shares highest sequence identity with CtrHb (34\%) and somewhat lower sequence identities with trHbN (24\%) and PtrHb (20\%) (25). The overlay of \textit{Syn}Hb and CtrHb is shown in Fig. 5A. \textit{Syn}Hb shares a 2-on-2 \(\alpha\)-helical fold more similar to that of CtrHb and PtrHb than trHbN (respective r.m.s.d. values are 1.53 Å over 94 C\(\alpha\), 1.65 Å over 85 C\(\alpha\), and 2.16 Å over 51 C\(\alpha\)). All four of these trHbs exhibit the characteristic three glycine motifs, the shortened one-turn A helix (though trHbN also has an N-terminal extension), and significant deletions in the CD-D region.

General differences between the \textit{Syn}Hb crystal structure and other trHbs include a shift of the \textit{Syn}Hb B helix away from the heme group, and movement of the E helix to a position closer to the heme that facilitates the bond between His46 and the distal site of the heme iron. The F
helices in other trHb crystal structures contain a wide "pre-F" loop preceding a one turn F helix hosting the proximal His(F8) ("F8" designates the helix position of the proximal His; see Table II). In contrast, the pre-F loop in SynHb is shorter and closer to the heme due to more extensive helical structure in this region. The F helix in the crystal structure of SynHb begins with one and a half turns (residues F1–F6 as opposed to residues F4–F8 in the others), is interrupted by a three residue loop hosting the His(F8), and ends in a one-turn loop of four residues. This is different from the NMR structure of SynHb, in which the pre-F loop swings out in the opposite direction from the three ligand bound structures, and the F helix is one long, 13-residue helix beginning at F2. Finally, the H helix in SynHb is interrupted and bent
by a loop (similar to that in PtrHb) that brings the end of the H helix close to the heme at the His117 covalent linkage.

Differences in the locations of specific, functionally important heme pocket amino acids in the crystal structure of SynHb compared with those of the other trHb proteins can be summarized as follows. 1) His(F8), Phe(CD1), and Tyr(B10) are conserved in all four structures (Table II). The Tyr(B10) side chain is found in the distal heme pocket and acts to stabilize the ligand in the other trHbs (13, 43–45) but extrudes out into the solvent in the SynHb structure. 2) Gln(E7) in CtrHb and PtrHb also stabilizes ligand binding (13), but in SynHb, Gln(E7) extrudes out into the solvent, up and away from the heme, and forms hydrogen bonds with two solvent molecules. 3) Thr(E11) in PtrHb and Gln(E11) in both CtrHb and trHbN participate in a network of hydrogen bonds that includes the side-chain hydroxyl of Tyr(B10) in each protein (13, 43). Gln(E11) in SynHb is located well above the heme and hydrogen bonds to a solvent molecule. 4) In the other trHb structures, the Lys(E10) side chain extrudes out into the solvent, interacting with the heme propionates in CtrHb and PtrHb (13). The corresponding His(E10) side chain in SynHb binds the distal site of the heme iron. In summary, besides hexacoordination, the major difference among the heme pockets of these structures is that several side chains important for stabilizing the bound ligand in other trHbs are found interacting with solvent in SynHb.

### 2.4.4. Comparison of SynHb with Other Hexacoordinate Hemoglobins

The structures of hemoglobins in the hemichrome state from five other organisms have been reported. These are the Hb-C chain from the sea cucumber Caudina arenicola (Hb-C) (46),
rice non-symbiotic Hb (riceHb1) (12), a tetrameric hemoglobin from the Antarctic fish Trematomus newnesi (HbTn) (47) (with the β-chains in a hemichrome state), horse methemoglobin (eHb) at pH 5.4 (with the α-chains in a hemichrome state) (48), and human neuroglobin (Ngb) with three Cys mutations (11). Backbone overlays of these hxHbs (none of which are trHbs) do not align well with SynHb. However, SynHb looks most like a truncated version of riceHb1, with an r.m.s.d. of 1.71 Å over 89 Ca atoms (Fig. 5B).

In the other hxHbs the hexacoordinating side chain is the distal His(E7), but in the two structures of SynHb the hexacoordinating distal residue is His(E10). Heme pocket comparisons reveal further interesting differences in stereochemistry among these hxHbs (Fig. 6 and Table III). For example, Fig. 6 shows the azimuthal angles for the above hxHbs and the two SynHb structures. Myoglobin (Mb) (49) and leghemoglobin (Lba) (50) are also shown, along with four cytochrome b5 proteins; bovine (bCYT) and rat (rCYT) cytochrome b5 (51, 52), and the cytochrome b5 domains from human (HSO) and chicken liver (CSO) sulfite oxidases (53, 54).

The azimuthal angle is defined as the intersection of the least squares plane of the proximal histidine with the least squares heme plane, relative to a line connecting diagonally opposed pyrrole nitrogens. The planar intersections are presented as colored lines through a representative heme molecule in Fig. 6A. The intersection for Mb (light cyan) eclipses the pyrrole nitrogens, whereas Hb-C (dark cyan) nearly eclipses the opposite pyrrole nitrogens, and Lba (orange), riceHb1 (purple), and Ngb (yellow), are staggered to various degrees. However, the azimuthal angles for eHb (pink), HbTn (gray), the crystal structure of SynHb
Figure 6. Proximal histidine azimuthal angles of selected hemoglobins. *Color-coded bars* indicate the line of intersection between the least squares planes of the heme and proximal histidine for each hxHb of known structure along with Mb, Lba, and four cytochrome $b_5$ proteins for comparison. A, Mb (*light cyan*) eclipses the pyrrole nitrogens, whereas Hb-C (*dark cyan*) nearly eclipses the opposite pyrrole nitrogens. Lba (*orange*), riceHb1 (*purple*), and Ngb (*yellow*) are staggered to various degrees. B, the crystal (*red*) and NMR (*blue*) structures of SynHb, eHb (*pink*), and HbTn (*gray*) are also staggered, but the inter-sections of their planes do not pass through the iron atom due to proximal histidine tilts. C, All four cytochrome $b_5$ proteins are staggered, with CSO (*dark green*) and bCYT (*dark blue*) slightly off-center compared with HSO (*light green*) and rCYT (*light blue*), also due to slight proximal histidine tilts.
(red), and the NMR structure (blue) are staggered but do not pass through the iron atom, as shown in Fig. 6B. This is due to the tilt of the proximal histidine imidazole plane. The tilts for all of the proteins in Fig. 6A, as delineated in the first column of Table III, are within about 2° of perpendicular to the heme plane. However, the proteins in Fig. 6B have tilts of 7° to 14° from the normal to the heme plane. Fig. 6C shows that all four of the cytochrome b₅ proximal histidines are also staggered. Again, plane intersections for CSO (dark green) and bCYT (dark blue) are somewhat off center due to small proximal histidine tilts, as opposed to HSO (light green) and rCYT (light blue).

Table 2-3. Heme pocket stereochemistry

<table>
<thead>
<tr>
<th>Protein</th>
<th>Proximal His tilt angle a (deg.)</th>
<th>Distal His tilt angle a (deg.)</th>
<th>Proximal Fe–His(Nε2) bond length (Å)</th>
<th>Distal Fe–His(Nε2) bond length (Å)</th>
<th>Fe-to-heme Plane b distance (Å)</th>
<th>PDB (chain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb</td>
<td>88.9</td>
<td>2.19</td>
<td>-0.16</td>
<td>2MBW (chain)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lba</td>
<td>87.9</td>
<td>2.21</td>
<td>0.10</td>
<td>1BIN (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb-C</td>
<td>89.3</td>
<td>77.0</td>
<td>2.07</td>
<td>2.08</td>
<td>0.07</td>
<td>1HLB</td>
</tr>
<tr>
<td>riceHb1</td>
<td>88.4</td>
<td>87.1</td>
<td>2.08</td>
<td>2.08</td>
<td>0.09</td>
<td>1D8U (A)</td>
</tr>
<tr>
<td>HbTn</td>
<td>76.8</td>
<td>73.0</td>
<td>2.00</td>
<td>1.99</td>
<td>0.08</td>
<td>1LA6 (B)</td>
</tr>
<tr>
<td>eHb</td>
<td>83.1</td>
<td>67.6</td>
<td>2.11</td>
<td>2.18</td>
<td>0.03</td>
<td>1NS6 (A)</td>
</tr>
<tr>
<td>Ngb</td>
<td>88.5</td>
<td>81.8</td>
<td>2.08</td>
<td>2.15</td>
<td>0.04</td>
<td>1OJ6 (A)</td>
</tr>
<tr>
<td>SynHb</td>
<td>79.0</td>
<td>72.7</td>
<td>2.11</td>
<td>1.99</td>
<td>-0.01</td>
<td>1RTX</td>
</tr>
<tr>
<td>bCYT</td>
<td>85.9</td>
<td>86.3</td>
<td>2.00</td>
<td>2.07</td>
<td>0.02</td>
<td>1CYO</td>
</tr>
<tr>
<td>rCYT</td>
<td>89.1</td>
<td>81.3</td>
<td>1.82</td>
<td>1.92</td>
<td>0.10</td>
<td>1B5M</td>
</tr>
<tr>
<td>HSO</td>
<td>87.0</td>
<td>86.3</td>
<td>2.01</td>
<td>2.01</td>
<td>-0.02</td>
<td>1MJ4</td>
</tr>
<tr>
<td>CSO</td>
<td>86.5</td>
<td>89.3</td>
<td>2.07</td>
<td>2.05</td>
<td>0.03</td>
<td>1SOX (A)</td>
</tr>
</tbody>
</table>

a Planes are defined as the least squares plane, determined using the 24 atom porphyrin macrocycle of the heme and the 5 atom imidazole ring of the histidine. The angle given is the acute angle between the two planes.

b Negative distance indicates that the iron is on the proximal side of the heme.
2.5. Discussion

The crystal structure of SynHb does not fall within the deviation of the NMR models and therefore reveals a separate conformation for this protein. Along with the difference in structural methods used, the covalent link between His117 and the heme prosthetic group is a likely cause for these conformational changes. Although the function of this covalent linkage is not yet understood, its structural implications are evident from a comparison of the crystal and NMR structures. Potential functional significance resulting from the covalent linkage includes inward movement of helices and heme pocket residues toward the heme resulting in decreased solvent accessibility and increased protein stability due to prevention of heme loss (27).

2.5.1. Implications for Ligand Binding in trHbs

SynHb is similar in overall fold to other trHbs with the obvious exception of intramolecular heme iron coordination. This is likely to account for many structural differences with other trHbs, such as the downward shift of the B helix that moves Tyr(B10) into position to stabilize the ligand in other trHbs. Because Tyr(B10) in SynHb is also predicted to stabilize bound ligands (24), a similar downward shift of the B helix is expected upon ligand binding in SynHb. His46 is predicted to be uninvolved in ligand stabilization (24, 25). This suggests that ligand binding in SynHb will cause a shift in the position of the E helix that moves this side chain outwards into the solvent. This shift could also move the Glu(E7) and Glu(E11) residues into position to contribute to a hydrogen bonding network around the bound ligand as seen in the ligand bound structures of other trHbs (2).
Although these structural changes would cause ligand-bound SynHb to resemble other ligand-bound trHbs, the ligand-free structures of CtrHb, PtrHb, and trHbN must be different than that of SynHb. Of these three, only CtrHb is capable of hexacoordination, which occurs only at alkaline pH by the Tyr(B10) side chain (55). This suggests that the B helix rather than the E helix in CtrHb moves closer to the heme in the absence of exogenous ligands. Structures of PtrHb and trHbN in the pentacoordinate state would be different from both SynHb and CtrHb in that neither exhibits hexacoordination.

The conventional "histidine-gate" path for ligand binding is blocked in trHbs (2). It has been proposed that the tunnel found in the previous trHb structures serves as an alternative diffusion path for ligands (43). However, there is no protein matrix tunnel in SynHb. Therefore, three possibilities exist for ligand entry and exit in SynHb. 1) Ligands enter and exit through the solvent face of the heme pocket. 2) Ligands enter through the solvent face of the heme pocket, and then a tunnel forms in the ligand bound state that serves as a route for exit. 3) Ligand entry and exit pathways are not evident in these structures. The solvent accessibility around the propionates would be an appropriate route for ligand access in support of the first two possibilities. In the second case, formation of a tunnel after ligand binding followed by its blockage when the ligand dissociates could serve to trap ligands near the heme pocket, providing a barrier to ligand escape, and enhancing geminate ligand recombination. This possibility is supported by the observation of room temperature CO geminate recombination in SynHb (24, 45).
2.5.2. Implications for Ligand Binding in hxHbs

The data in Table III show some of the stereochemical factors likely to contribute to ligand affinity in hxHbs. In general, it is thought that staggering of the azimuthal angle of the proximal histidine away from the heme pyrrole nitrogens increases ligand affinity by allowing the iron to move into the plane of the porphyrin ring (56, 57). This holds true for SynHb, riceHb1, and Lba, which are all staggered and which all have relatively high oxygen affinities compared with Mb (17, 22, 58).

Columns 4 and 5 in Table III indicate differences in Fe-His(Nε2) bond lengths, with the cytochromes in most cases exhibiting shorter bond lengths between the heme iron and the proximal and distal histidines. The longer bond lengths found in the hxHbs decrease the bond strength and are therefore likely to be a factor enabling distal histidine dissociation and subsequent ligand binding in the hxHbs. Bond strength is also influenced by the degree of histidine tilt (with larger tilts decreasing the strength of the bond) (48). The hxHbs also have generally larger tilt angles for both the proximal and distal histidines compared with cytochrome b₅, which is likely to be another element contributing to their ability to bind ligands.

2.6. Conclusions

Gross features of the crystal structure of SynHb are similar to the NMR structure, but many specific differences are also evident. Many of these differences could be due to the covalent linkage between His117 and the heme-2-vinyl group that is present in the crystal structure
but not in the NMR structure. In comparison to other trHbs, hexacoordination in SynHb forces many side chains important for ligand stabilization to reside in contact with solvent in the absence of exogenous ligands. Hexacoordination in SynHb results from a different tertiary arrangement of the B and E helices compared with other hxHbs, as well as many specific differences in heme pocket stereochemistry, leading to alternative mechanisms for regulating ligand binding. Examination of SynHb in the context of other hxHbs and cytochrome b5 reveals that destabilization of the distal His-heme iron bond to allow for exogenous ligand binding could be indicated by increased bond lengths and tilt angles characterizing intramolecular coordination in the hxHbs.

2.7. Abbreviations

Hb, hemoglobin;
Mb, myoglobin;
SynHb, Synechocystis sp. hemoglobin;
hxHb, hexacoordinate hemoglobin;
trHb, truncated hemoglobin;
r.m.s.d., root mean square deviation;
CtrHb, C. eugametos truncated hemoglobin;
PtrHb, P. caudatum truncated hemoglobin;
trHbN, M. tuberculosis truncated hemoglobin;
riceHb1, rice hemoglobin 1;
Ngb, neuroglobin;
Hb-C, hemichrome chain from *C. arenicola*;
eHb, hemichrome chain from horse methemoglobin;
HbTn, hemichrome chain from *T. newnesi*;
bCYT, bovine cytochrome *b*$_{5}$;
rCYT, rat cytochrome *b*$_{5}$;
HSO, cytochrome *b*$_{5}$ domain from human sulfite oxidase;
CSO, cytochrome *b*$_{5}$ domain from chicken liver sulfite oxidase;
MES, 4-morpholinoethanesulfonic acid;
Lba, leghemoglobin.

### 2.8. Acknowledgments

This work was supported by the National Science Foundation (Grant MCB-0077890), the U. S. Department of Agriculture (Grant 2002 35318 1217), and the Iowa State University Plant Sciences Institute. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U. S. Department of Energy, Office of Basic Energy Sciences. The Stanford Synchrotron Radiation Laboratory, Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. Portions of this research were carried out at the University of Iowa Protein Crystallography Facility. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be
hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank David S. Doty for computational assistance.

## 2.9. References


CHAPTER 3. CONFORMATIONAL CHANGES UPON LIGAND BINDING IN SYNECHOCYSTIS HEMOGLOBIN

This chapter contains work contributing to a paper published in the Journal of Molecular Biology. The author’s contributions are described and include refinement and analysis of the structure of cyanide bound Synechocystis hemoglobin.

3.1. Introduction

Since the characterization of myoglobin (Mb), a wide diversity in form and function has been discovered in new hemoglobins (Hbs) from many different species. For example, truncated hemoglobins (trHbs) are a distinct family of Hbs that have been found in both eubacteria and eukaryotes. They are 20-40 residues shorter than non-vertebrate hemoglobins, and have a 2-on-2 α-helical fold, rather than the typical 3-on-3 fold [1]. Hexacoordinate hemoglobins (hxHbs) are another unique group of proteins in which an endogenous amino acid coordinates the iron atom on the distal side of the heme. However, hxHbs are still capable of exogenous ligand binding. Such reversible ligand binding distinguishes the hxHbs from cytochrome $b_5$ and denatured Hbs that cannot bind ligands in their hexacoordinate state [2]. Such unusual characteristics are likely to result in differing functional roles for these groups of Hbs compared to traditional oxygen storage and transport proteins like myoglobin [3].

---

The Hb from the cyanobacteria *Synechocystis* sp. (SynHb) is interesting in that it is both hexacoordinate and truncated [4]. In addition to the distal histidine and the proximal histidine, SynHb also has a third histidine coordinating the heme [5]. His$^{117}$ is covalently linked to the heme 2-vinyl, presenting another novel and intriguing aspect to this protein. In this study, the cyanide ligand bound structure of *Synechocystis* hemoglobin (SynHb:CN) has been solved to complement the study of the ligandless structure [6] discussed in Chapter 2, and to provide information on the mechanism of ligand binding in this protein. Determination of the x-ray crystal structures of SynHb in both the endogenous (His$^{46}$) and exogenous (CN) ligand bound states allows for the first comparison of corresponding conformational changes in a trHb or hxHb.

### 3.2. Methods

The protein was expressed and purified as described previously [3, 7]. The oxidized protein was crystallized by the vapor diffusion method using a well buffer consisting of 1.5 M sodium citrate (pH 6.0), 0.2 M ammonium citrate, 20% (w/v) sucrose, 3% (v/v) dioxane and 25 mM sodium cyanide at 4 °C. Protein crystals took several weeks to grow and were not stable at room temperature. Diffraction data were collected at 100 K on a Rigaku/MSC home source generator, and processed using d*TREK [8]. The structure was solved by molecular replacement using the *C. eugametos* hemoglobin (CtrHb) structure (1DLY.pdb) and the program CNS [9]. Refinement and model building were performed with CNS, CCP4 [10], and O [11], with a final R-factor of 17.2% and R-free of 19.9%. The final model has been
deposited in the Protein Data Bank (1S69.pdb) and statistics are provided in Table 1. A ribbon diagram of the protein and representative density are shown in Figure 1.

<table>
<thead>
<tr>
<th>Table 3-1. Data collection and refinement statistics for SynHb:CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Reflections (unique/total)</td>
</tr>
<tr>
<td>R_{merge} (%)</td>
</tr>
<tr>
<td>Ave I/σ(I)</td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit Cell (Å)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Resides/Waters</td>
</tr>
<tr>
<td>R_{crystal} (%)</td>
</tr>
<tr>
<td>R_{free} (%)</td>
</tr>
<tr>
<td>r.m.s.d. from ideal geometry</td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
</tr>
<tr>
<td>Bond angles (deg.)</td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
</tr>
<tr>
<td>Most favored</td>
</tr>
<tr>
<td>Addn allowed</td>
</tr>
</tbody>
</table>

Numbers in parentheses are outer-shell statistics.
Figure 1. Ribbon diagram and electron density of the *Syn*Hb:CN structure. Stereoviews of A) the trHb fold, and B) representative electron density in the heme pocket of cyanide bound *Syn*Hb.
3.3. Results and Discussion

Comparison of endogenously coordinated SynHb with exogenously liganded SynHb:CN reveals a conformational shift unprecedented among hemoglobins [12]. Upon exogenous ligand binding of cyanide, the A- and upper B-helices shift towards the heme, while the E-helix and EF loop swing outward away from the heme. Figure 2 illustrates the magnitude of this change by showing the rms deviation between the two structures of SynHb. The heme position also differs in the ligand bound form, where it is tilted and recessed within the distal pocket.

Figure 2. Structural comparison of SynHb and SynHb:CN. Cross-eyed stereo overlay of SynHb (grey) and SynHb:CN (colored by Ca rms deviation from SynHb, with blue indicating low rmsd and red indicating high rmsd). The heme and proximal and distal histidines are also shown (light grey for SynHb, dark grey for SynHb:CN). Note that the distal histidine swings out of the distal pocket into the solvent, as seen on the right hand side of each panel.
pocket as compared to the hexacoordinate form. These movements bring Glu(E7), Glu(E11), and Tyr(B10) into position to stabilize the bound ligand, as seen in Figure 3. In addition, the unusual covalent link is seen in both crystal structures, as can also be seen in Figure 3.

Figure 3. **Heme pocket overlay.** Cross-eyed stereo overlay of selected heme pocket residues of SynHb (red) and SynHb:CN (blue), aligned using the heme group and labeled by helix position.

The hexicoordinate state of SynHb is stabilized by hydrogen bonds between the heme D-propionate and Lys\(^{42}\) of the E-helix and Tyr\(^{65}\) of the F-helix, as shown in Figure 4A. The F-helix is stabilized by hydrogen bonding between Arg\(^{64}\) of the E-helix and Asp\(^{62}\) of the EF loop.
Figure 4. **Hydrogen bonding networks in SynHb and SynHb:CN.** Cross-eyed stereo view of the hinge residues involved in hydrogen bonding. **A.** SynHb (red) before ligand binding, with H-bonds shown in green and residues labeled according to amino acid and protein position. **B.** SynHb:CN (blue), after ligand binding with H-bonds shown in orange.

The displacement of the distal histidine, His$^{46}$, which makes ligand binding possible, occurs with a rotation and outward movement of the E-helix, accompanied by a partial unraveling of
the F-helix. This conformation is stabilized by a network of hydrogen bonds between the heme D-propionate, His$^{46}$, and Tyr$^{61}$ of the EF loop, as depicted in Figure 4B.

Meanwhile, the unraveling of the F-helix allows the heme D-propionate to hydrogen bond with the side chain and backbone of Arg$^{64}$, while Glu$^{68}$ of the F-helix hydrogen bonds with the backbone of Tyr$^{65}$. Thus, the EF loop and N-terminal end of the F-helix act as a sort of “hinge,” flexibly allowing the globin movements needed to accommodate ligand binding, yet regulating this flexibility through a system of hydrogen bonds to stabilize the end points [12].

In addition, a solvent tunnel, though not observed in the endogenously hexacoordinate structure, is observed in the SynHb:CN structure. The tunnel is similar to, but less extensive than the tunnels observed in other trHbs such as *Chlamydomonas eugametos* hemoglobin (CtrHb), as shown in Figure 5 [13, 14]. It is possible that this tunnel may function to increase ligand accessibility to the heme pocket, contributing to the high rate of ligand binding characteristic of the trHbs. However, it appears that there is also direct solvent accessibility along the propionyl face of the heme (Figure 6).
Figure 5. **Solvent tunnels in trHbs.** The solvent tunnel in A) SynHb:CN is discontinuous, while the tunnel in B) CtrHb is continuous.
Figure 6. Solvent accessibility near the heme propionates. Solvent accessibility increases upon ligand binding as seen in this stereoview comparison of A) SynHb and B) SynHb:CN.
3.4. Conclusions

The use of x-ray crystallography affords the ability to visualize conformational changes due to ligand-binding reactions at a molecular level. This provides a context for understanding complicated kinetic data and insight into the functional roles of new and unusual Hbs. One such hemoglobin, $Syn$Hb, undergoes a conformational shift upon exogenous ligand binding unprecedented among hemoglobins for which both structures with and without ligands are known. However, the flexibility of this protein is carefully controlled by a network of hydrogen bonds and a novel third covalent link to the heme.

The structural observations made here reveal the large differences that can occur among proteins with the same general fold but differing primary sequences. In myoglobin, there is little conformational change upon ligand binding, and the bound ligand is coordinated by the distal histidine. In $Syn$Hb, on the other hand, the large conformational change that must occur to move the distal histidine out of the way of the ligand, coupled with the strong network of distal pocket residues that coordinate the ligand once bound, result in an entirely different mode of ligand binding that will impact the functioning of this hemoglobin. The magnitude of the conformational change that takes place in $Syn$Hb upon exogenous ligand binding may enable interaction with another protein as a signaling mechanism. Alternatively, the large structural changes may serve to open up ligand tunnels, thereby favorably affecting ligand binding rates. The extensive ligand coordination might facilitate ligand storage, or chemical reactions at the heme iron. For now, all that is certain is that $Syn$Hb is an example of unprecedented hemoglobin folding changes upon ligand binding.
3.5. References


CHAPTER 4. COVALENT HEME ATTACHMENT IN SYNECHOCYSTIS HEMOGLOBIN IS REQUIRED TO PREVENT FERROUS HEME DISSOCIATION

A paper submitted to Protein Science

Julie A. Hoy\textsuperscript{8,9}, Benoit J. Smagghe\textsuperscript{9}, Puspita Halder\textsuperscript{9}, and Mark S. Hargrove\textsuperscript{9,10}

4.1. Abstract

\textit{Synechocystis} hemoglobin contains an unprecedented covalent bond between a non-axial histidine side chain (H117) and the heme 2-vinyl. This bond has been previously shown to stabilize the ferric protein against denaturation, and also to affect the kinetics of cyanide association. However, it is unclear why \textit{Synechocystis} hemoglobin would require the additional degree of stabilization accompanying the His\textsuperscript{117}-heme 2-vinyl bond because it also displays endogenous \textit{bis}-histidyl axial heme coordination, which should greatly assist heme retention. Furthermore, the mechanism by which the His\textsuperscript{117}-heme 2-vinyl bond affects ligand binding has not been reported, nor has any investigation of the role of this bond on the structure and function of the protein in the ferrous oxidation state.

\textsuperscript{7}Manuscript in press.
\textsuperscript{8}Primary Author and Researcher
\textsuperscript{9}Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011.
\textsuperscript{10}To whom correspondence should be addressed
Here we report an investigation of the role of the *Synechocystis* hemoglobin His$^{117}$-heme 2-vinyl bond on structure, heme coordination, exogenous ligand binding, and stability in both the ferrous and ferric oxidation states. Our results reveal that hexacoordinate *Synechocystis* hemoglobin lacking this bond is less stable in the ferrous oxidation state than the ferric, which is surprising in light of our understanding of pentacoordinate Hb stability, in which the ferric protein is always less stable. It is also demonstrated that removal of the His$^{117}$-heme 2-vinyl bond increases the affinity constant for intramolecular histidine coordination in the ferric oxidation state, thus presenting greater competition for the ligand binding site and lowering the observed rate and affinity constants for exogenous ligands.

### 4.2. Introduction

The past few decades have revealed that hemoglobins are found not only in vertebrates, but also in archaea, bacteria, animals, and plants (Weber and Vinogradov 2001; Wittenberg et al. 2002; Kundu 2003b; Wu et al. 2003). A common function of vertebrate Hb is to transport oxygen from the lungs to tissues through the vascular system. However, Hbs have many other potential functions including ligand scavenging, detoxification, small molecule sensing, signal transduction, and gene regulation (Appleby 1984; Minning et al. 1999; Burr et al. 2000; Hou et al. 2000; Poole and Hughes 2000; Wittenberg et al. 2002; Hankeln et al. 2005), while the functions of many Hbs remain to be identified.

A number of newly discovered Hbs display unorthodox structural characteristics. Plants, animals, and some bacteria have Hbs that carry out reversible intramolecular coordination of
the heme iron by a "distal" histidine side chain (Burmester et al. 2000; Trent et al. 2001b; Weber and Vinogradov 2001; Burmester et al. 2002; Trent and Hargrove 2002; Wittenberg et al. 2002; Kundu 2003b; Wu et al. 2003). These "hexacoordinate" Hbs (hxHbs) are likely the scaffold from which oxygen transport Hbs evolved in plants and animals (Vinogradov et al. 2006). The “proximal” histidine is tightly coordinated to the heme iron (as in oxygen transport Hbs), but the distal histidine reversibly coordinates the ligand binding site and competes with the binding of exogenous ligands like oxygen and nitric oxide. Recent studies in plant and animal systems have described structural components that regulate affinity constants for distal histidine binding, showing that the strength of histidine coordination greatly affects affinity and kinetic constants for exogenous ligand binding (Smagghe 2006; Smagghe et al. 2006).

Another nontraditional group of Hbs, the "truncated” Hbs (trHbs), is found in eubacteria, bacteria, single-celled eukaryotes, and plants (Couture et al. 2000; Scott and Lecomte 2000; Hvitved et al. 2001; Watts et al. 2001; Wittenberg et al. 2002). They can be 20–40 residues shorter than the majority of Hbs and comprise an abbreviated globin architecture surrounding the heme that forms a “2-on-2” fold (Pesce et al. 2000; Wittenberg et al. 2002). Cyanobacteria contain a trHb that is unique even within this family. The Hb from *Synechocystis* sp. PCC 6803 (SynHb) is hexacoordinate, truncated, and bears an unprecedented covalent bond between the non-axial His₁₁⁷ and a heme porphyrin vinyl group (Vu et al. 2002).
Extensive research has investigated the role of this covalent heme-protein link since its 2002 discovery (Vu et al. 2002). Prior to identification of this bond, the existence of two forms of the protein had been noted, and experiments were carefully performed only on the form of the protein that, in retrospect, lacked the covalent bond (Lecomte et al. 2001). These experiments included observation of slow heme release (Scott and Lecomte 2000; Lecomte et al. 2001), thermal denaturation (Lecomte et al. 2001), and measurement of the electrochemical midpoint potential (Lecomte et al. 2001). After the covalent bond was discovered as the source of heterogeneity, partial NMR structural analysis indicated little global conformational change among the ferric tertiary structures of the two forms of wild type protein and the H117A mutant protein (which is incapable of forming the covalent bond) (Falzone et al. 2002; Vu et al. 2004a; Vu et al. 2004b). Thermal and acid denaturation studies of the ferric forms of these three proteins found the wild type protein with the covalent bond to be more stable than the mutant or wild type lacking the bond decreased compared to wild type protein with an undetermined covalent bond contribution (Vu et al. 2004a).

The above studies suggested that the His^{117}\text{--}heme vinyl bond does not affect the overall conformation of the ferric protein, but does affect protein stability, heme affinity, and the ability of SynHb to bind exogenous ligands in the ferric oxidation state. The most plausible explanation for the lowered cyanide affinity seen in the mutant lacking the covalent bond is increased competition for the binding site due to an increase in the equilibrium constant for distal histidine coordination to the ferric heme iron. Since heme affinity (and thus protein stability) would increase with this tighter heme coordination, it is odd that the His^{117}\text{--}heme
vinyl bond would be necessary to stabilize the protein in the ferric oxidation state. Furthermore, small molecule model-heme studies affirm that bis-histidyl heme coordination may be tighter in the ferric than the ferrous oxidation state (Nesset 1996; Safo and Debrunner 1997). If this chemical principle holds true in SynHb, the protein should be more susceptible to heme loss and subsequent unfolding in the ferrous oxidation state compared to the ferric. Consequently, the largest effects of removal of the His^{117}-heme vinyl covalent bond might be exposed in the ferrous protein, a factor not investigated previously.

Here we demonstrate that removal of the His^{117}-heme vinyl covalent bond in SynHb affects distal histidine coordination of the heme iron by facilitating hexacoordination in the ferric oxidation state by a factor of ~10 compared to the wild-type protein. This is the root of the slower binding kinetics for ferric ligands in SynH117A, and also results in lowered equilibrium affinity constants for azide and cyanide. Surprisingly, the effects of the H117A mutation on protein stability and heme affinity are much more pronounced in the ferrous oxidation state compared to the ferric. This extraordinary observation is counter to our understanding of heme affinity in pentacoordinate hemoglobins (Hargrove and Olson 1996), but is consistent with studies of coordination affinity in bis-histidyl model heme compounds. Finally, crystal structures of wild type and H117A SynHb affirm that the effects of this covalent bond on protein chemistry are not obvious in protein structure in either the ferric or ferrous oxidation states.
4.3. Results

4.3.1. Crystal structures of ferric and ferrous SynHb and SynH117A

Figure 1 presents an overlay of the crystal structures of SynHb (with the covalent bond) and SynH117A in both ferric and ferrous oxidation states. The backbone superposition of the ferric structures in panel A shows minimal backbone deviation (rmsd = 0.45 Å), but small changes in the orientation of the heme and axial histidines, in agreement with the partial structural observations obtained by NMR (Vu et al. 2004b). The superposition in panel B shows that, relative to the heme, mutation of H117 causes outward movement of the H-helix with minor movements of the F- and G- helices (rmsd = 1.08 Å). Similar results are found in the comparisons of the ferrous forms as shown in Figure 1C (backbone superposition) and Figure 1D (heme superposition), with additional movement of the short A-helix. Therefore, the His$^{117}$-heme vinyl bond has minimal effects on overall structure compared to the wild-type protein in both the ferric and ferrous oxidation states. However, the H117A mutation and the consequent movement of the H-helix opens space on the proximal side of the heme that allows solvent entry into heme pocket from the protein surface in both oxidation states (Figure 1E and F). A closer examination of the changes in heme stereochemistry is summarized in the Supplementary Materials. It appears that there are no obvious structural differences between wild type and SynH117A in either the ferrous or ferric oxidation states that rationalize the observed functional differences between these two proteins or the presence of the covalent bond between His$^{117}$ and the heme 2-vinyl.
Figure 1. *SynHb* (grey) versus *SynH117A* (colored by rmsd from *SynHb*, with dark blue corresponding to smallest deviation). Ferric structures overlaid by A) backbone, and B) heme. Ferrous structures overlaid by C) backbone, and D) heme. Axial histidines are shown, along with the amino acid at position 117. Mutation causes a solvent pocket to open on the proximal side of the heme, as seen in the superposition of *SynHb* (red) with *SynH117A* (blue) in the E) ferric, and F) ferrous forms. The amino acid at position 117 is shown for each protein. The x-ray crystal structure of the ferric form of *SynHb* was solved previously (Hoy et al. 2004).
4.3.2. Ligand binding in the ferrous oxidation state

The reactions associated with ligand binding in HxHbs are shown in Scheme 1. For the case of negligible ligand dissociation, the equation for the observed rate of the exogenous ligand binding reaction assuming a steady state equilibrium between Hb_H and Hb_P is given by Equation 1.

\[
Hb_H \xrightleftharpoons{k_{-H}} \xleftrightarrow{k_H} Hb_P \xrightleftharpoons{k_L} Hb_L
\]

Scheme 1

\[
k_{obs,L} = \frac{k_{-H}k'_L[L]}{k_{-H} + k_{-H} + k'_L[L]} \quad \text{Equation 1}
\]

Here, H, P, and L refer to Hb in the hexacoordinate, pentacoordinate, and ligand bound states respectively. The rate constants k_{-H} and k_H are the dissociation and association rates for the distal histidine, while k'_L[L] and k_L are the ligand bimolecular association and dissociation rate constants.

Flash photolysis was used to measure the CO bimolecular association rate constant of the SynH117A mutant protein for comparison to that published previously for wild type SynHb (Hvitved et al. 2001; Smagghe et al. 2006); this value, k_Co = 150 \mu M^{-1}s^{-1}, is only slightly greater than that of the wild type protein (see Table 1 for values and Supplemental Data for more information). The influence of intramolecular histidine coordination becomes apparent when comparing kinetic experiments initiated by flash photolysis to those measured by rapid mixing (Smagghe et al. 2006). Figure 2A shows time courses for SynH117A binding with CO following rapid mixing. These data are consistent with those of wild type SynHb in that the total expected amplitude is observed for the reaction at each [CO], and the observed rate
constant increases with [CO] (Smagghe et al. 2006). This indicates that $k'_{\text{CO}} > k_{\text{H2}}$ and $k_{\text{H2}}$ (where $k_{\text{H2}}$ and $k_{\text{H2}}$ are the dissociation and association constants for the distal histidine in the ferrous state), and that the fraction of hexacoordinate Hb is near 1. Both of these observations are consistent with wild type ferrous SynHb (Smagghe et al. 2006).
The [CO] dependence of the observed rate constants (from rapid mixing) in Figure 2A is plotted in Figure 2B. The asymptote approached in this curve is \( k_{\text{H2}} \), and a fit to Equation 1 provides \( k_{\text{H2}} \) (Hargrove 2000). These values are reported in Table 1, and reflect only modest (~3-fold) increases for each in the mutant protein. Since both \( k_{\text{H2}} \) and \( k_{-\text{H2}} \) increase to approximately the same degree in the mutant compared to the wild type protein, the resulting values of \( K_{\text{H2}} \) (the association equilibrium constant for the distal histidine, equivalent to \( k_{\text{H}}/k_{-\text{H}} \)) are nearly the same for each. These results demonstrate that the His\(^{117}\)-heme vinyl bond has only minor effects on hexacoordination and exogenous ligand binding in ferrous SynHb.

### 4.3.3. Spectroelectrochemistry

Electrochemistry can be used to measure changes that occur in the strength of hexacoordination upon change in oxidation state (Halder 2006; Smagghe 2006). In hexacoordinate Hbs, a shift to more-negative reduction potentials (compared to pentacoordinate Hbs) reveals that intramolecular histidine coordination of the heme iron is tighter in the ferric compared to the ferrous oxidation state (Dewilde et al. 2001; Halder 2006; Smagghe 2006). Since previous work reported the midpoint reduction potential for only the wild type protein without the covalent bond (Lecomte et al. 2001), we examined the midpoint potential for SynHb with the covalent bond and the SynH117A protein to determine if the His\(^{117}\)-heme vinyl bond influences hexacoordination differentially in the ferric versus the ferrous oxidation state. The fraction of reduced protein for both SynHb and the H117A mutant protein as a function of the observed cell potential are shown in Figure 3A. The midpoint potential value (given in Table 1) is vastly lower for SynH117A (-280 mV) than for SynHb (-195 mV).
The relationship between the equilibrium constant for hexacoordination ($K_{H}$) in the ferric (H3) and ferrous (H2) forms and the midpoint reduction potential ($E_{\text{mid}}$) for the pentacoordinate (pent) and hexacoordinate (hex) forms is given by Equation 2 (Halder 2006).

$$\left(\frac{1+K_{H,3}}{1+K_{H,2}}\right) = e^{-\frac{nF}{RT}\left(E_{\text{mid},\text{hex}}-E_{\text{mid},\text{pent}}\right)}$$  \hspace{1cm} \text{Equation 2.}

If we assume that the difference in midpoint potentials between the wild type and H117A proteins is due only to histidine coordination (that is, $E_{\text{mid},\text{pent}}$ is the same in both proteins), an estimate of the effect of the covalent bond on hexacoordination ($K_{H,3}/K_{H,2}$) can be calculated by Equation 3, which is derived by division of Equation 2 for the wild type protein by that for $\text{SynH}117\text{A}$ (Halder 2006).

$$\left(\frac{1+K_{H,3,\text{SynH}117\text{A}}}{1+K_{H,2,\text{SynH}117\text{A}}}\right) = e^{-\frac{nF}{RT}\left(E_{\text{mid,\text{SynH}117\text{A}}}-E_{\text{mid,\text{SynHb}}}\right)}$$  \hspace{1cm} \text{Equation 3.}

This equation correlates the effects of hexacoordination in each oxidation state with the observed reduction potential in each protein. An increased value of $K_{H,3}/K_{H,2}$ is predicted for the mutant based on the large negative shift in reduction potential observed in Figure 3A. In fact, using the measured values given in Table 1 to calculate the right hand side of Equation 3 gives a value of 27, affirming the increase in ferric hexacoordination upon removal of the covalent bond.
Figure 3. Electrochemistry and ferric binding data for SynHb (filled circles) and SynH117A (open circles). A) SynHb has a higher reduction midpoint potential than SynH117A. B) Optical cyanide binding kinetics showing slower ligand binding to SynH117A. C) SynHb and SynH117A equilibrium binding curves. Values found from the fits (solid lines) in each panel are given in Table 1.
As the covalent bond is absent in both oxidation states of the H117A mutant protein, the large negative shift resulting from His<sup>117</sup> mutation results from either a change in overall globin electrostatic properties, or from a targeted effect on heme coordination by the two axial histidines. The former is difficult to judge since previous work has demonstrated that $E_{\text{mid,pent}}$ is not readily obtainable by mutation of the $\text{SynHb}$ distal histidine<sup>(34)</sup>. However, since the crystal structures of $\text{SynHb}$ and the H117A mutant show only minor deviations in comparison to each other in either oxidation state, the assumption will be made that the covalent bond exhibits only the latter affect, on heme coordination.

This assumption has the benefit of allowing $E_{\text{mid,pent}}$ to cancel out of the equation that results from the division of Equation 4 as applied to the wild type protein by the same equation as applied to the mutant. This is given by Equation 5:

$$
\frac{1 + K_{H3,\text{SynH}117A}}{1 + K_{H2,\text{SynH}117A}} = e^{-\left\{ \frac{nF(E_{\text{mid,hex}} - E_{\text{mid,pent}})}{RT} \right\}}
$$

Equation 5.

The midpoint potential in this equation reflects the supposedly unchanged reduction potential of the "pentacoordinate" version of the protein augmented by hexacoordination in each oxidation state. An increased value of $K_{H3}$ is predicted for the mutant based on the large negative shift in reduction potential. In fact, using the measured values given in Table 2 to
calculate the right hand side of Equation 5 gives a value of 27, affirming the increase in ferric hexacoordination upon mutation of the covalent bond.

<table>
<thead>
<tr>
<th>Protein</th>
<th>k\textsubscript{CO} M\textsuperscript{-1}s\textsuperscript{-1} (x10\textsuperscript{-6})</th>
<th>k\textsubscript{H\textsubscript{2}} s\textsuperscript{-1}</th>
<th>k\textsubscript{H\textsubscript{2}} s\textsuperscript{-1}</th>
<th>K\textsubscript{H\textsubscript{2}}</th>
<th>E\textsubscript{mid} mV</th>
<th>k\textsubscript{CN,obs} M\textsuperscript{-1}s\textsuperscript{-1} (x10\textsuperscript{-3})</th>
<th>K\textsubscript{CN,obs} M\textsuperscript{-1} (x10\textsuperscript{-3})</th>
<th>K\textsubscript{N\textsubscript{3},obs} M\textsuperscript{-1} (x10\textsuperscript{-3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>SynHb</td>
<td>90</td>
<td>14</td>
<td>4,200</td>
<td>300</td>
<td>-195</td>
<td>0.7</td>
<td>0.6</td>
<td>0.005</td>
</tr>
<tr>
<td>SynH117A</td>
<td>150</td>
<td>27</td>
<td>24,00</td>
<td>890</td>
<td>-280</td>
<td>0.05</td>
<td>0.06</td>
<td>NA</td>
</tr>
</tbody>
</table>

Values for wild type SynHb are taken from Smagghe et al (Smagghe et al. 2006).

### 4.3.4. Ferric ligand binding

The increase in K\textsubscript{H\textsubscript{3}}/K\textsubscript{H\textsubscript{2}} predicted above can be tested using equilibrium and kinetic ligand binding experiments. The effect of a change in K\textsubscript{H\textsubscript{3}} on these rate constants has been described previously in other systems (Trent et al. 2001b; Smagghe et al. 2006), and the equations are shown here for cyanide serving as the ligand.

\[
K_{CN,\text{observed}} = \frac{K_{CN,\text{pent}}}{1 + K_{H\textsubscript{3}}} \quad \text{Equation 4.}
\]

\[
k_{CN,\text{observed}} = \frac{k_{CN,\text{pent}}[CN]}{1 + K_{H\textsubscript{3}}} \quad \text{Equation 5.}
\]

K\textsubscript{CN,obs} and k\textsubscript{CN,obs} are the observed equilibrium affinity and association rate constants (respectively). The right-hand side of each equation relates these values to the equilibrium affinity and association rate constants for the pentacoordinate state of the protein (K\textsubscript{CN,\text{pent}} \textsubscript{CN,\text{pent}}).
and \( k'_{\text{CN, pent}} \) along with the effect of hexacoordination under conditions of prior equilibrium (this assumes that \( k'_{\text{CN}}[\text{CN}] \ll k_{\text{H3}} \) and \( k_{\text{H3}} \)) (Espenson 2002).

Kinetic and equilibrium constants for azide and cyanide binding were measured for the H117A and wild type proteins (Figures 3B and C). The observed rate constants for cyanide association are linear and very slow, indicating that the bimolecular rate constant for the reaction with the pentacoordinate form of the Hb (\( k'_{\text{CN}} \)) is \( << k_{\text{H3}} \) and \( k_{\text{H3}} \), as required in the equations above (Smagghe et al. 2006). A linear fit to these data provides \( k_{\text{CN, obs}} \) for each protein showing that the value for \( \text{Syn} \text{H117A} \) is reduced \( \sim 10 \)-fold compared to the wild type protein (Table 1), which is consistent with the value reported previously (Vu et al. 2004a). This is accompanied by a proportionate decrease in the association equilibrium constant (\( K_{\text{CN, obs}} \)) for cyanide. Likewise, the wild-type protein binds azide (albeit with a much smaller \( K_{\text{N3, obs}} \) than for cyanide), but the mutant protein displayed no binding until concentrations higher than 4 M azide (data not shown).

### 4.3.5. GdmCl unfolding

The foremost hypothesis for the role of the His\(^{117}\)-heme vinyl bond is stabilization of the protein through heme retention. Vu et al. (Vu et al. 2004a) showed that in fact the ferric protein is less stable without the covalent bond (Nesset 1996; Safo and Debrunner 1997). However, while there is reason to believe that hxHbs might be even less stable in the ferrous oxidation state than the ferric, ferrous protein stability has not been examined. Therefore, we have measured GdmCl-induced unfolding of these proteins in both oxidation states. These
experiments are shown in Figures 4 and 5, with unfolding midpoints and \( \Delta G_0 \) values listed in Table 2.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Midpoint GdmCl (M)</th>
<th>( \Delta G_0 ) kcal/mol</th>
<th>Final Coordination State</th>
</tr>
</thead>
<tbody>
<tr>
<td>hhMb ferric</td>
<td>1.6</td>
<td>8.3</td>
<td>free heme</td>
</tr>
<tr>
<td>hhMb ferrous</td>
<td>2.4</td>
<td>7.9</td>
<td>free heme</td>
</tr>
<tr>
<td>SynHb ferric</td>
<td>3.0</td>
<td>11.5</td>
<td>pentacoordinate</td>
</tr>
<tr>
<td>SynHb ferrous</td>
<td>3.0</td>
<td>10.9</td>
<td>pentacoordinate</td>
</tr>
<tr>
<td>SynH117A ferric</td>
<td>1.6</td>
<td>6.1</td>
<td>free heme</td>
</tr>
<tr>
<td>SynH117A ferrous</td>
<td>1.1</td>
<td>3.8</td>
<td>free heme</td>
</tr>
</tbody>
</table>

Figure 4A provides a control with pentacoordinate horse heart Mb (hhMb), demonstrating that the ferrous protein is more stable than the ferric; mainly due to the tighter heme-proximal histidine bond in the ferrous protein (Hargrove et al. 1996a; Hargrove and Olson 1996; Hargrove et al. 1996b). GdmCl unfolding of Mb yields free heme and unfolded apoglobin (Hargrove and Olson 1996; Tang et al. 1998), evident from the absorbance spectra for ferrous and ferric hhMb at 5 M GdmCl (Figure 5A and B). Wild type SynHb, on the other hand, is equally stable in either oxidation state (Figure 4B); in fact, the spectral transition associated with GdmCl titration of wild type SynHb does not yield free heme, but rather pentacoordinate complexes in both oxidation states (Figure 5C and D). This is most obvious in the ferrous oxidation state (Figure 5D, inset), in which the spectrum in 7 M
GdmCl is a single broad absorbance band at 550 nm, characteristic of a ferrous pentacoordinate heme complex (Arredondo-Peter et al. 1997; Smagghe et al. 2006).

Figure 4. Guanidinium chloride denaturation. Ferric (3+, filled circles, left axis) versus ferrous (2+, open circles, right axis) protein denaturation curves for A) horse heart myoglobin, B) SynHb, and C) SynH117A. Values found from the fits (solid lines) in each panel are given in Table 2.
Figure 5. Spectral analysis of GdmCl denaturation. Optical spectra are shown without addition of GdmCl (solid lines) and at high [GdmCl] (dashed lines) for the A) ferric and B) ferrous forms of pentacoordinate horse heart myoglobin, the C) ferric and D) ferrous forms of hexacoordinate SynHb with the covalent bond, and E) ferric and F) ferrous forms of hexacoordinate SynH117A. Panels E) and F) also contain the high [GdmCl] spectra (dotted lines) from the hhMb graphs of panels A) and B) for comparison.
The results for the SynH117A mutant protein are quite surprising (Figure 4C). This protein is less stable in both oxidation states compared with wild type SynHb, but unlike pentacoordinate hhMb, is most unstable in the ferrous oxidation state. Furthermore, the final absorbance spectra at high [GdmCl] are nearly identical to those of hhMb, indicating that heme is dissociating from this protein which lacks the His$^{117}$-heme vinyl bond (Figure 5E and F). This suggests that the His$^{117}$-heme vinyl covalent bond is the dominant contributor to heme affinity compared to the axial His-heme coordinate bonds.

4.3.6. Heme dissociation

If heme affinity is the principal factor affecting protein stability, it should be possible to directly measure its dissociation. This is the case for ferric (but not ferrous) Mb, in which transfer to the heme-scavenger protein (apoH64Y/V68F Mb) has been used to measure rate constants for heme dissociation (Brantley et al. 1993; Hargrove et al. 1996b), demonstrating that this protein loses heme at a rate of ~ 0.007 h$^{-1}$ at pH 7.0. Previous work demonstrated very slow heme loss from ferric wild type SynHb without the covalent bond (Scott and Lecomte 2000; Lecomte et al. 2001). Figure 6 shows heme dissociation experiments for SynHb (containing the covalent bond) and the SynH117A mutant protein in both the ferrous and ferric oxidation states. In each case, absorbance spectra are shown just after mixing with apoH64Y/V68F Mb and 2000 minutes after mixing. For the wild type protein in both oxidation states, and for the ferric SynH117A mutant protein, no appreciable heme transfer is observed (Figure 6 A-C). However, the ferrous SynH117A mutant protein loses heme at a rate of 0.012 min$^{-1}$, approximately 100-times faster than ferric sperm whale myoglobin under
the same conditions (Figure 6D and E). Thus the ferrous form of the protein is susceptible to heme loss, and therefore loss of function, without the added protection of the covalent bond.

Figure 6. Heme loss from holoprotein to mutant horse heart apomyoglobin. Optical spectra initially (lines) and after 2000 s (dotted lines) for A) ferric SynHb, B) ferrous SynHb, C) ferric SynH117A, and D) ferrous SynH117A reveal heme loss only for E) ferrous SynH117A.
4.4. Discussion

Interest in hemoglobin stability originated with efforts to understand hemoglobin pathologies that resulted in heme release and protein degradation. Thus, the vast majority of research in hemoglobin stability has focused on myoglobin and red blood cell hemoglobin, which are pentacoordinate with only the proximal His coordinating the heme iron. The earliest studies linked heme release with the ferric oxidation state, which led to the view that the ferrous-His bond is much stronger (Banerjee 1962; Bunn and Jandl 1968). Later work established that contributions to stability comprise globin folding and heme affinity in a relatively complex, interdependent relationship between the strength of the coordinate bond, hydrophobic interactions between the globin and the porphyrin, and solvent exclusion from the heme pocket (Hargrove et al. 1994a; Hargrove et al. 1996a; Hargrove and Olson 1996; Hargrove et al. 1996b; Tang et al. 1998).

Stability and heme affinity in hexacoordinate hemoglobins is not yet understood at this level, though a logical extension from work with Mb would suggest that having two coordinate His-heme iron bonds would further stabilize bound heme and foster protein stability. It is therefore surprising that heme retention in hexacoordinate SynHb would require additional support from the His\textsuperscript{117}-heme vinyl bond. The present results indicate that this bond is necessitated by lowered heme affinity in the ferrous oxidation state; this is counterintuitive to the situation with pentacoordinate Hbs, where it is the ferric oxidation state that has a lower affinity for heme. Possible reasons for this fundamental difference between SynHb and
pentacoordinate Hbs, along with the consequences of this unusual covalent bond on ligand binding are discussed below.

4.4.1. The role of the His\textsuperscript{117}-vinyl bond in protein stability

The His\textsuperscript{117}-heme vinyl bond is the dominant force holding heme in SynHb in both the ferrous and ferric oxidation states. This conclusion is most evident from the fact that even when the wild type protein unfolds (presumably leading to the pentacoordinate spectral transition in Figure 5C and D), the heme is retained. But if the covalent bond is broken, heme dissociation accompanies unfolding, which occurs under much milder conditions (Figures 5E and F). The cause of the decrease in stability in the absence of the His\textsuperscript{117}-heme vinyl bond could have both enthalpic and entropic origins. The loss of hydrophobic clustering in the heme pocket upon heme dissociation could lower the enthalpy of the folded state, and heme dissociation upon unfolding could increase the entropy of the unfolded state. However, the experiments presented here do not distinguish between these contributions.

Vu et al. (Vu et al. 2004b) report that decoordination of both axial histidines accompanies protein unfolding in both ferric wt SynHb forms (with and without the covalent bond) due to the presence of isosbestic points in the spectral transitions. However, their ferric, unfolded spectra are similar to those reported here, including the difference seen with and without the covalent bond. The similarity of the spectrum lacking the covalent bond to that of unfolded hhMb, in which it has been demonstrated that heme is dissociated, supports heme dissociation in SynHb lacking the covalent bond. The spectra of unfolded SynHb containing the covalent bond however, does not look like that of hhMb, suggesting that heme is retained.
Furthermore, the presence of isosbestic points does not imply any particular final coordination state, but only suggests a two-state transition. However, differences in unfolded states could also result from the different methods of denaturation used in these two reports.

The most surprising discovery from the experiments illustrated in Figures 4, 5, and 6 is that, in the absence of the His\textsuperscript{117}-heme vinyl bond, the ferrous hexacoordinate complex is significantly less stable than the ferric. This suggests the possibility that the influence of oxidation state on heme affinity in hexacoordinate Hbs is opposite to that in pentacoordinate Hbs like Mb, at least in the case of \textit{Syn}Hb. Coordination studies of model heme compounds are consistent with the behavior of \textit{Syn}Hb. Coordination of the heme iron is stronger in the ferric form than the ferrous due to the associated change in spin state of the iron upon binding of the second ligand (Nesset 1996; Safo and Debrunner 1997).

These results raise questions about the chemical nature of Hb stability in pentacoordinate versus hexacoordinate hemoglobins. Belief that the ferrous heme-His bond is stronger in Mb originated with the observation that heme transfer does not occur in this oxidation state (Hargrove et al. 1994a; Hargrove et al. 1996a; Hargrove and Olson 1996; Hargrove et al. 1996b; Hargrove et al. 1997). However, there is no obvious explanation for why the single ferrous His-heme bond is stronger than the ferric bond in pentacoordinate Hbs, or why the opposite might be the case in hxHbs. It will be important to investigate whether this phenomenon is specific to \textit{Syn}Hb, or if it is also seen in larger hexacoordinate Hbs lacking a covalent bond between the porphyrin ring and globin. It is possible that the truncated nature of \textit{Syn}Hb requires the His\textsuperscript{117}-heme vinyl bond in lieu of sufficient hydrophobic
"waterproofing" of the proximal heme pocket (Liong et al. 2001). This hypothesis is consistent with the increase in solvation of the H117A proximal heme pocket seen in the crystal structures of the mutant protein in both oxidation states (Figure 1E and F).

4.4.2. Effects of the His^{117}-vinyl bond on hexacoordination and exogenous ligand binding

While the details of the effects of the His^{117}-heme vinyl bond on protein stability were surprising, a role related to heme affinity was not unanticipated. A more curious result is the effect of the covalent bond on ligand binding. We have shown here that removal of this bond is felt more acutely on the ferric heme than the ferrous due to a higher affinity constant for hexacoordination in the ferric oxidation state. The ferrous hexacoordination equilibrium constant (K_{H2}) is only ~3-fold larger in the H117A mutant protein, mainly due to an increase in the rate constant for His coordination (k_{H2}) (Table 1). A much larger effect is observed on the redox potential, which drops by 85 mV in the H117A protein compared to wild type SynHb. A shift of this magnitude suggests the possibility of much tighter hexacoordination in the ferric oxidation state (K_{H3}); specifically, a 27-fold higher ratio of K_{H3}/K_{H2} compared to wild type SynHb. In fact, this 27-fold increase in K_{H3}/K_{H2} combined with the 3-fold increase in K_{H2} for SynH117A predicts that K_{H3} for the mutant should be ~80-times greater than that of SynHb according to Equation 3.

The effect on K_{H3} can also be evaluated by measuring rate and affinity constants for exogenous ligand binding in the ferric oxidation state. The influence of K_{H3} on these reactions is described by Equations 4 and 5, respectively. In each case, the observed binding
constant is reduced by a factor $\sim K_{H3}$. Therefore, $K_{H3+H117A}/K_{H3+SynHb}$ should equal the ratio of $k_{CN,obs,SynHb}/k_{CN,obs,H117A}$, which should also be equal to $K_{CN,obs,SynHb}/K_{CN,obs,H117A}$. In fact, these values (14 and 10, respectively) are quite similar, and their correspondence supports assignment of this value to the increase in $K_{H3}$ associated with breaking the His$^{117}$-heme vinyl bond in SynHb.

However, these values are much lower than the value of 80 predicted by Equation 3, suggesting that the shift in redox potential for SynH117A compared to the wild type protein derives from factors other than just effects of the covalent bond on axial histidine heme coordination. This conclusion is also supported by the fact that the redox potential for wild type SynHb in which the protein is produced without the His$^{117}$-heme vinyl bond is -150 mV (Lecomte et al. 2001), 45 mV higher than when this bond is present (Table 1), suggesting that the H117A mutation has additional effects beyond simply preventing this covalent bond. These effects might include movement of the H-helix (Figures 1B and D) and the resulting increase in solvent accessibility to the proximal heme pocket (Figures 1E and F).

In conclusion, the His$^{117}$-heme vinyl bond in SynHb is not required to achieve the global structural architecture of the wild type protein in either the ferric or ferrous oxidation states, although minor structural perturbations do affect hexacoordination of the heme. Removal of the covalent linkage also affects exogenous ligand binding, particularly in the ferric oxidation state. The different kinetic and equilibrium constants exhibited by SynHb with and without this bond can both be explained by changes in the affinity constant for intramolecular distal His coordination for the ferric heme ($K_{H3}$). Furthermore, the ferrous protein lacking this
bond is much less stable against denaturation than the ferric, and both oxidation states differ from wild type SynHb in that heme is lost during unfolding of the H117A mutant protein. In contrast, when this bond is present in SynHb, even very high concentrations of GdmCl (7M) do not dissociate the heme from the globin in either oxidation state. Therefore, this covalent bond is likely present to prevent heme dissociation from the ferrous hexacoordinate complex of SynHb.

4.5. Materials and Methods

4.5.1. Protein expression and purification

The expression and purification methods used for SynHb (containing the covalent bond) and SynH117A have been described in detail (Hvitved et al. 2001; Trent et al. 2001a; Smagghe et al. 2006). Briefly, both proteins were cloned into a pET-29a (Novagen, Germany) vector and over-expressed in Escherichia coli BL21-CP (Invitrogen, Carlsbad, CA) for wt SynHb, and C41 (Avidis, France) for SynH117A. Following cell lysis, purification included ammonium sulphate fractionation, phenyl Sepharose hydrophobic exchange, DEAE Sepharose anion-exchange, and G-75 Sephadex size exclusion chromatography. The purified protein was oxidized with an excess of potassium ferricyanide followed by desalting on a G-25 Sephadex column in 0.01 M potassium phosphate, pH 7.0. The purity of the resulting ferric hemoglobin was found to have a Soret/A280 absorbance ratio greater than 4.7. The wild type SynHb protein purified and used in all experiments below contains the covalent bond (Hoy et al. 2004; Trent et al. 2004).
4.5.2. Structural determination

Crystals of ferric SynHb and SynH117A were grown as described previously (Hoy et al. 2004). Ferrous crystals were produced by reducing ferric crystals for 5 minutes in 10 µl of well buffer containing 100 µM sodium dithionite. Before mounting, the crystals were extensively washed in well buffer lacking sodium dithionite to remove excess reductant. To confirm that the crystal remained in the ferrous oxidation state during diffraction, the crystal was removed from the goniometer after data collection and immediately dissolved in 100 µl of 100 mM potassium phosphate buffer, pH 7.0, equilibrated with CO. The presence of CO-bound Hb in this sample was observed spectrally with an Ocean-Optics USB2000 UV-VIS spectrophotometer (Dunedin, FL), confirming that the crystal was still reduced, as ferric SynHb does not bind this ligand. As a control, this procedure (minus the initial treatment with dithionite prior to data collection) was repeated on ferric crystals and no CO-bound Hb absorption spectrum was observed. An additional check on oxidation state is the observation that ferrous crystals shatter in the presence of CO while ferric crystals do not. Finally, sulfur dioxide molecules (the reducing agent arising from sodium dithionite (Weiland et al. 2004)) are seen in complex with the protein in the electron density of the reduced crystals (see Supplementary Materials for comparative solvent electron density demonstrating the presence of sulfur dioxide).

Diffraction data for all three structures were collected on a Rigaku/MSC (Houston, TX) home source generator and solved by molecular replacement using the ferric SynHb structure as a model (1RTX.pdb). All three crystals were in space group P212121 with four
monomers in the asymmetric unit. A table of data collection and refinement statistics for the datasets of ferrous SynHb, ferric SynH117A, and ferrous SynH117A in included in the Supplementary Materials. Atomic coordinates have been deposited in the Protein Data Bank (www.rcsb.org, PDB ID 2HZ1, 2HZ2, and 2HZ3 for ferrous SynHb, ferric SynH117A, and ferrous SynH117A respectively). Figure 1 was created using the programs SPDBV (ExPASy, Switzerland), POV-Ray (Persistence of Vision Raytracer Pty. Ltd., Australia), and SURFNET (United Kingdom) using a 1.4 Å probe (Laskowski 1995; Guex and Peitsch 1997).

4.5.3. Carbon monoxide binding

Flash photolysis was used to measure bimolecular carbon monoxide association rate constants as described previously (Hargrove 2000). CO binding was measured by rapid mixing using a BioLogic SFM 400 stopped flow reactor coupled to a MOS 250 spectrophotometer (France). Experiments were conducted at 20ºC in 100 mM potassium phosphate, pH 7.0, using the method of Smagghe et al. (Smagghe et al. 2006). Time courses were collected at different CO concentrations, with at least three kinetic traces collected and averaged for each CO concentration. Data analysis and figure generation for ligand binding experiments were performed with the software Igor Pro (Wavemetrics, Inc., Portland, OR).

4.5.4. Spectroelectrochemistry

Potentiometric titrations were performed using the method of Altuve et al (Altuve et al. 2004) as described in detail previously (Halder 2006; Smagghe 2006). In brief, ferric Hbs (~30 µM) were titrated stepwise with a sodium dithionite solution (~40 mM) under anaerobic
conditions. All reactions used saturated calomel (reference) and platinum (working) electrodes, and were carried out at 25°C in argon-saturated 0.1M potassium phosphate, pH 7.0. Reduction was monitored by recording absorbance spectrum in the visible region (500-700 nm), and the corresponding cell potential was noted for each addition of sodium dithionite after equilibrium was reached. All midpoint potentials are reported with respect to a standard hydrogen electrode (SHE).

Midpoint potentials were extracted from the change in absorbance by fitting to the following equation (Bogumil et al. 1994):

\[
F_{\text{reduced}} = \frac{e^{-\left(\frac{nF(E_{\text{obs}} - E_{\text{mid}})}{RT}\right)}}{1 + e^{-\left(\frac{nF(E_{\text{obs}} - E_{\text{mid}})}{RT}\right)}}
\]

Equation 6.

\(F_{\text{reduced}}\) is the fraction of reduced protein, given as the normalized change in absorbance at 560 nm, \(E_{\text{obs}}\) is the observed cell potential, and \(E_{\text{mid}}\) is the midpoint potential obtained by fitting the experimental data to Equation 6.

4.5.5. Cyanide binding kinetics

Time courses for cyanide binding were measured using a Varian Cary 50 spectrophotometer (Palo Alto, CA). Cyanide concentrations ranging from 0 M to 0.4 M were used, and each cyanide concentration was prepared using 100 mM potassium phosphate, pH 7.0, in a separate cuvette before protein addition, and the reaction was monitored for up to twelve hours. The change in Soret absorbance was plotted versus time and fitted to a single exponential to obtain \(k_{\text{obs}}\) at each [CN]. The linear dependence of this value is the observed bimolecular (pseudo-first order) rate constant \((k_{\text{CN}}')\).
4.5.6. Equilibrium ferric ligand binding

Since the kinetic schemes for ligand binding in hxHbs are complex, it is important to measure affinities using equilibrium methods as well (Kundu 2003a). Equilibrium association constants for cyanide and azide were measured by equilibrium titration. For both experiments, the protein concentrations were ~7 µM (well below the respective K_D values). The samples were allowed to sit for 30 minutes following mixing to ensure equilibrium. Absorbance spectra were measured during ligand binding, and the position of the Soret peak was monitored to calculate fractional saturation. The peak shifted from ~ 408/410 nm (for the unliganded ferric protein) to ~ 415/418 nm (for the ligand-bound species) for SynHb/SynH117A, and the inflection of the first derivative of each spectrum was used to find the peak maximum. Equilibrium binding curves were fit to the following equation to extract equilibrium association constants:

\[
\lambda = \frac{[L]}{K_{CN,obs} + [L]} \lambda_L + \left(1 - \frac{[L]}{K_{CN,obs} + [L]} \right) \lambda_{3+}
\]

Equation 7.

where [L] is the ligand concentration, K_{CN,obs} is the equilibrium association constant, and \lambda_L and \lambda_{3+} are the Soret peak wavelengths for the ferric ligand bound and unliganded ferric forms of the protein, respectively.

4.5.7. Guanidinium chloride unfolding

Guanidinium Chloride (GdmCl) unfolding experiments were performed according to the method of Pace et al. (Pace 1990). A stock solution of 7.5 M GdmCl, 30 mM Mops, pH7.0 was used to prepare cuvettes varying in [GdmCl] from 0 to 5 M. 10 µM Hb was added to
each cuvette. For ferrous experiments, all solutions contained 100 µM sodium dithionite and were purged with nitrogen. Samples were equilibrated for 1-2 hours, then absorbance spectra were measured. Soret absorbance values were plotted as a function of GdmCl concentration (see Equation 8), and the midpoints of the unfolding curves and ∆G_O values (Table 2) were determined according to a two-state model of denaturation (\( F \leftrightarrow U \)) (Pace 1990; Hargrove et al. 1994a; Hargrove and Olson 1996; Mok et al. 1996).

\[
A_{obs} = A_F - \frac{A_F - A_U}{1 + e^{(\Delta G_O - m[GdmCl]) / RT}} \quad \text{Equation 8}
\]

where A is absorbance, F refers to folded protein, U refers to unfolded protein, ∆G_O is the Gibbs free energy, and m is the number of denaturant molecules bound in the reaction.

4.5.8. Heme Loss Assay

Rate constants for heme dissociation were measured using the H64Y/V68F apomyoglobin method described by Hargrove et al (Hargrove et al. 1994b). For experiments with ferrous Hbs, all cuvettes and solutions were purged with nitrogen and the reaction buffers contained 100 µM sodium dithionite. Spectral changes were monitored over the course of 2000 minutes. For ferric proteins, heme loss is accompanied by an increased absorbance at 600 nm. For ferrous proteins, heme loss shifts the Soret peak to 434 for MbH64Y/V68F. The change in absorbance at the Soret peak as a function of time was fit with a single exponential to extract the rate constant for heme loss.
4.6. Supplemental Materials

4.6.1. Solvent Electron Density

Sulfur dioxide molecules are found in the ferrous structures of SynHb and SynH117A. Supplemental Figure 1 shows an example of this density in comparison to that of sulfate, cadmium, and water. All of these molecules were found in the crystallization buffer (30% PEG MME 5000, 0.2 M ammonium sulfate, 0.01 M cadmium chloride, 0.1 M MES at pH 6.5, and ~100 µM sodium dithionite).

Figure S1. Comparative solvent electron density. Electron density from A) sulfur dioxide, B) water, C) sulfate, and D) cadmium. Taken from the structures of ferric and ferrous SynHb (1RTX and 2HZ1).
4.6.2. Crystallographic Data

Data collection and refinement statistics are provided in Table 3.

<table>
<thead>
<tr>
<th>Table 4-3. Crystallographic Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ferrous SynHb</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>R-merge (%)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit cell (Å)</td>
</tr>
<tr>
<td>R&lt;sub&gt;cryst&lt;/sub&gt; (%)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)</td>
</tr>
<tr>
<td>Avg. B factor (Å²)</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Heme</td>
</tr>
<tr>
<td>Solvent</td>
</tr>
</tbody>
</table>

4.6.3. Heme Pocket Stereochemistry

In the ferric form, mutation of H117 increases the bond lengths between the axial histidines and the heme iron atom. However, the histidine tilts are decreased, resulting in the iron atom being pulled further out of the heme plane to the proximal side. In the ferrous state, only the proximal bond length increases, but both tilt angles change, resulting in no shift in the location of the iron relative to the heme plane. (See Table 4.)
Table 4-4. Heme stereochemistry of the ferric and ferrous structures of SynHb and the H117A mutant.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Proximal His tilt angle&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Distal His tilt angle&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Proximal His(Nε2) - Fe bond length</th>
<th>Distal His(Nε2) - Fe bond length</th>
<th>Fe-to-heme plane distance&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>SynHb ferric</td>
<td>79.0</td>
<td>72.7</td>
<td>2.11</td>
<td>1.99</td>
<td>-0.01</td>
</tr>
<tr>
<td>SynHb ferrous</td>
<td>86.8</td>
<td>79.6</td>
<td>2.07</td>
<td>1.99</td>
<td>0.01</td>
</tr>
<tr>
<td>SynH117A ferric</td>
<td>86.6</td>
<td>76.2</td>
<td>2.17</td>
<td>2.06</td>
<td>-0.03</td>
</tr>
<tr>
<td>SynH117A ferrous</td>
<td>89.3</td>
<td>76.5</td>
<td>2.22</td>
<td>1.99</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<sup>a</sup>Planes are defined as the least squares plane, determined using the 24-atom porphyrin macrocycle of the heme and the 5-atom imidazole ring of the histidine. The angle given is the acute angle (in degrees) between the two planes.

<sup>b</sup>Negative distance indicates that the iron is on the proximal side of the heme. Distances are in Angstroms.

4.6.4. Flash Photolysis Data

Flash photolysis (Figure S2) was used to measure bimolecular carbon monoxide association rate constants as described in the text.
Figure S2. Flash photolysis data for SynH117A. A) Time courses for 6 concentrations of CO. B) CO concentration plotted versus the sum of $\gamma_1$ and $\gamma_2$ provides $k_{CO}$ from the slope (Hargrove 2000).

4.7. Abbreviations

Hb, hemoglobin;

hxHbs, hexacoordinate hemoglobins;

trHbs, truncated hemoglobins;

SynHb, *Synechocystis* hemoglobin;

His, histidine;

SynH117A, *Syn*Hb mutant H117A;

CO, carbon monoxide;

CN, cyanide anion;

N3, azide anion;

GdmCl, Guanidinium Chloride;
Mb, myoglobin;

rmsd, root mean square deviation;

obs, observed;

hhMb, horse heart myoglobin

4.8. References


Hargrove, M.S., Barrick, D., and Olson, J.S. 1996a. The association rate constant for heme binding to globin is independent of protein structure. *Biochemistry* **35**: 11293-11299.


CHAPTER 5. CRYSTAL STRUCTURES OF PHENYLALANINE B10 MUTANTS IN A NONSYMBIOTIC PLANT HEMOGLOBIN

This section details the structures of two mutants of riceHb1 as published in Biochemistry\textsuperscript{11}. The author was responsible for protein production, purification, and crystallization, as well as structure solving, refinement, and analysis.

5.1. Introduction

Nonsymbiotic hemoglobins (nsHbs) are a family of proteins that differ from "classical" hemoglobins like myoglobin both structurally and in biophysical behavior as a result of reversible intramolecular coordination of the heme iron. This characteristic defines a group of Hbs now called "hexacoordinate hemoglobins" (hxHbs). The physiological role of these proteins remains unknown, although their expression has been linked with the plant response to hypoxia, to germination, respiration, somatic embryogenesis, metabolism of nitric oxide, antioxidation, and to events associated with the plant disease resistance response [1-18]. The ubiquitous presence of nsHbs in plants, coupled with the physiological significance of the events triggering their expression, indicate the need for a more comprehensive understanding of these proteins.

The investigation of strategically placed heme pocket amino acids within nsHbs is one method of learning more about their ligand regulation. The amino acid at position B10 (based on structural homology to myoglobin) has been found to be an important distal residue in Hbs. In myoglobin and leghemoglobin, the residue at this position interacts with the distal histidine to orient it with respect to the bound ligand [19-22], while in several truncated hemoglobins like the one from *Synechocystis*, this residue moves into the distal pocket to directly coordinate with the bound ligand [23-28]. A phenylalanine at position B10 is conserved in all known nsHbs except for the one found in *Nostoc commune* in which it is a histidine [29]. In the crystal structure of rHb1, it is located very close to the C$\gamma$ atom of the hexacoordinated distal histidine [30]. Thus, Phe(B10) is likely to play a functional role by affecting hexacoordination and ligand binding. A series of Phe(B10) mutants of riceHb1 were made and studied extensively, revealing that a phenylalanine is the only residue in this position capable of achieving of a stable oxyferrous protein complex and high affinity ligand binding in the ferric state[31].

### 5.2. Methods

The mutant riceHb1 proteins were expressed and purified as described previously [32, 33]. The purified proteins were oxidized with a molar excess of potassium ferricyanide, desalted on a Sephadex G-25 column in 0.01 M potassium phosphate, pH 7.0, concentrated to 3 mM, and stored at -80 °C until use. Crystal growth was achieved by hanging-drop vapor diffusion. Drops were produced by mixing 2 µl of 3 mM protein with 2 µl of well buffer containing 1.9 M Ammonium Phosphate, 20% Sucrose, and 3% Dioxane at pH 7.0. Single crystals grew
overnight for F40L and within a week for F40W, both at room temperature. Representative crystals are shown in Figure 1.

![Figure 1. Nonsymbiotic hemoglobin crystals. Crystals of the rHb1 mutants F40L and F40W respectively.](image)

Diffraction data were collected at 100 K on a Rigaku/MSC home source generator, and processed using d*TREK [34]. Molecular replacement was performed using CNS [35] and the structure of riceHb1 (1D8U.pdb) as the starting model. Refinement of F40L was performed using both CNS and CCP4 [36] with manual rebuilding in O [37]. The final model is a dimer containing a total of 319 amino acids (residues A1-7 and B1-10 were left out due to lack of density and to improve statistics) and 222 water molecules, with $R = 20.3\%$, $R_{\text{free}} = 24.7\%$, and 99.8% completeness from 20-2.3 Å. The crystals of F40W were twinned, with twinning fraction 0.47. The twinning scripts in CNS were used to refine the structure. The final model is a dimer containing a total of 330 amino acids and 97 waters, with $R = 17.5\%$,
$R_{\text{free}} = 20.4\%$, and $97.5\%$ completeness from 20-2.4 Å. The final models have been deposited in the Protein Data Bank (2GNV.pdb and 2GNW.pdb for F40L and F40W respectively) and statistics are provided in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>riceHb1 F40L</th>
<th>riceHb1 F40W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P3$_1$21</td>
<td>P3$_1$21</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td>a=b=125.3 c=56.3 $\alpha=\beta=90^\circ$ $\gamma=120^\circ$</td>
<td>a=b=125.7 c=56.7 $\alpha=\beta=90^\circ$ $\gamma=120^\circ$</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0-2.3</td>
<td>20.0-2.4</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>32270</td>
<td>22580</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.3</td>
<td>5</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.8%</td>
<td>97.5%</td>
</tr>
<tr>
<td>Average $I/\sigma_I$</td>
<td>8.5</td>
<td>15.9</td>
</tr>
<tr>
<td>$R_{\text{merge}}$</td>
<td>0.065</td>
<td>0.069</td>
</tr>
<tr>
<td>Refined residues/waters</td>
<td>330 / 222</td>
<td>330 / 97</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$ (%)</td>
<td>20.3</td>
<td>17.4</td>
</tr>
<tr>
<td>$R_{\text{free}}$ (%)</td>
<td>24.7</td>
<td>20.4</td>
</tr>
<tr>
<td>Average B value</td>
<td>42.5</td>
<td>49.8</td>
</tr>
<tr>
<td>Ramachandran statistics (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored</td>
<td>91.3</td>
<td>90</td>
</tr>
<tr>
<td>Additionally allowed</td>
<td>7.6</td>
<td>11.4</td>
</tr>
<tr>
<td>Generously allowed</td>
<td>1.1</td>
<td>1.6</td>
</tr>
<tr>
<td>RMS Deviations from ideal geometry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.015</td>
<td>0.045</td>
</tr>
<tr>
<td>Bond angles (Å)</td>
<td>1.4</td>
<td>2.77</td>
</tr>
</tbody>
</table>
5.3. Results

There are no major structural perturbations resulting from these substitutions compared to the structure of wild type riceHb1. However, subtle changes in the heme pocket are observed. Figure 2A demonstrates that Phe(B10) in wild type riceHb1 packs against distal His(E7) and, along with Val(E11) and Phe(CD1), likely blocks major movement or rotation of the distal histidine. The F40L substitution creates a cavity above the His(E7) side chain and allows rotation in that direction. However, the large Trp(B10) side chain of the F40W mutant protein fills the cavity area and forces His(E7) to rotate in the opposite direction (Figure 2B). These rotations are accompanied by a 0.1 Å shift in the heme iron from the distal side of the heme in the wildtype protein to the proximal side in each mutant.

Both of these structures are homodimers, as is the wildtype (Figure 2C). The two molecules are similar for the wild type and F40L proteins, but the F40W structure shows two very different conformations for the Trp(B10) side chain. One molecule of the dimer crowds the distal pocket, as shown in Figure 2B. In the other, the Trp(B10) side chain moves out of the heme pocket into the solvent and creates a cavity similar to that in the structure of the F40L protein. In this conformation, the His(E7) side chain relaxes and rotates toward the cavity as in F40L (Figure 2C) [31].
Figure 2. X-ray crystal structures of ferric, hexacoordinate F40L and F40W riceHb1 compared to the wild type protein. A) A space filling model of wild type riceHb1 (left) shows that Phe(B10), Phe(CD1), and Val(E11) pack tightly against the coordinating His(E7) side chain. In the F40L protein (right), the mutation creates a cavity near the His(E7) side chain. B) The structures of F40L (blue) and F40W (chain A, purple) are overlaid with wild type riceHb1 (red). His(E7) rotates in each mutant protein compared to its position in wild type riceHb1. C) An overlay of the two molecules in the asymmetric unit of the F40W mutant protein shows two conformations for the Phe(40) side chain.
5.4. Conclusions

Mutational studies of residue Phe(B10) in riceHb1 have revealed that phenylalanine decreases hexacoordination compared to other amino acids placed at that position [31]. There appears to be a dynamic relationship between Phe(B10) and His(E7) that allows hexacoordination, but keeps the distal histidine from binding tightly enough to inhibit exogenous ligand binding. The rate of autooxidation of oxy-ferrous riceHb1 increases upon mutation of Phe(B10) [31] indicating that this residue’s close proximity provides stabilization to the distal histidine, allowing it to effectively coordinate the bound ligand. In the structure of the F40L mutant, the loss of the large phenylalanine conveys greater freedom of movement to the histidine, preventing such fine-tuning of its orientation and binding. Similarly, in the F40W structure, the mutated residue either leaves the pocket to convey the same freedom of movement, or crowds the pocket and hinders histidine movement. Selection pressure has favored phenylalanine in this position because it can achieve the necessary balance of confining the histidine enough to assist ligand binding, but not so much as to prevent hexacoordination.

5.5. References


14. Ross, E.J., Shearman, L., Mathiesen, M., Zhou, Y.J., Arredondo-Peter, R., Sarath, G.,
    and Klucas, R.V. (2001). Nonsymbiotic hemoglobins in rice are synthesized during
    germination and in differentiating cell types. Protoplasma 218, 125-133.

15. Ross, E.J., Stone, J.M., Elowsky, C.G., Arredondo-Peter, R., Klucas, R.V., and
    promoter by the cytokinin-regulated transcription factor, ARR1. J Exp Bot 55, 1721-
    1731.

    hemoglobins in rice (Oryza sativa var. Jackson) plants growing in normal and stress


    genes in plants. Plant Mol. Bio. 47, 677-692.

19. Carver, T.E., Brantley, R.E., Jr., Singleton, E.W., Arduini, R.M., Quillin, M.L.,
    with an unusually high O2 affinity and low autooxidation rate. J Biol Chem 267,
    14443-14450.

    (1993). The mechanism of autooxidation of myoglobin. J. Biol. Chem. 268, 6995-
    7010.


    (2004). Tyrosine B10 inhibits stabilization of bound carbon monoxide and oxygen in

    hemoglobin. Heme pocket structure and reactions with ligands. J Biol Chem 274,
    6898-6910.

    heme of Chlamydomonas chloroplast hemoglobin: evidence for ligation of tyrosine-
    63 (B10) to the heme. Biochemistry 38, 15360-15368.


CHAPTER 6. THE CRYSTAL STRUCTURE OF A LIGANDED NON-SYMBIOTIC PLANT HEMOGLOBIN REVEALS STRUCTURAL CHANGES IN THE EVOLUTION OF OXYGEN TRANSPORT

A paper for publication in the journal Structure.

Julie A. Hoy, Howard Robinson, James T. Trent III, Smita Kakar, Benoit J. Smagghe, Mark S. Hargrove

6.1. Summary

The crystal structure of the cyanide bound hemoglobin from ferric barley has been solved to 1.8 Å. This structure presents the first opportunity to examine the conformational changes that take place a nonsymbiotic plant hexacoordinate hemoglobin upon ligand binding. Comparison to the structure of riceHb1 reveals a large “piston” motion of the E-helix,

12 This work was made possible by the National Institutes of Health Award R01-GM065948, and support from the Iowa State University Plant Sciences Institute. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886.

13 Manuscript in preparation.

14 Primary Author and Researcher.

15 Department of Biochemistry, Biophysics, and Molecular Biology; Iowa State University; Ames, Iowa, 50011; USA

16 Biology Department; Brookhaven National Laboratory; Upton, New York, 11973; USA

17 To whom correspondence should be addressed
accompanied by coiling of the CD loop and uncoiling of the EF loop and resulting in
decoordination of the distal histidine side chain to allow exogenous ligand binding. Analysis
of this structure compared to other hexacoordinate and plant hemoglobins provides insight
into the mechanism of reversible hexacoordination in nonsymbiotic hemoglobins (nsHbs)
and the structural origins of the evolutionary specialization of pentacoordinate plant
symbiotic hemoglobins for oxygen transport.

6.2. Introduction

The evolutionary history of hemoglobins (Hbs) predates the divergence of prokaryotes and
eukaryotes, and descendants are present in nearly every living organism (Hardison, 1996;
Hardison, 1998; Hardison, 1999). Early Hbs were not oxygen transport proteins; rather, their
likely function was protection of cells from oxygen toxicity (Falkowski, 2006; Raymond and
Segre, 2006). Many organisms still contain non-oxygen transport Hbs as exemplified by the
"nonsymbiotic" Hbs (nsHbs) in plants (Appleby, 1988), neuroglobin (ngb) (Dewilde et al.,
2001; Trent et al., 2001a) and cytoglobin (cgb) (Burmester et al., 2002; Trent and Hargrove,
2002) in mammals, Ascaris Hb in invertebrates (Wittenberg, 1966), the Escherichia coli
oxygen sensor in bacteria (Delgado-Nixon et al., 2000), and the protoglobins in archaea
(Freitas et al., 2004).

Distinct structural differences are apparent when comparing oxygen transport and non-
oxygen transport Hbs in plants and animals (Kundu, 2003b). The oxygen transport Hbs
contain a heme linked to the globin via a single histidine side chain, leaving the heme group
"pentacoordinate" with an open binding site for oxygen. A second histidine is held in place near this site to stabilize the bound ligand, but does not coordinate the heme iron (Olson et al., 1988; Olson and Phillips, 1996). In the non-oxygen transport Hbs, this second histidine reversibly coordinates the ligand binding site (Duff et al., 1997; Hargrove, 2000b; Dewilde et al., 2001; Trent and Hargrove, 2002). These Hbs are thus referred to as "hexacoordinate" Hbs (hxHbs), and their structures and chemistry have become the subject of increasing attention due to potential roles in sensing and detoxifying NO and other environmental challenges (Gardner, 2005; Hankeln et al., 2005; Smagghe et al., 2006b). It is believed that the oxygen transport Hbs evolved from hexacoordinate Hbs in both plants and animals (Appleby, 1988; Taylor et al., 1994; Burmester et al., 2002).

Hexacoordination of the heme active site is energetically favorable due to the change in spin state of the iron d-shell electrons upon coordination by the second histidine side chain (Nesset, 1996; Safo and Debrunner, 1997). Localizing a potential ligand such as the distal histidine near a pentacoordinate heme iron without allowing it to bind presents a formidable thermodynamic challenge. Thus oxygen transport Hbs must provide a protein scaffold that can offset this coordination energy and stabilize an “open,” pentacoordinate, heme center.

Leghemoglobins are plant oxygen transport proteins found in symbiotic root nodules where they facilitate oxygen transport to nitrogen fixing bacteria (Appleby, 1984). Leghemoglobins, also called "symbiotic Hbs", evolved from nsHbs fairly recently (~ 200 million years ago) compared to oxygen transport Hbs in animals (~ 500 million years ago) (Trevaskis et al., 1997; Hunt et al., 2001; Vinogradov et al., 2006). The resulting sequence identity is higher
between leghemoglobins and nsHbs (~ 40 %) than between red blood cell Hb and non-oxygen transport Hbs in animals (< 20 %). The high degree of sequence identity in plants presents an opportune system for dissecting the protein structural features that changed during evolution of the capacity for oxygen transport.

Structures of several leghemoglobins have been reported, as has the structure of a nsHb in the hexacoordinate state (Harutyunyan et al., 1995; Hargrove, 1997; Hargrove, 2000b). Here we report the structure of barley nsHb bound to cyanide (barHb:CN), which provides the first opportunity to explore the structural mechanisms of ligand binding in nsHbs. The difference between this structure and that of nonsymbiotic hexacoordinate riceHb1 centers on a “piston” movement of the E-helix dependent on rearrangements in the "CD" and "EF" regions of the globin, resulting in a large conformational change that dislocates the histidine side chain to allow exogenous ligand binding. A comparison of this structure to leghemoglobins provides insight into the evolution of oxygen transport Hbs in plants, demonstrating such proteins were made possible by locking a nsHb into the pentacoordinate conformation.

6.3. Results and Discussion

6.3.1. The structure of barHb:CN and comparison to riceHb1

The structure of ferric barley hemoglobin bound to cyanide was solved to 1.8 Å. The backbone of one molecule and sample electron density for the barHb:CN crystal structure are shown in Figure 1A and B, with data collection and refinement statistics listed in Table 1.
The electron density is well-defined for all amino acids except for the first 7 at the N-terminus and the last at the C-terminus, and these are not included in the final model. The asymmetric unit contains 8 molecules, arranged as four pairs of dimers. The 8 chains are very similar, with a main-chain rmsd of only 0.38 Å. The density clearly shows a cyanide ligand bound 1.97 Å from the heme iron and 2.62 Å from the N$_{\varepsilon2}$ atom of the distal histidine.

Ideally a comparison of the hexacoordinate and ligand-bound structures of nsHbs would involve the same polypeptide. However, neither a hexacoordinate structure of barHb nor a ligand-bound structure of riceHb1 have been solved to date. Efforts to soak ligands into crystals of hexacoordinate riceHb1 result in crystal destruction, as would be expected for even a modest conformational change (Trent et al., 2004). Fortunately, barHb and riceHb1 are homologues sharing 82% sequence identity. One potential concern about comparing these proteins is the difference in rate constants as reported for ligand binding to barHb (Duff et al., 1997) compared to those of riceHb1 (Arredondo-Peter et al., 1997). However, we have recollected the kinetic data for barHb using updated methods and find that they are in fact consistent with the values for other nsHbs (see Supplemental Data for additional information).

In addition to primary sequence similarity, the two proteins also share the same dimer contacts, despite the conformational changes occurring with ligand binding. The dimer contacts are as reported for riceHb1 (Hargrove, 2000b; Goodman and Hargrove, 2001), the
Figure 1. BarHb:CN structure and electron density. A) Stereoview of the barHb:CN structure, showing ribbon backbone, heme, proximal histidine below the heme, distal histidine above the heme, and cyanide ligand. Helix labels are included for clarity. B) Electron density for barHb:CN and C) riceHb1 shows the breakdown of density around the CD loop region in the later.
Table 6-1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data Collection Statistics</th>
<th>Native</th>
<th>SAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (nm)</td>
<td>1.1</td>
<td>1.72</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>141-1.8 (1.86-1.8)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50-2.01 (2.08-2.01)</td>
</tr>
<tr>
<td>R-merge (%)</td>
<td>0.045 (0.378)</td>
<td>0.056(0.273)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>87.4 (47.8)</td>
<td>87.8 (52.4)</td>
</tr>
<tr>
<td>Unique, Total Reflections</td>
<td>123785, 452780</td>
<td>89616, 619984</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.7 (2.6)</td>
<td>7.0 (3.8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement/Quality Statistics</th>
<th>barHb:CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Group</td>
<td>P1</td>
</tr>
<tr>
<td>Unit Cell (Å)</td>
<td>a = 45.74, b = 60.56, c = 145.80</td>
</tr>
<tr>
<td></td>
<td>α = 99.17, β = 96.97, γ = 92.43</td>
</tr>
<tr>
<td>Molecules in Asymmetric Unit</td>
<td>8</td>
</tr>
<tr>
<td>Refined Residues, Waters</td>
<td>1227, 876</td>
</tr>
<tr>
<td>(R_{\text{cryst}}) (%)</td>
<td>20.23 (30.9)</td>
</tr>
<tr>
<td>(R_{\text{free}}) (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>24.96 (38.3)</td>
</tr>
<tr>
<td>Average B factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>35.1</td>
</tr>
<tr>
<td>Ramachandran Plot</td>
<td></td>
</tr>
<tr>
<td>Most favored (%)</td>
<td>94.3</td>
</tr>
<tr>
<td>Additionally allowed (%)</td>
<td>4.8</td>
</tr>
<tr>
<td>Generously allowed (%)</td>
<td>0.9</td>
</tr>
<tr>
<td>r.m.s.d. from Ideal Geometry</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.018</td>
</tr>
<tr>
<td>Bond angles (deg)</td>
<td>1.564</td>
</tr>
</tbody>
</table>

<sup>a</sup> Outer shell statistics are shown in parentheses.

<sup>b</sup> Calculated using 5% of reflections.

The only exception being a single substitution of valine in riceHb1 to isoleucine in barHb. The molecules are held together by a cluster of complementary hydrophobic side chains (I43, V117, and F120) and hydrogen bonds (between P45, S46, and E116) in the BC loop and G-helix regions of both molecules. These results suggest that a comparison between the structures of barHb:CN and riceHb1 will provide insight into the switch between the pentacoordinate and hexacoordinate conformations.
6.3.2. Structural changes upon ligand binding in nsHbs

The structures of both the ligand-bound and ligand-free conformations have been reported for only three hexacoordinate hemoglobins. The first was the truncated hemoglobin from the cyanobacterium *Synechocystis* (SynHb). The conformation change upon ligand binding in this protein is the largest observed to date (Hoy et al., 2004; Trent et al., 2004) and involves a hinge movement of the E-helix around the EF loop that swings the distal histidine completely out of the distal pocket. On the contrary, the structures of mouse neuroglobin reveal little conformational change upon ligand binding (Vallone et al., 2004b; Vallone et al., 2004a). However, the ligand-bound structure of neuroglobin resulted from soaking a hexacoordinate crystal in a CO solution, which could limit conformational flexibility. The cyanide-bound structure of a hxHb from *Drosophila melanogaster* reveals both an internal shift of the heme and a rearrangement of the CD loop and N-terminal region of the E-helix (de Sanctis, 2005; de Sanctis, 2006). The conformational change reported here for plant nsHbs is different from all three previous structures.

Upon ligand binding, barHb undergoes a large conformational change as compared to riceHb1 involving the E-helix, facilitated by changes in the CD and EF loop regions. This can be seen in the main chain rmsd between barHb and riceHb1 overlaid with the primary sequence alignment in Figure 2A. The rmsd of the E-helix between the two structures is 3.98 Å compared to the best overall protein fit of 1.07 Å over 115 amino acids. This is due to a large shift along the helical axis in the N-terminal direction and a small upward shift away from the heme. This moves the distal histidine Ne$\varepsilon_2$ atom from 2.08 Å to 5.14 Å away from the heme iron, allowing room for the incoming ligand. This large movement is concomitant
Figure 2. Comparison of plant Hb structures. A) The primary sequence alignment is provided for barHb, riceHb1, and soybean Lba, color coded by conservation using the BLOSUM62 score. Above the alignment, helices are labeled according to their position in barHb, and the rmsd between barHb and riceHb1 is shown per residue. B) The backbone alignment of barHb:CN (blue), hexacoordinate riceHb1 (1D8U.pdb, red), and acetate bound Lba (1BIN.pdb, green) reveals the large shift in the barHb:CN E-helix upon ligand binding. Included are the heme, ligand, proximal, and distal histidines.
with an increase in structural definition in the CD loop, and a decrease in definition of the EF loop. A D-helix forms in the CD region, though a proline found in the barley Hb sequence (serine in riceHb1) just prior to the D-helix may affect the formation of this helix, and its presence in the hexacoordinate form cannot be ruled out. A kink forms in the C-terminal end of the E-helix and a turn of the F-helix is lost in order to increase the length and flexibility of the EF loop. The global changes in conformation between the barHb:CN and riceHb1 proteins are shown in Figure 2B.

In the CD loop region, a network of hydrogen bonds forms among the side chains, backbone, and solvent molecules (Figure 3A). In addition, a cluster of hydrophobic residues forms within the B- to E-helix sections of the protein. Both serve to hold the newly formed helical structures intact. riceHb1, on the other hand, has no such network of H-bonds and many of the hydrophobic side chains can be seen pointing away from the interior of the protein (Figure 3B). Hence, clear electron density is seen for the CD loop region in barHb:CN but not in riceHb1 (Figure 1B, C). In the EF loop region, a similar network of hydrogen bonds involving solvent is much more extensive in barHb:CN than in riceHb1, and likely functions to control the movements of the elongated loop, as seen in Figure 4A and B.

6.3.3. A mechanism to achieve reversible ligand binding

The previous section described the folding of the CD and EF loop regions responsible for stabilizing the "open" protein conformation to which exogenous ligands can bind. Another important structural mechanism required for reversible intramolecular coordination is a
Figure 3. Hydrogen bonds and hydrophobic networks in the CD-loop region. The hydrogen bonds (pink) and hydrophobic residues (yellow) in the CD-loop region are shown for A) barHb:CN (blue), B) riceHb1 (red), and C) Lba (green) in divergent stereoview. Helices are labeled along with some familiar amino acids for reference.
Figure 4. Hydrogen bonds in the EF-loop region. The hydrogen bonds (pink) in the EF-loop region are shown for A) barHb:CN (blue), B) riceHb1 (red), and C) Lba (green) in divergent stereoview. Helices are labeled for reference.

means of dislodging and properly orienting the indigenous histidine side chain. Previous studies have revealed the importance of the conserved Phe(B10) in nsHbs. The side chain in this position in Hbs from other organisms often plays a role in regulating ligand binding
(Carver et al., 1992; Kundu, 2003a; Trent et al., 2004). In nsHbs, Phe(B10) is required to maintain a stable oxyferrous complex and to prevent tight histidine coordination, particularly in the ferric oxidation state (Smagghe et al., 2006a). A comparison of barHb:CN and riceHb1 allows for a structural explanation of the role of Phe(B10) in regulating histidine coordination.

The interaction between His(E7) and Phe(B10) in riceHb1 is unique compared to other Hbs. The distal histidine adopts a planar orientation very unusual for axial histidines coordinating the heme iron. In most circumstances, the imidazole plane is positioned so that the N\(\delta_1\) atom points in the direction of the heme propionates (Zaric et al., 2001). In riceHb1 and other hxHbs, the imidazole plane is positioned so that the N\(\delta_1\) atom points directly away from the propionates, in order for the N\(\varepsilon_1\) atom to coordinate the heme iron. This orientation results in an interaction between the N\(\delta_1\) atom of the distal histidine with the aromatic ring of Phe(B10).

Extensive studies of the interactions between phenylalanine and histidine observed in high-resolution x-ray crystallographic structures have established the most common Phe-His orientations (Thomas et al., 2002; Meurisse et al., 2003). According to this analysis, the orientation of these two side chains is very unusual in riceHb1. In most cases, Phe-His pairing does not occur between helices, as in riceHb1, and N\(\delta_1\) is almost never the atom of closest approach. Furthermore, only 1% of Phe-His pairs adopt the \(\chi_1\) torsion angles seen in riceHb1 (approximately \(-90^\circ\) for Phe and trans for His). Figure 5 shows the configuration
adopted by these side chains in riceHb1 with respect to the heme. The low occurrence of such a Phe-His pairing means it is likely to be energetically unfavorable, thus preventing the distal histidine from binding the heme iron too tightly for de-coordination and subsequent ligand binding (Sippl, 1995).

Upon ligand binding, the distal histidine side chain rotates to a more favorable orientation with respect to the heme propionates, and the atom of closest approach to the Phe(B10) side chain becomes C\(\delta_2\), a more common orientation. The distance of closest approach also increases from 3.39 Å to 3.68 Å, and the B-factors decrease from 47 Å\(^2\) and 35 Å\(^2\) (for Phe and His respectively) in riceHb1 to 24 Å\(^2\) and 30 Å\(^2\) in barHb:CN. In addition, the rmsd between molecules in the asymmetric unit (8 for barHb:CN and 2 for riceHb1) for these amino acids decreases from 0.86 Å and 0.89 Å (for Phe and His respectively in riceHb1) to 0.15 Å and 0.16 Å (for Phe and His respectively in barHb:CN). It is evident that the arrangement of these two side chains is more stable in the ligand-bound conformation, likely contributing to the lower degree of hexacoordination seen in nsHbs compared to other hxHbs (Smagghe et al., 2006b).

The flexibility of the CD and EF loop regions, and the structural changes that take place in these regions, make possible the “piston” motion of the E-helix that accompanies dissociation of the distal histidine and subsequent ligand binding. This flexibility, along with the interaction of Phe(B10) and the distal histidine, promote reversible hexacoordination. These observations begin to address the question of why bis-histidyl coordination in other
Figure 5. Heme pocket comparison between barHb and riceHb1. Phe(B10) and the distal histidine His(E7) in riceHb1 (red) and barHb:CN (blue).

protein folds is irreversible. For example, *bis*-histidyl heme proteins such as cytochrome b$_5$ (cyt b$_5$) are not capable of reversible distal histidine binding and therefore cannot bind exogenous ligands. In the case of cyt b$_5$, the heme is held in place by four short helices interrupted by short loops (Durley, 1996). This rigid structure would not accommodate the large coordinated His side chain movement that is essential for reversible coordination in the nsHb mechanism. Furthermore, modifications that allow ligand binding to occur in cyt b$_5$ (Rodriguez and Rivera, 1998; Ihara et al., 2000; Hirota et al., 2001) result in increased heme loss since cyt b$_5$ relies heavily on axial coordination for heme retention, due to heme cavity polarity and solvent exposure (Durley, 1996; Falzone et al., 1996; Falzone et al., 2001; Wang et al., 2003). Therefore, it is evident that the structure of nsHb is specially suited to disrupt
His coordination, allowing the flexibility in the E-helix required for a conformation shift that stabilizes the resulting pentacoordinate protein.

6.3.4. **The evolution of oxygen transport in plants**

NsHbs are ancestral to leghemoglobins, the latter evolving only in select plants for a specialized role in oxygen transport (Trevaskis et al., 1997; Hunt et al., 2001; Guldner et al., 2004a; Guldner et al., 2004b). A comparison of available leghemoglobin structures (such as 1GDJ, 2GDM, and 1BIN) shows very little overall structural change upon ligand binding, unlike the nsHbs. The backbone alignment of Lba with barHb:CN (Figure 2B) demonstrates a high degree of structural similarity between these proteins, with an rmsd of 1.45 Å over 100 of the 144 Cα atoms in Lba.

Comparison of barHb:CN and Lba provides insight into the structural changes involved in the conversion between hexa- and pentacoordinate Hbs. A telling difference between these two structures is in the loops that border the E helix. The CD loop region of Lba is two amino acids shorter than in barHb, and its structure is held together with a series of hydrogen bonds internal to the protein and less dependent on solvent. This network in Lba is multihelical, as in barHb:CN (comprising the B-, D-, and E-helices), and serves as an anchor for this corner of the globin. In Lba, H-bonds among S45, K64, and the heme propionates also contribute to rigidity in this loop region.

Similarly, the EF loop region is four amino acids shorter in Lba than in barHb, including an N-terminal portion of the F-helix that uncoils in the process of ligand binding in barHb:CN.
The hydrogen bond network in this region of Lba (Figure 4C) is limited to the corner of the EF loop near the C-terminal end of the E-helix, but also forms significant bonds with portions of the A- and H-helices. The combination of these factors limit the flexibility of the EF loop, inhibiting the E-helix shift required for reversible hexacoordination.

In addition to the rigid loops surrounding the E-helix in Lba, other factors that inhibit hexacoordination are present in this protein compared to nsHbs. A number of amino acid substitutions that stabilize the position of the E helix are apparent in Lba compared to barHb (Figure 6). Five of these (Asp$^8$, Tyr$^{30}$, Ser$^{45}$, Arg$^{70}$, and Ser$^{72}$) result in five additional inter-helical hydrogen bonds between the E-helix and the A-, B-, D-, and H-helices (Figure 6A). Three others (Ph$^{18}$, Leu$^{65}$, and Val$^{69}$) increase the inter-helical hydrophobic cluster involving the E-helix and amino acids throughout the distal portion of the globin fold (Figure 6B). The majority of these mutations change amino acids that are conserved within the barHb and riceHb1 sequences (Figure 2A). Thus, Lba exhibits fixed structural features which stabilize the pentacoordinate state and which are lacking in nsHbs.

It appears that evolutionary selection has generated pentacoordinate oxygen transport Hbs in plants by "freezing" nsHb into the “open” pentacoordinate conformation and eliminating reversible hexacoordination. Based on the comparison of the structures of pentacoordinate Lba and hexacoordinate barHb:CN, the evolution of oxygen transport in plant Hbs required sequence deletions limiting the flexibility of the loops surrounding the E-helix. Meanwhile,
Figure 6. **Primary sequence mutations in Lba.** Amino acid substitutions in Lba increase both the A) hydrogen bonds (pink) and B) hydrophobic contacts in Lba compared to barHb:CN. Mutated amino acids are shown in orange, while other important residues involved in the interactions are shown in dark green. Both panels contain the heme in grey and the acetate ligand and the distal and proximal histidines in green. Helices are labeled for reference.

Selection occurred for amino acid side chains that provided structure to the CD corner and EF regions, and increased inter-helical hydrogen bonding and hydrophobic contacts involving the E-helix. The combination of these factors produced a plant Hb locked into a rigid
pentacooordinate fold with enough free energy to overcome potential coordination by the distal His.

In animals, it is believed that oxygen transport myoglobins and hemoglobins evolved from hexacoordinate hemoglobins like neuroglobin and cytoglobin (Burmester et al., 2002). Because of the lower sequence identity between hexacoordinate and pentacoordinate animal sequences, it is difficult to tell whether the same structural changes occurred in animal Hbs as in plant Hbs. However, primary sequence alignments of the hemoglobins found in several animals (frog, zebrafish, chicken, mouse, human) all reveal extra amino acids in the CD and EF loop regions of the hxHbs compared to oxygen transport Hbs (data not shown). Only two structures of animal hxHbs in both the hexacoordinate and ligand-bound states have been solved. Flexibility in the CD loop and E-helix region is observed in *Drosophila* Hb, but not in mouse ngb. In the latter, ligand binding is accommodated by a shift in the position of the heme rather than movement of the E-helix. Thus it is possible that different structural mechanisms for the production of oxygen transport Hbs have occurred independently. However high sequence similarity between leghemoglobins and nsHbs, resulting from the more recent conversion to oxygen transport in plants, has left us with a clear structural "fossil record" for this particular evolutionary event.
6.4. Experimental Procedures

6.4.1. Protein expression and purification

Barley hemoglobin was cloned into pET29a vector (Novagen) and expressed in host strain BL21-CP cells (Invitrogen) as described previously (Trent et al., 2001b), with the following exceptions: the cells were grown for 24 hours without hemin addition and with kanamycin as the sole selective antibiotic. Purification included ammonium sulfate fractionation and phenyl-sepharose, DEAE-cellulose, and CM50-sephadex column chromatography. The Soret/A$_{280nm}$ ratio was at least 2.5 for kinetics and 3.1 for crystallography. The protein was reduced with sodium dithionite, which was removed with a G25 sepharose column, then oxidized with ferricyanide and run over the G25 column again before use.

6.4.2. Crystallization and data collection

Crystals were grown using the vapor diffusion method at 4°C with a well buffer consisting of 20% PEG 3350, 200mM ammonium citrate, 4% polypropylene glycol, 100mM sodium cyanide, pH5.3. Drops consisting of 2µl ~3mM ferric protein in 10mM sodium hepes, pH 7 and 2µl buffer produced crystals in less than a week. Native and single wavelength anomalous dispersion datasets were both collected at the National Synchrotron Light Source at Brookhaven National Laboratory on beamline X29. The SAD data was collected using anomalous scattering of the heme iron.
6.4.3. Structure determination and refinement

Structure determination depended on the anomalous scattering from eight iron atoms (for 162 residues plus one heme each). The crystal structure was determined by the SAD method, and the data were integrated and reduced using DENZO and SCALEPACK (Otwinowski, 1997). Atomic positions for eight iron atoms were located by HKL2MAP (SHELXD) (Pape, 2004), and the phases were calculated with SOLVE (Terwilliger and Berendzen, 1999). A solvent-flattened map calculated with RESOLVE (Terwilliger, 2000) at 2.4 Å resolution was of high quality. Model building was performed using RESOLVE (Terwilliger, 2002) and COOT (Emsley and Cowtan, 2004), while positioning of the other 7 monomers in the asymmetric unit was done with Molrep (Vagin, 1997), a program from the CCP4 suite (CCP4, 1994). Manual rebuilding was done in O (Jones et al., 1991), and the atomic model was refined using REFMAC5 (Murshudov et al., 1997), also from the CCP4 suite. The resolution was extended to 1.8 Å using the native data set and further refinement was carried out using REFMAC5. Structural calculations and figure creation were performed with LSQMAN (Kleywegt, 1996) and SPDBV (Guex and Peitsch, 1997). The primary sequence alignment in Figure 2 was performed with DIALIGN (Morgenstern, 2004) and analyzed with Jalview (Clamp et al., 2004).

6.4.4. Kinetic experiments

Flash photolysis was used to measure the bimolecular CO and O₂ association rate constants as described previously (Hargrove, 2000a; Trent et al., 2001a; Smagghe et al., 2006b). Rapid mixing experiments were carried out as previously described (Trent et al., 2001b) using a
BioLogic SFM 400 stopped-flow reactor coupled to a MOS 250 spectrophotometer. Reactions were monitored at the CO-bound heme absorbance maximum for barley, 417 nm. See Supplemental Data for more information.

6.5. References


6.6. Supplemental Information

The reaction associated with ligand binding in hexacoordinate Hbs is shown in Scheme 1 below.

\[
\begin{align*}
Hb_H & \rightleftharpoons Hb_P \xrightarrow{k'_L[L]} Hb_L \\
& \text{Scheme 1.}
\end{align*}
\]

By assuming a steady state equilibrium between Hb\(_H\) and Hb\(_P\), the following equation can be derived for the observed rate of the reaction (Hargrove, 2000):

\[
k_{\text{obs,CO}} = \frac{k_{-H}k'_L[L]}{k_H + k_{-H} + k'_L[L]} \quad \text{Equation 1.}
\]

Duff et al. reported rate constants for k\(_{CO}'\), k\(_{O2}'\), k\(_{O2}\), and k\(_{H}\) (Duff et al., 1997). The values of k\(_{CO}'\), k\(_{O2}'\) were estimated from rapid mixing experiments varying the concentrations of either O\(_2\) or CO. These association rate constants cannot be measured accurately by rapid mixing if k\(_{H}\) limits the reaction (Hargrove, 2000; Smagghe et al., 2006). Therefore, we have recollected these data using flash photolysis to define the bimolecular rate constants for O\(_2\) and CO binding to the pentacoordinate complex of the protein, and then fitting the observed reaction of CO initiated by rapid mixing to extract not only k\(_{H}\) (which is very similar to the value reported by Duff), but k\(_{H}\) as well. These data are shown in Supplemental Figure 1, and are provided in Table 2 to demonstrate that in fact barHb is very similar to riceHb1 and the other class 1 nsHbs that have been characterized since the first report by Duff et al.
The concentrations of CO used were 100, 200, 400, 600, 800, and 1000 µM. Flash photolysis was used to measure the bimolecular $O_2$ rate constant in a similar manner, as previously described (Trent et al., 2001). The concentrations of $O_2$ used were 131, 262, 625, and 1250 µM. Rapid mixing experiments were carried out as previously described (Trent et al., 2001) using a BioLogic SFM 400 stopped-flow reactor coupled to a MOS 250 spectrophotometer. Reactions were monitored at the CO-bound heme absorbance maximum for barley, 417 nm.

**Table 6-2. Kinetic Data**

<table>
<thead>
<tr>
<th>Protein</th>
<th>$k'_\text{CO}$ µM(^{-1}) s(^{-1})</th>
<th>$k_{\text{CO}}$ s(^{-1})</th>
<th>$K_{\text{CO}}$ µM</th>
<th>$k'_\text{O}_2$ µM(^{-1}) s(^{-1})</th>
<th>$k_{\text{O}_2}$ s(^{-1})</th>
<th>$K_{\text{O}_2}$ µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>BarHb:CN</td>
<td>2</td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>riceHb1(^a)</td>
<td>7.2</td>
<td>0.001</td>
<td>7200</td>
<td>68</td>
<td>0.038</td>
<td>1800</td>
</tr>
<tr>
<td>Soybean Lba(^b)</td>
<td>16</td>
<td>0.0084</td>
<td>1900</td>
<td>130</td>
<td>5.6</td>
<td>23</td>
</tr>
</tbody>
</table>

\(^a\) (Arredondo-Peter et al., 1997), \(^b\) (Hargrove et al., 1997)
Supplemental Figure 1. Ferrous ligand binding kinetics in cyanide bound barley Hb. Fitted rate constants from flash photolysis are plotted versus A) [O₂] and B) [CO] and fit to a line to provide k′O₂ and k′CO. C) Fitted rate constants from rapid mixing at different CO concentrations are fitted to Equation 1 to extract k⁻¹Η and kΗ.


Hemoglobins are ancient proteins that predate the divergence of prokaryotes and eukaryotes[1, 2]. Although the most thoroughly studied hemoglobins are mammalian oxygen transporters, it has become clear that this function is a recent evolutionary advancement coinciding with the “oxygen revolution” and the development of multicellular organisms[3]. Hemoglobins capable of oxygen transport are highly specialized structurally to maintain pentacoordination of the heme and appropriate ligand rate and equilibrium constants, while preventing autooxidation due to bound oxygen. Their predecessors however, adopt a more energetically favorable hexacoordinate conformation that precludes oxygen transport.

Despite the fact that hxHbs are ubiquitous in biological organisms, little is known about their purpose and mechanisms of function. Endogenous reversible hexacoordination in these proteins raises questions about the structural arrangements that must occur within the protein for exogenous ligand binding to occur. The goal of this research has been to increase understanding of ligand binding in hexacoordinate hemoglobins by studying the structural factors involved in allowing and regulating ligand binding.
7.1. The hexacoordinate hemoglobin from *Synechocystis* cyanobacteria

Since *Synechocystis* is a model organism for photosynthesis and a cyanobacteria that took part in the modification of the atmosphere 3.5 to 2 billion years ago[4], *Syn*Hb provides possible insight into the evolution of hemoglobin structure and function. The crystal structures of *Syn*Hb and *Syn*H117A presented in this work provide a detailed picture of the dynamic features of the heme environment and the hydrogen bonding networks that control the reactivity of this protein. The unusually versatile fold of this trHb is carefully constrained to control heme chemistry and ligand binding properties.

*Syn*Hb is unique in many ways. It is both truncated and hexacoordinate, with the distal histidine at position E10 rather than the typical E7 position and a novel covalent bond between the heme and a non-axial histidine. Unlike other trHbs, it is hexacoordinate in both the ferric and ferrous oxidation states over a wide pH range. [5-10] The work put forth here demonstrates that this protein also displays the largest conformational change yet seen upon ligand binding in hxHbs and that the ferrous state is less stable than the ferric state, a phenomenon opposite of what occurs in pentacoordinate hemoglobins[11-13].

Kinetic studies of the protein predicted a highly accessible heme pocket with a reactive heme and a porous protein matrix that might trap ligands[7]. The structures confirm these predictions, revealing high solvent accessibility around the heme propionates and a ligand tunnel seen in the ligand bound form that might slow ligand dissociation. Also revealed are
heme stereochemical factors contributing to ligand affinity, such as long bond lengths between the heme iron and axial histidines and large tilt angles for both histidines. However, the flexibility that provides the framework for this strong ligand affinity also seems to weaken the ferrous form of the protein, requiring the assistance of the covalent link between the heme and third histidine from within the protein.

The function of SynHb is not yet known, although some functions can be ruled out. The oxygen dissociation rate constant of SynHb is too low for oxygen transport and Synechocystis does not produce a nitrogenase complex like the cyanobacterium Nostoc commune, so cannot be involved in nitrogen fixation as is the Hb from N. commune[5-7]. Due to the variance among trHb association and dissociation rates[7], the differences in iron and ligand coordination[9], and the many unique features of SynHb, it is possible that the function of this particular protein might also be particular to Synechocystis.

What is known is that for ligand binding to occur, extremely large conformational changes must occur. This makes it possible that SynHb might function as a signal molecule. However, the need for large helical and loop movements may also be an artifact of the shortened globin fold. The tight ligand binding involving several distal pocket residues may implicate the protein in a ligand scavenging or storage role. Further physiological studies are needed to find out what this protein does in this organism, while the structural studies here provide insight into how the protein performs its function.
Further studies should also investigate the stability of other trHbs and hxHbs to determine if the instability and heme loss seen in the ferrous oxidation state of SynHb is typical of either trHbs or hxHbs, or if this is a phenomenon unique to Synchocystis. Since other hxHbs and most trHbs do not have the extra covalent bond found in Synchocystis, it is possible that the extraordinary ductility of SynHb is the root of this instability. On the other hand, perhaps ferrous instability is common to hxHbs but larger hxHbs have other mechanisms to compensate for it.

7.2. Hexacoordinate nonsymbiotic hemoglobins

The hxHbs from barley and rice have served as model proteins for understanding the function of nsHbs by studying ligand affinity and biophysical capabilities. The exogenous ligand-bound structure of barHb presented here complements the endogenously hexacoordinate structure of riceHb1, and provides information on how function is dictated by structure in this class of Hbs. A comparison of these structures provides insight into how reversible hexacoordination and protein conformational changes regulate ligand binding, and answers the question of how small gaseous ligands can bind the heme iron when the binding site is blocked by the distal histidine.

While ligand binding in barHb requires a large conformational change, the change is not as large as that seen in SynHb. This suggests that the movements required for ligand binding in the latter protein are exaggerated by the shortened, more flexible fold in SynHb. As seen in SynHb, the conformational changes that occur in barHb upon ligand binding are carefully
controlled by hydrogen bond networks. BarHb also enlists a network of hydrophobic residues and solvent molecules to control the flexibility of the CD and EF loop regions, while unique interactions with the Phe(B10) residue keep the distal histidine in the right position for reversible hexacoordination and stabilization of the bound ligand. It appears that the evolutionary advancement from the trHb 2-on-2 fold to the nsHb 3-on-3 fold, has conveyed a more finely-tuned and subtle mechanism for ligand binding and globin fold regulation.

The ligand bound conformation in barHb is very similar in fold to soybean leghemoglobin (Lba), which exhibits very little structural change upon ligand binding[14, 15]. Since the leghemoglobins are thought to have evolved from nonsymbiotic hemoglobins[16, 17], this suggests that specific mutational changes have in effect “frozen” the leghemoglobins into the ligand bound conformation seen in barHb so that large conformational changes are no longer necessary.

In fact, because the evolution of oxygen transport hemoglobins occurred relatively recently in plants, it is possible that future studies of selected mutations in barHb or Lba might reveal the specific structural elements that distinguish the hxHbs from the pentacoordinate oxygen transporters. This would facilitate understanding of the elements that predispose a hemoglobin towards a particular function as well as provide a model for how environment dictates evolution. In addition, this knowledge could be used in the rational design of hemoglobins for desired functions.
Rational design of this nature will require an understanding of the function of nsHbs. Future research will have to answer this question and determine if hxHbs function similarly in all plants or if there are multiple functions. Currently, evidence points to a function in nitric oxide scavenging and detoxification for possibly both plant and animal hxHbs. Further structural comparisons between the ligand free and ligand bound states of additional hxHbs will aid in determining whether all hxHbs are capable of similar function or if several varied functions are more plausible.

7.3. References


APPENDIX: CRYSTAL STRUCTURE OF CORN 1 HEMOGLOBIN

As the crystallographer for the Hargrove lab, I have been responsible for performing crystallographic analyses for work published in articles for which I am primary author as well as in articles for which lab members are primary authors. It has also been my responsibility to assist lab members in the tasks of crystal growth and screening, data collection and processing, and structure solving, refinement, and analysis. One structure to result from this collaboration that has not been discussed in previous chapters will be discussed here. Corn is an important crop plant that contains two nsHbs (Hbm1 and Hbm2)[1]. The crystal structure of ferric, hexacoordinate Hbm1 was solved to 2.2 Å resolution. The manuscript of which this structure is a part is in preparation.

The crystals were grown by vapor diffusion in 20% PEG MME 1900, 0.1 M sodium acetate, 0.2 M ammonium sulfate, pH 5.6. Data was collected at 100K on a Rigaku/MSC home source generator and processed using d*TREK[2], while the structure was solved using MOLREP[3] and Amore[4] from the CCP4 suite[5] and riceHb1 (1D8U.pdb) as a molecular replacement model. The structure was then refined using REFMAC5[6] from CCP4 and manual rebuilding in O[7]. The final tetramer contains 584 amino acids and 584 waters with $R = 20.3\%$ and $R_{\text{free}} = 25.3\%$. The structure is shown in Figure 1.

The program LSQMAN[8] from the Uppsala Software Factory was used to find the central monomer of the tetramer (chain B) and the root mean square deviation among the four
monomers (rmsd = 0.571 Å). The structure is very similar to the ferric, hexacoordinate structure of riceHb1, with an rmsd of 0.70 Å and 0.64 Å between Hbm1 chain B and riceHb1 chains A and B respectively. The dimer contacts in Hbm1 are also very similar to those in riceHb1. Hydrogen bonds are formed between Ser$^{49}$ and Glu$^{119}$, and the hydrophobic residues Phe$^{123}$, Val$^{120}$, and Val$^{46}$ are in close proximity to each other. The tetramer contacts are much weaker and rely on two sulfate ions at each interface.

In corn and rice, there are a number of H-bonds between helices throughout the protein that help to maintain the stability of the globin. However, there are twice as many in corn than in rice. This may limit the flexibility of corn to a degree, contributing to the kinetic differences seen between the two proteins. Despite the global structural similarity, changes in heme stereochemistry may also contribute to these differences. The proximal histidine tilt angle is smaller in Hbm1 than in riceHb1, while both the proximal and distal histidine to iron bond lengths are longer. The result is that the iron is pulled out of the heme plane 0.09 Å in the distal direction in riceHb1 and –0.04 Å in the proximal direction in Hbm1. These observations are in agreement with the decreased hexacoordination seen in Hbm1 compared to riceHb1[9].
Figure 1. The structure of nsHb corn Hbm1. The corn Hbm1 structure is a tetramer of two dimers, shown in blue and green. The heme in each monomer is colored in red, while the dimer contact residues are shown in orange (hydrogen bonding) and yellow (hydrophobic interactions).


