An examination of hexacoordinate hemoglobins using the techniques of biochemistry, biophysics and molecular biology

James Thomas Trent III
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Biochemistry Commons, Biophysics Commons, and the Molecular Biology Commons

Recommended Citation
http://lib.dr.iastate.edu/rtd/1896

This Dissertation is brought to you for free and open access by 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.
An examination of hexacoordinate hemoglobins using the techniques of biochemistry, biophysics and molecular biology

by

James Thomas Trent III

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Biochemistry

Program of Study Committee:
Mark S. Hargrove, Major Professor
Amy H. Andreotti
Gloria M. Culver
Reuben J. Peters
Jacob W. Petrich

Iowa State University
Ames, Iowa
2006

Copyright © James Thomas Trent III, 2006. All rights reserved.
# TABLE OF CONTENTS

**ABSTRACT** iv

**CHAPTER 1: GENERAL INTRODUCTION** 1
  Thesis Organization 3
  References 5

**CHAPTER 2: A MODEL FOR LIGAND BINDING TO HEXACOORDINATE HEMOGLOBINS** 7
  Abstract 7
  Introduction 8
  Materials and Methods 9
  Results 12
  Discussion 20
  Acknowledgements 24
  References 25

**CHAPTER 3: HUMAN NEUROGLOBIN, A HEXACOORDINATE HEMOGLOBIN THAT REVERSIBLY BINDS OXYGEN** 36
  Abstract 36
  Introduction 37
  Materials and Methods 39
  Results 42
  Discussion 44
  Acknowledgements 48
  References 49

**CHAPTER 4: A UBIQUITOUSLY EXPRESSED HUMAN HEXACOORDINATE HEMOGLOBIN** 56
  Abstract 56
  Introduction 57
  Materials and Methods 58
  Results 62
  Discussion 67
  Acknowledgements 73
  References 74

**CHAPTER 5: CRYSTALLOGRAPHIC ANALYSIS OF SYNECHOCYSTIS CYANOglobin REVEALS THE STRUCTURAL CHANGES ACCOMPANYING LIGAND BINDING IN A HEXACOORDINATE HEMOGLOBIN** 89
  Abstract 89
  Introduction 90
  Results 91
Discussion 101
Materials and Methods 105
Acknowledgements 107
References 108

CHAPTER 6: GENERAL CONCLUSIONS 123

ACKNOWLEDGEMENTS 126
ABSTRACT

Hexacoordinate hemoglobins are found in a truly diverse array of organisms, ranging from cyanobacteria to humans. The prevalence of these proteins in nature coupled with very high sequence identity between species homologues suggests they have a role vital to life. The high sequence identity also implies that hxHb structural features, including the hexacoordination phenomenon, are critical aspects of the physiological function. A detailed understanding of the ligand binding behavior of these proteins will not only facilitate efforts towards discovering their physiological role(s), but will also help explain how hxHbs perform their function. This dissertation consists of four published papers which examine several hexacoordinate hemoglobins from a diverse array of species. The studies conducted in these papers range from gene discovery and mapping, tissue expression and protein biochemistry through kinetic and structural analysis.
CHAPTER 1: GENERAL INTRODUCTION

Hemoglobins (Hbs) are perhaps the most well known family of proteins, due in part to a prominent role in respiration and distinctive red coloration. The abundance and ready availability of Hbs made them a favorite of early biochemical inquiry, and so today the breadth and scope of the accumulated body of work involving these proteins is probably unequaled by any other protein system. Despite this wealth of information it is becoming ever more clear that there is much we still do not know about these proteins. Recent studies suggest that many fundamental questions regarding Hb functionality remain to be answered by modern researchers.\textsuperscript{1; 2; 3; 4; 5} This dissertation presents work that strives to address some of these questions arising from a particular branch of the hemoglobin family known as the “hexacoordinate” hemoglobins.

Although they demonstrate a diverse array of primary and quaternary structures, members of the Hb protein family can be readily identified by a conserved alpha-helical "globin" fold and covalently attached heme b cofactor. The ligand binding reactions that comprise the typical physiological role of Hbs occur at the heme iron atom, which normally exists in one of two redox states, ferrous (Fe\textsuperscript{2+}) or ferric (Fe\textsuperscript{3+}). The octahedral coordination geometry of the heme iron presents six potential binding sites; pyrrole nitrogens occupy four of these sites and anchor the metal atom within the porphyrin ring. Of the two remaining binding sites, one (the "proximal" site) coordinates an invariant His sidechain, and the other (the "distal" site) is where "functional" ligand binding reactions occur.\textsuperscript{6} The terms “pentacoordinate” and “hexacoordinate” are used to describe the occupation of the heme iron; if the distal site is empty, then five of six binding sites are occupied and the Hb is referred to as pentacoordinate (see Figure A).
Until recently, it was widely accepted that a ferrous, deoxygenated Hb should exist in a pentacoordinate form.\textsuperscript{6, 7} This idea was centered on the belief that unhindered access to the available iron-binding site was necessary for rapid ligand binding and stabilization. Discoveries of numerous Hbs exhibiting reversible intramolecular coordination of the heme iron have shown this concept to be false.\textsuperscript{3,4} In the reduced, deoxygenated form these proteins have a ligand-binding site filled by the internal coordination of an amino acid side chain from the globin (see Figure B).\textsuperscript{8, 9} Logically this should present an impediment to entering ligands; in fact studies of other bis-coordinate hemoproteins (such as cytochrome b\textsubscript{5}) find they do not bind diatomic gases. Yet hxHbs are not only capable of reversibly binding oxygen and other heme ligands, but they also appear to do so with high affinity.\textsuperscript{10, 11, 12} The molecular details of these reactions are not known however, nor are the physiological function(s) they enable.

The work presented in this dissertation has previously been published in peer-reviewed journals. Chapters 2 through 5 represent four papers whose underlying theme is an interest in discerning the mechanism of ligand binding in hxHbs. A chief means of seeking this understanding is through kinetic analysis of ligand binding in various representative proteins from the hxHb family. But additional insight is sought through expansion of the numbers of known hxHbs through gene discovery, as well as acquiring a molecular picture
with x-ray crystallographic studies. My contribution to, and a brief discussion regarding the content of, each of these chapters is found in the following paragraphs.

**Thesis Organization**

Chapter 2 consists of a kinetic analysis of several different hxHbs, chiefly from plants but including a cyanobacterial example. This work was published in *Biochemistry*¹², and established a foundation method for our approach to subsequent kinetic studies. It describes a previously unknown phase of ligand binding that occurs over much longer time scales than that typically associated with hemoglobins. A kinetic model is presented that more accurately accounts for this phase of binding than previously existed. I conducted all of the kinetic experimental work, as well as cloned and expressed many of the proteins used in the study. The derivation of the equations used to describe the ligand binding reactions were done by Mark Hargrove and the cyanobacterial protein was provided by Angela Hvitved.

Chapter 3 is a paper which takes our kinetic analysis of hxHbs into the animal kingdom. This work was published in the *Journal of Biological Chemistry*¹³, and identified human neuroglobin as the first known instance of a hxHb in vertebrates. All experimental work was performed by myself, with some of the kinetic analysis done by Mark Hargrove. Richard Watts contribution consisted of the equation representing impact of the hexacoordination equilibrium constant on overall ligand equilibrium affinity.

Chapter 4 represents the first paper to be wholly initiated and conceptualized by myself. Published in the *Journal of Biological Chemistry*¹⁴, it identifies a novel human gene coding for a hxHb that is expressed in a wide range of tissues. This gene is examined from both a nucleic acid perspective as well as through protein biochemistry. I did all work associated this paper.
Chapter 5 arises from a project conceptualized independently by both myself and Suman Kundu. Upon discovering our duplication of effort, we pooled resources with respect to protein generation and concentrated our efforts separately with respect to solving the unliganded (Suman) and liganded (Me) structures of the cyanobacterial hxHb from *Synechocystis*. This resulted in a paper published in the *Journal of Molecular Biology*\textsuperscript{15}, and allowed for the first time a comparative study between the ligand bound and unbound states of a hxHb. Most experimental work was performed by myself, with Suman Kundu sharing in the protein generation and Julie Hoy doing the final refinement of the structure.
References


CHAPTER 2: A MODEL FOR LIGAND BINDING TO HEXACOORDINATE HEMOGLOBINS

A paper published in *Biochemistry*

James T. Trent III, Angela Hvitved, and Mark S. Hargrove

**Abstract:** Hexacoordinate hemoglobins are heme proteins capable of reversible intramolecular coordination of the ligand binding site by an amino acid side chain from within the heme pocket. Examples of these proteins are found in many living organisms ranging from bacteria to humans. The plant nonsymbiotic hemoglobins (nsHbs) are a class of hexacoordinate hemoglobins present in all plants. The nsHb from rice (rHb1) has been used as a model system to develop methods for determining rate constants characterizing binding and dissociation of the His residue responsible for hexacoordination. Measurement of these reactions exploits laser flash photolysis to initiate the reaction from the unligated, pentacoordinate form of the heme protein.

A model for ligand binding is presented that incorporates the reaction following rapid mixing with the reaction starting from the pentacoordinate hemoglobin. This model is based on results indicating that ligand binding to hexacoordinate Hbs is not a simple combination of competing first order (hexacoordination) and second order (exogenous ligand binding) reactions. Ligand binding following rapid mixing is a multi-phasic reaction displaying time courses ranging from ms to minutes. The new model incorporates a "closed", slow reacting form of the protein which is not at rapid equilibrium with the reactive conformation. It is also demonstrated that formation of the closed protein conformation is not dependant on hexacoordination.

**Abbreviations:** Mb-Myoglobin (Sperm Whale), Hb-hemoglobin, nsHb-nonsymbiotic plant hemoglobin, rHb1-rice nonsymbiotic plant hemoglobin I, H73L-rHb1-rice leucine substituted for the distal coordinating histidine, synHb-synechocystis hexacoordinate hemoglobin, ckHb1-cichorium hexacoordinate hemoglobin I, msHb-moss hexacoordinate hemoglobin.

---

Introduction

Recent discoveries of new hemoglobins in archea, bacteria, protozoa, plants, and vertebrate animals demonstrates the wide diversity of these proteins and suggests that they are found in most, if not all organisms (1-6). Hemoglobins serve many physiological roles, from oxygen storage and transport to catalytic activation and signal transduction (7). New physiological functions will certainly be discovered over the next few years. Many of these newly discovered heme proteins show the unusual ability to form reversible, intramolecular hexacoordinate complexes. Examples of this group include the carbon monoxide dependant transcription activator CooA, the *E. coli* oxygen sensor Dos, nonsymbiotic plant hemoglobins (nsHbs), *clamydomonous* Hb, *Synechocystis* Hb, and human neuroglobin (4,6, 8-11). The hexacoordinating amino acid side chain exerts considerable influence on ligand binding to these Hbs. Instead of a simple bimolecular reaction as is observed with traditional pentacoordinate Hbs, ligand binding to hexacoordinate Hbs is a complex competition between the binding of the exogenous ligand and intramolecular coordination by a local amino acid side chain. The large number of newly discovered heme proteins exhibiting this behavior, coupled with wide variation in hexacoordination kinetics, indicates that this is a fundamentally new and ubiquitous method for regulation of ligand binding to heme proteins.

Methods have recently been developed for the kinetic characterization of hexacoordination and ligand binding in hexacoordinate hemoglobins (12). These methods were developed using the nsHb from rice (rHb1) in which hexacoordination in the ferric and deoxy ferrous forms results from the binding of a histidine side chain to the sixth coordination site of the heme iron (Figure 1A). When ligand binding is initiated by flash photolysis of the ligand bound protein (Figure 1C), the reaction starts with the protein in the
pentacoordinate state (Figure 1B). Under these conditions, rate constants for hexacoordination and bimolecular ligand binding can be determined by monitoring bi-exponential ligand rebinding (12). However, under physiological conditions it is unlikely that ligand binding starts from a completely pentacoordinate population of protein molecules. It is much more likely that the protein exists in an equilibrium mixture of the hexacoordinate and pentacoordinate forms. When a suitable ligand becomes available, the rate of binding would then be determined by the fraction of protein in the pentacoordinate form in conjunction with the rate constants associated with ligand binding and hexacoordination.

The work described here uses rapid mixing and kinetic analysis to investigate the ligand binding reactions of several hexacoordinate hemoglobins. The rate constants for ligand binding and hexacoordination determined by flash photolysis are evaluated against time courses for ligand binding obtained from rapid mixing experiments. These comparisons demonstrate that the simple hexacoordination reaction scheme shown in Figure 1 is not sufficient to fully describe ligand binding to these proteins. A general model describing ligand binding to hexacoordinate hemoglobins is developed that accounts for reactions following rapid mixing and flash photolysis.

Materials and Methods

Bacterial Strains and Culture Conditions- All of the proteins examined here were expressed recombinantly. *E. coli* BL21(λDE3) – CodonPlus-RP cells (*E. coli*-CP-RP) (Stratagene, La Jolla, CA) were used as the protein expression host. These cells were grown in a fermentation apparatus consisting of a 20L polypropylene carboy resting in a 40L open rectangular water bath that regulates the growth culture temperature. Aeration was accomplished through mild stirring with a 6 cm magnetic stir bar accompanied with O₂ that
was bubbled through autoclave durable aquarium air stones at 60 psi. Expression media consisted of Terrific Broth supplemented with 50 μg/ml kanamycin, 50 μg/ml chloramphenicol and 1.0 ml antifoam (Sigma).

**Cloning, Expression, and Purification of Recombinant Proteins**—The cDNAs coding for each protein were inserted into pET29a (Novagen) between the NdeI and EcoRI restriction sites in the multi-cloning region. The rHb1 H73L mutant protein was generated as described previously (13). The source of the cDNAs were: rHb1 and H73L rHb1 (13), synHb [Hvitved et al. unpublished], ckHb1 (14), and msHb (15).

Each plasmid was grown in 20 liters of the expression media at 37°C until the optical density at 600 nm was 1.5, at which time expression was induced by the addition of IPTG (Sigma) to a final concentration of 500 μM. In some cases 5 mg/l hemin chloride (solubilized in 0.1M NaOH) was added at induction to increase the fraction of soluble protein. However, the addition of hemin chloride was not required for the formation of soluble, heme-bound protein and was not added at later stages of protein purification. Protein expression was allowed to continue for five hours, and the culture was harvested through pelleting of the cells by centrifugation (~5000g for 10 minutes.) The harvested cells and lysis supernatants all displayed bright red color, as typically seen in soluble heme-bound hemoglobins.

All proteins were then purified using a three-stage procedure consisting of ammonium sulfate fractionation, and phenyl-sepharose and DEAE-cellulose column chromatography (13). The protein solutions were dialyzed into 20 mM Tris (pH 8.5) after the phenyl-sepharose separation and prior to the DEAE-cellulose purification step. This three stage procedure typically results in proteins with a Soret/A₂₈₀nm ratio > 3.0. Horse heart myoglobin was obtained commercially from Sigma and not subjected to further purification.
**Kinetic Experiments**—Rapid mixing experiments were carried out with a Durrum stopped flow apparatus interfaced with a personal computer using previously described methods \((16, 17)\). Absorbance spectra were collected during CO binding using an On Line Instruments (OLIS) Rapid Scanning Monochromator (RSM) stopped flow apparatus. Samples were prepared as follows. Perfektum gas tight syringes containing 100mM potassium phosphate (pH 7.0) were bubbled with either CO or N\(_2\) for 20 minutes. Dry sodium dithionite was then added to each syringe to a concentration of \(~200\ \mu\text{M}\). Solutions of CO prepared in this way are 1000 \mu{M}, with lower concentrations of CO obtained through dilution of the 1000 \mu{M} solution into secondary syringes containing buffer equilibrated in N\(_2\). The protein sample syringe consisted of N\(_2\) equilibrated buffer with \(~200\ \mu\text{M}\) sodium dithionite and \(~1\ \mu\text{M}\) protein. Reactions were monitored at the corresponding CO peak for each protein sample.

**Data Analysis**—Rate constants for CO binding to rHb1 were extracted from least squares fits to binding time courses between 0.003 s (the dead time of the stopped flow apparatus) and 0.2 s. Below [CO] of 50 \mu{M}, a two exponential decay expression was necessary to describe the time courses (Figure 4). The faster of these rate constants are shown in Figure 5A (filled circles), and the slower are shown in Figures 5A (open circles) and Figure 6A. Above 50 \mu{M} [CO], data between 0 and 200 ms fit well to a single exponential decay (Figure 4C).

The H73L rHb1 rate constants were determined using single exponential fits to the rapid-mixing time courses between 0.003 s and 0.2 s at [CO] less than 50 \mu{M}. At concentrations above 50 \mu{M} [CO], data were fit to a double exponential expression and the faster rate constants interpreted in Figure 7 along with the single exponential fits for the time at [CO] \(< 50 \mu{M}\). The presence of a second rate constant in the higher [CO] time courses
was attributed to the slower binding phenomenon demonstrated in Figures 2 and 3, which is the subject of ongoing investigation. The kinetic simulation in Figure 6B was generated using the program KINSIM(18). All graphic representations and data analysis were carried out using the program Igor Pro (Wavemetrics, Inc).

Results

Time courses for ligand binding following rapid mixing- The reaction scheme for hexacoordination shown in Figure 1 can be simplified to the following equation if the dissociation rate constant for the bound ligand is very small compared to all other rate constants, and the reaction is carried out under pseudo first order conditions with the ligand in excess.

$$Hb_H \leftrightarrow k_H^\prime \rightarrow Hb_L$$  \hspace{1cm} Equation 1.

Based on this equation and using an improved steady state approximation for HbP (19), the following equation can be derived for the observed rate constant of ligand binding following rapid mixing.

$$k_{obs,H} = \frac{k_{-H} k_L^\prime [L]}{k_{-H} + k_H^\prime + k_L^\prime [L]}$$  \hspace{1cm} Equation 2.

Equation 2 predicts a single exponential decay time course that is dependent upon the concentration of the ligand. However, if $k_L^\prime [L] >> k_{-H} + k_H^\prime$, the observed rate of reaction will be $k_{-H}$ regardless of the ligand concentration (20).

Rapid mixing time courses for carbon monoxide binding to several hexacoordinate hemoglobins are shown in Figure 2A along with horse heart myoglobin (Mb) for comparison. Unlike the hexacoordinate hemoglobins, Mb is strictly pentacoordinate in the reduced, unligated form and its time course fits well to a single exponential decay. However, none of
the hexacoordinate hemoglobin time courses fit well to single exponential terms. This disagrees with the behavior predicted by Equation 2, and indicates that ligand binding following rapid mixing is more complex than the reaction shown in Figure 1 and described by Equation 1. The various hexacoordinate hemoglobins exhibit differing degrees of non-ideality with the most extreme deviance from single exponential binding displayed by the nsHb from moss.

To rule out slow drifts in absorbance or slow equilibration of the final absorbance spectra as the cause of non-ideality in Figure 2A, complete spectra were measured during CO binding using a rapid scanning stopped flow apparatus. An example of these spectra is given in Figure 2B for CO binding to synHb. All reactions exhibited discrete isosbestic points, demonstrating that the only reaction occurring on these time scales is the conversion of deoxy Hb to CO bound Hb.

The time course for CO binding to rHb1 at low [CO] (10 µM) is examined in more detail in Figure 3. The fitted curves in this figure have been isolated to demonstrate graphically the different fractions of binding present in the time course. Thus the fitted curve in Figure 2A for rHb1 represents the sum of the fitted curves presented in Figure 3. Of these three distinct fractions of ligand binding, two exist in the 0 to 200 ms region, and the third heterogeneous fraction of rebinding occurs on a much slower time scale. The percent of ligand binding which occurs in each fraction for rHb1 is listed in Table 1. As the majority of binding to rHb1 occurs during the first 200 ms, our current effort to model ligand binding concerns only these two phases of binding. The slow binding fraction (that occurring after 200 ms) will be analyzed later in the context of a hexacoordinate hemoglobin that exhibits a large amount of rebinding during this time.
Figure 4 demonstrates that two exponential terms are necessary for fitting CO binding to rHb1 at low [CO] (< 50 µM). Figure 4A is a time course for CO binding to rHb1 at 20 µM [CO]. Single and double exponential decays were fit by least squares to this time course and are shown along with the raw data. Residuals from each fit are shown in Figure 4B. In this case, the fit to the double exponential decay provides a significantly better model than the single exponential decay. At [CO] > 50 µM, a single exponential decay successfully accounts for all binding amplitude on this time scale. Figure 4C demonstrates this with a time course for CO binding to rHb1 at 400 µM [CO]. Single and double exponential decays were fit to the time course and overlaid with the raw data. The residuals for each fitted curve are plotted above the time course. In this case no benefit is derived from inclusion of a second exponential when fitting the time course as the residuals from the two fits are functionally identical. These results indicate that the more rapid rate constant for binding in Figure 4A is lost in the mixing dead time (~0.003 s) at the higher [CO] giving rise to Figure 4C.

Comparison of $k_{obs,H}$ by flash photolysis and rapid mixing—Equation 2 predicts a single exponential binding event if the reaction in Equation 1 is at rapid equilibrium. The rHb1 rate constants for each of the terms in Equation 2 have been measured independently using flash photolysis and are 1911 s$^{-1}$, 517 s$^{-1}$, and 6.0 µM$^{-1}$s$^{-1}$ for $k_{H}$, $k_{H}$, and $k_{CO}$, respectively (12). Therefore, the [CO] dependence of $k_{obs,H}$ predicted from Equation 2 can be compared to that of the two rate constants extracted from the 0-200 ms time courses for binding following rapid mixing.

This comparison is shown in Figure 5A with the expected values of $k_{obs,H}$ predicted from Equation 2 shown as a line, the faster of the two rapid mixing rate constants shown as filled circles, and the slower as open circles. The faster of the two rate constants (filled
circles) is lost in the mixing dead time at [CO] > 50 µM. Another comparison of the CO binding predicted from Equation 2 and that observed following rapid mixing is shown in Figure 5B. A simulated time course for CO binding (400 µM) to rHb1 following rapid mixing was generated using the mechanism in Equation 1 and the rate constant calculated from Equation 2. Figure 5B illustrates the poor correlation between this simulation and the corresponding experimentally acquired time course.

There are two important points made in Figure 5. 1) CO binding following rapid mixing is more complex than what is predicted from Equations 1 and 2. 2) The faster of the two observed rate constants corresponds with those predicted by Equation 2. Thus rHb1 exists in a mixture of states, with a fraction of the protein behavior described by Equations 1 and 2. As shown in Table 1, this fraction comprises approximately 30% of binding to rHb1. Evidently this population of molecules exists in rapid equilibrium between HbH, and HbP, behaving in accordance with the mechanism shown in Figure 1.

This suggests that the remaining 70% of rHb1 molecules exist as a discrete population, indicating that the overall binding reaction is not at rapid equilibrium and must therefore contain a slowly reacting species. We have defined the population not described by the reaction in Equation 1 to be a "closed" form of the protein. In rHb1, CO binding to the majority of the closed protein (50% of total binding) is relatively rapid and dependent on ligand concentration. The remaining binding (20% of total binding) is the slow heterogeneous phase mentioned above.

**A model for ligand binding**—There are two realistic models that could explain this closed, slow binding population of protein. 1) Hexacoordination by His73 is heterogeneous and results in two phases of ligand binding. 2) The closed form of the protein involves a conformation that must shift to an "open" form (that associated with the 30% binding
corresponding with Figure 1 and Equation 2) prior to His\(^{73}\) dissociation and formation of the reactive pentacoordinate species. The time courses for hexacoordination and ligand binding following flash photolysis are exactly those expected for a mechanism comprising a single hexacoordinating species \((12)\). Furthermore, the structure of rHb1 suggests a homogenous hexacoordinate conformation \((10)\). With no support for heterogeneous hexacoordination, the latter of the two models was selected as a hypothesis for ligand binding to rHb1.

Equation 3 describes this kinetic mechanism:

\[
\begin{align*}
Hb_C & \leftrightarrow Hb_{O,H} \leftrightarrow Hb_{O,P} \rightarrow Hb_{O,L} \\
& \quad \text{Equation 3.}
\end{align*}
\]

where Hb\(_C\) is the closed form of the protein; Hb\(_{O,H}\) is the open hexacoordinate form (corresponding with Figure 1A); Hb\(_{O,P}\) is the open pentacoordinate form (corresponding with Figure 1B); and Hb\(_{O,L}\) is the open ligand-bound form of the protein (corresponding with Figure 1C). Since Hb\(_C\) is not at rapid equilibrium with the open forms of the protein, the rate constant for opening (\(k_O\)) must be slow in comparison with the rate constants for hexacoordination and bimolecular ligand binding. Under these conditions, ligand binding following flash photolysis will involve only the open forms of the protein, and ligand binding following rapid mixing will be heterogeneous. This interpretation is in agreement with both the data presented here (Figure 2) and that determined previously from flash photolysis \((12)\).

**Kinetic analysis of Hb\(_C\)** Equation 3 can be used in combination with steady state approximations for Hb\(_{O,H}\) and Hb\(_{O,P}\) to calculate the observed rate constant for ligand binding to Hb\(_C\) following rapid mixing.

\[
k_{obs,C} = \frac{k_{-H}k_Ok'_L[L]}{k_Hk_{-O} + k_O[k_{-H} + k_H] + [k_O + k_{-O} + k_{-H}]k'_L[L]} \quad \text{Equation 4.}
\]
In this equation, each rate constant is defined as in Equation 3. Similar to Equation 2, Equation 4 predicts concentration dependent ligand binding unless \( k_O \) is very small in comparison to the other rate constants. Under these circumstances the limiting value of \( k_{obs,c} \) will be exactly equal to \( k_O \).

To test the model described in Equation 3, the dependence of \( k_{obs,C} \) on CO concentration was measured. These data are shown in Figure 6A. Because \( k_{obs,C} \) approaches an asymptote, it was possible to extract values for the rate constants \( k_O \) and \( k_{-O} \) from a least squares fit to Equation 4 with \( k_{-H}, k_{H}, \) and \( k_{',L} \) fixed at the values previously determined by flash photolysis. The fitted values for \( k_O \) and \( k_{-O} \) are 64 s\(^{-1}\) and 640 s\(^{-1}\), respectively.

Using these values and the rate constants determined from flash photolysis, numerical integration of the kinetic scheme in Equation 3 was used to simulate rapid mixing time courses for CO binding to rHb1. Figure 6B shows simulated time courses for CO binding at concentrations of 5 and 500 µM along with the corresponding observed time courses. The correlation between the simulated and observed traces at high and low [CO] in Figure 6B is in sharp contrast to Figure 5B, demonstrating the ability of the model in Equation 3 to predict CO binding to rHb1 with greater accuracy than the model in Equation 1.

The values of \( k_O \) and \( k_{-O} \) measured for rHb1 predict an equilibrium constant for opening, \( K_O (k_O/k_{-O}) \), of 0.1. Our analysis of this reaction scheme and the fitted values for \( k_O \) and \( k_{-O} \) can be tested by comparing this kinetically determined value of \( K_O \) with the equilibrium value calculated from the fraction of protein which exists in the closed form (70%, Table 1). The relationship between the fraction of closed protein and \( K_O \) is derived beginning with Equation 5, which defines \( \gamma_c \) as the fraction of protein in the closed form prior to rapid mixing.
\[ \gamma_C = \frac{Hb_C}{Hb_{O,P} + Hb_{O,H} + Hb_C} \]  \hspace{1cm} \text{Equation 5.}

When the concentration of each species in Equation 5 is normalized to \( Hb_{O,P} \), the following equilibrium relationships exist; \( Hb_{O,P} = 1 \), \( Hb_{O,H} = K_H \), and \( Hb_C = K_H/K_O \) (where \( K_H \) is \( k_H/k_{H,-} \)). Solving for \( K_O \) yields the following relationship between \( K_O \) and \( \gamma_C \):

\[ K_O = \frac{K_H (1 - \gamma_C)}{\gamma_C (1 + K_H)} \]  \hspace{1cm} \text{Equation 6.}

Applying Equation 6 to rHb1 (\( \gamma_C = 0.70 \); Table 1) with \( K_H \) determined from flash photolysis (\( K_H = k_H/k_{H,-} = 0.27 \)) results in a value for \( K_O \) of 0.1. This provides equilibrium corroboration of our kinetically determined value of \( K_O \), and further supports the model proposed in Equation 3.

**Determination of the necessity of hexacoordination in the formation of Hb\(_C\)-** To investigate whether the closed protein form is contingent upon hexacoordination, ligand binding was measured using a pentacoordinate mutant rHb1 protein in which His\(^{73}\) was replaced with Leu (H73L rHb1). This protein has previously been shown to be pentacoordinate, and has a very rapid bimolecular rate constant for CO binding following flash photolysis (\( k_{CO} = 161 \) \( \mu \text{M}^{-1}\text{s}^{-1} \)) (12, 13).

If ligand binding to H73L rHb1 was a simple bimolecular event like that of traditional pentacoordinate hemoglobins, the rapid mixing reaction would be finished in the mixing dead time at [CO] > 5 \( \mu \text{M} \). However, binding to H73L rHb1 is observed for as long as 100 s at much higher [CO] (Figure 2). This implies that hexacoordination is not required for formation of Hb\(_C\).
As the H73L rHb1 mutant protein is capable of existing in a closed conformation and yet incapable of hexacoordination, the following reaction scheme serves to describe ligand binding to this protein.

\[
Hb_C \overset{k_o}{\leftrightarrow} k_{-O} \overset{k_L[L]}{\rightarrow} Hb_L \quad \text{Equation 7.}
\]

Equation 7 is of the same form as Equation 1, and using a steady state approximation for \( Hb_{O,P} \) provides a similar relationship between \( k_{\text{obs}} \) and the rate constants for opening and closing in the H73L rHb1 mutant protein.

\[
k_{\text{obs,H73L}} = \frac{k_{-O}k_L[L]}{k_{-O} + k_O + k'_L[L]} \quad \text{Equation 8.}
\]

Rate constants for opening and closing in H73L rHb1 were determined from the concentration dependence of \( k_{\text{obs,H73L}} \) on the binding of CO following rapid mixing. These data are shown in Figure 7 along with a non-linear least squares fit to Equation 8 with \( k'_L = 161 \ \mu\text{M}^{-1}\text{s}^{-1} \). The resulting values of \( k_O \) and \( k_{-O} \) for H73L rHb1 are 80 s\(^{-1}\) and 8,000 s\(^{-1}\), respectively.

**Ligand binding to other hexacoordinate hemoglobins**—The amplitudes for the three fractions of binding extracted from each time course in Figure 2A are reported in Table 1. The percent of total ligand binding which occurs in each fraction varies widely amongst the hexacoordinate hemoglobins. Binding of CO to synHb and H73L rHb1 on the < 200 ms time scale can be described by single exponential values because binding to the open fraction is lost in the dead time associated with rapid mixing (12) (Hvitved et al. unpublished). Therefore, no "open" fraction binding is reported.
Discussion

Structural and kinetic investigations of pentacoordinate hemoglobins have been carried out for decades. The resulting mechanisms for ligand binding to these proteins indicate that the distal heme pocket uses specific amino acid side chains to stabilize the ferrous heme iron against oxidation and to create an electrostatic environment appropriate for ligand binding and selectivity (21). The use of reversible intramolecular hexacoordination to regulate ligand binding is a more recent discovery, and is now known to exist in a large and phylogenetically diverse group of proteins.

The ligand binding model presented here is undoubtedly applicable to many of these hexacoordinate hemoglobins. In demonstration of this, we discuss two hexacoordinate hemoglobins that have been previously investigated; the nsHb from barley, and an oxygen sensing heme protein (Dos) from *E. coli* (9, 22). Barley nsHb and rHb1 share significant sequence homology and would be expected to have similar rate constants for ligand binding. However, in contrast to the bimolecular rate constant for CO binding to rHb1 (6 µM⁻¹ s⁻¹), the barley rate constant has been reported as 0.57 µM⁻¹ s⁻¹ (22). This value was estimated from stop flow measurements at low [CO]. Values of $k'_{CO}$ were estimated from the slope of $k_{obs}$ at low [CO] with the assumption that this CO concentration range provided a direct measurement of bimolecular binding. Equation 2 (and Equation 4) demonstrate that this can not be the case. When we examine the data for CO binding to rHb1 in Figure 6A with the same methods used by Duff *et al.* (22) we find a similar value of 0.58 µM⁻¹ s⁻¹. Flash photolysis experiments have shown that this does not represent the true value of $k'_{CO}$ (12).

Dos, the oxygen sensing heme protein from *E. coli* described by Delgado-Nixon *et al.* (9) is also hexacoordinate although it is predicted to be a PAS protein instead of a globin. This protein shows slow and remarkably similar binding rates for O₂, CO, and NO when
these reactions are initiated by rapid mixing. It is probable that these experiments are not measuring ligand binding that occurs according to a mechanism like that described in Equations 1 and 2. Instead it is likely they are measuring the reactions described in Equations 3 and 4, with the "opening" event limiting the observed rate of reaction. This hypothesis would explain the very unusual, ligand-independent kinetic binding constants that have been reported. To determine the validity of this hypothesis, the rate constants reported from stopped flow experiments should be compared with values determined from laser flash photolysis. The model presented here resolves the unexpected discrepancy of different ligand binding rates displayed by two very similar nsHbs, and also supplies a possible explanation for the unusual ligand binding behavior of Dos.

**Structural determinants of heterogeneous ligand binding**—Hexacoordinate hemoglobins exist in at least two different states. The term “open” has been applied to the conformation of the protein that is capable of reversible hexacoordination, formation of a pentacoordinate complex, and ligand binding. A “closed” conformation has been attributed to the fraction of protein that reacts more slowly following rapid mixing. In this model, the closed conformation is incapable of ligand binding. The open and closed forms of the hexacoordinate hemoglobins examined here are not at rapid equilibrium on the time scales during which ligand binding is traditionally measured. A conversion between the two forms of the protein may involve structural rearrangement. The structure of ferric rHb1 when compared with leghemoglobin suggests that folding of the rHb1 D helix region could accompany E helix rearrangement following ligand binding (10). It is possible that this reorganization in structure is linked to the slow conformational change observed here.

Surprisingly, hexacoordination is not required for formation of the closed protein conformation. If this were the case, ligand binding to H73L rHb1 following rapid mixing
would be extremely fast and not measurable on conventional stopped flow time scales. Figure 2 and Table 1 demonstrate that this is not the case for at least part of the binding reaction. These results suggest that the ability to exist in both conformations is coded in the polypeptide backbone and is not simply a consequence of a well positioned distal His side chain. Therefore, this conformational change could confer regulation of ligand binding in concert with hexacoordination. It is also possible that ligand binding is predominately regulated by varying the ratio of open and closed fractions of protein. A great deal of regulatory control over ligand binding is possible through a combination of reversible hexacoordination and variation in fractions of open and closed protein.

Additional evidence supporting folding instead of hexacoordination in the regulation of open and closed conformations comes from the fact that the ratios of the closed fractions for rHb1 and H73L are the same (Table 1). In both proteins, ~75% of the closed fraction is fast and mono-exponential. Future work should investigate whether the same phenomenon occurs with the moss protein, where the fraction of this fast-reacting closed conformation is much smaller (20%; Table 1).

Physiological significance of hexacoordination—Traditional pentacoordinate hemoglobins have presumably evolved to bind ligands rapidly (i.e. > 100 s\(^{-1}\) in air) due to their roles in oxygen transport and scavenging. To ensure this ability, myoglobins and mammalian hemoglobins have developed rigid, "open" heme pocket conformations that create minimal resistance to entering ligands. Regulation of overall affinity is thus a function of the electrostatic environment in the heme pocket and the innate reactivity of the heme iron (which can be affected by electrostatics and proximal coordination) (21, 23). These biophysical properties vary among hemoglobins and are tailored to the specific physiological needs of the organism. For example, the oxygen affinity of human hemoglobin is highly
regulated for proper transfer between the lungs and other tissues. In plant leghemoglobins, the affinity and kinetics of the reaction with oxygen are ideally suited for the unique function of scavenging and delivery required during symbiotic nitrogen fixation (24). It is probable that both reversible hexacoordination and slow conformational changes are linked to physiological function in the hexacoordinate hemoglobins.

There are two fundamental possibilities for the relationship of hexacoordination and slow conformational changes to physiological function. 1) These reactions may exist to regulate the biochemical reactivity of hexacoordinate hemoglobins. 2) Hexacoordination and conformational changes may be moderators in a signal transduction response. These two possibilities are not mutually exclusive. In support of the first possibility are the unique properties of nsHbs. These proteins have unusually low oxygen dissociation rate constants in combination with moderately rapid bimolecular association rate constants. They also discriminate unusually well against the binding of CO in favor of oxygen (25). It is possible that these properties are a result of hexacoordination and slow conformational changes.

In support of the second possibility, it has been shown that rHb1 quaternary structure is affected by ligand binding (26). Furthermore, the discovery of hemoglobins with catalytic domains solidifies the potential for involvement in signal transduction pathways (7). These “heme-based sensors” may serve to coordinate an organism's response to varying levels of a number of possible heme ligands. The ligand binding heterogeneity reported here indicates that different fractions of protein are present at the same time. As these forms are differentially reactive, changes in $K_O$ could regulate the fraction of protein present in an “active” form. The variability in $K_O$ observed for the different hexacoordinate hemoglobins may reflect the different responses required by each organism to a particular ligand. Slow conformational regulation has been demonstrated in other protein systems. One example
which bears similarity to that proposed here is the kinetic model describing ligand binding to bacterial histidine permease (27). In this system, the histidine binding protein HisJ can exist in different conformations which affects histidine transport specificity.

The variety of organisms containing hexacoordinate hemoglobins suggests these proteins serve important roles in all living organisms. The complexity in their mechanism for ligand binding indicates that their reactivity has the potential for a large degree of regulation. A thorough understanding of these reactions in a number of different hemoglobins will be invaluable for correlating biophysical and physiological function. The model presented here provides a mechanistic basis on which to build an understanding of the structure and function of this class of hemoglobins.

Acknowledgements

We thank Drs. Cyril Appleby and John S. Olson for careful readings of this manuscript and many helpful suggestions. We also appreciate the use of the rapid scanning stopped flow spectrophotometer in the laboratory of Dr. James Espenson at Iowa State University.
References


Table 1. The relative percentage associated with each fraction of CO binding to the hexacoordinate hemoglobins in Figure 1. These fractions were determined using reactions at 30 µM [CO]. H73L and synHb exhibit only the slower fractions because binding to the open fraction is lost in the dead time of the reaction (Hvitved et al., unpublished).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Open fraction</th>
<th>Closed fraction <em>fast</em></th>
<th>Closed fraction <em>slow</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>rHb1</td>
<td>30%</td>
<td>50%</td>
<td>20%</td>
</tr>
<tr>
<td>msHb</td>
<td>10%</td>
<td>20%</td>
<td>70%</td>
</tr>
<tr>
<td>ckHb</td>
<td>20%</td>
<td>30%</td>
<td>50%</td>
</tr>
<tr>
<td>synHb</td>
<td>-</td>
<td>80%</td>
<td>20%</td>
</tr>
<tr>
<td>H73L</td>
<td>-</td>
<td>75%</td>
<td>25%</td>
</tr>
</tbody>
</table>
Figure 1. Ligand binding to rHb1. (A) The ferrous deoxy form of this protein has a His side chain coordinated to the ligand binding site which must dissociate to form the reactive pentacoordinate species. This heme pocket is that of rHb1 (10). (B) The heme pocket of deoxy lupin leghemoglobin (28) serves as a model for deoxy rHb1. (C) CO bound lupin leghemoglobin (29) serves as a model for the heme pocket of ligand bound rHb1. The rate constant for each of these reactions is titled along with the arrow indicating the reaction. The relative position of the highly conserved phenylalanine side chain is shown in each of the heme pockets.
Figure 2. CO binding time courses for several hexacoordinate hemoglobins and myoglobin. (A) The change in absorbance at the CO absorbance peak following rapid mixing of each protein at 30 µM [CO] is plotted against time. To facilitate comparison, the time axis is linear for the first 200 ms and log scale for the remainder. The fitted curves are a single exponential fit for Mb, double exponential fits for synHb and H73L rHb1, and three exponential fits for the rest. This figure serves to illustrate the diversity and heterogeneity in ligand binding exhibited by hexacoordinate hemoglobins. (B) Absorbance spectra following rapid mixing of synHb at 30 µM [CO]. Individual spectra are labeled with their time of acquisition following the rapid mixing event. This figure demonstrates a single transition from the deoxy protein to the CO bound form.
Figure 3. Different fractions of ligand binding to rHb1. This is a time course for CO binding to rHb1 at 30 µM [CO]. The three exponential fitted curve has been partitioned into its component single exponential fitted curves which are overlaid on corresponding regions of the time course. The bottom curve represents the open fraction, the middle curve is the faster of the “closed” fractions, and the top curve shows the heterogeneous, “slow” fraction of ligand binding. Fractions of binding resulting from fits to the other hexacoordinate hemoglobins in Figure 1 are given in Table 1.
Figure 4. Time courses for CO binding to rHb1 at high and low [CO].  (A) A time course for CO binding to rHb1 at 20 µM [CO].  The data are plotted along with double (solid line) and single (dotted line) exponential fits to the time course.  (B) Residuals from the single (dotted line) and double (solid line) exponential fits from (A).  (C) CO binding to rHb1 at 400 µM [CO].  As in (A), the data were fit to single (dashed line) and double (solid line) exponential expressions.  The residuals from each fit are plotted above the time course.  These data illustrate that time courses below 50 µM [CO] are biphasic on this time scale.  At [CO] > 50 µM, the time courses are monophasic.
Figure 5. A comparison of rate constants for rapid mixing to those predicted from the individual rate constants for hexacoordination. (A) The fast (filled circles) and slow (open circles) rate constants obtained from the double exponential fits to the time courses for CO binding to rHb1 are shown as a function of [CO]. The faster of the two observed rate constants becomes unmeasurable at [CO] > 50 µM due to the mixing dead time. The predicted values of $k_{\text{obs,H}}$ from Equation 2 are simulated as a function of [CO] and shown as a solid line. There is a very good correlation between the faster of the two observed rate constants and the values associated with hexacoordination and bimolecular ligand binding. (B) A time course resulting from rapid mixing of rHb1 at 400 µM [CO] (squares) is plotted with its single exponential fitted curve (solid line running through the data points) and a simulated data trace predicting rHb1 binding at 400 µM in accordance with Equation 2. A clear difference exists between the predicted and observed time courses.
Figure 6. The concentration dependence of CO binding to “closed” rHb1. (A) The slower rate constant for CO binding in Figure 5 is shown again here with the ordinate axis scaled appropriately. Each data point is an average of several replications of rapid mixing experiments at each concentration of CO. Error bars are the variance in each data point. The solid line is a least squares fit for $k_O$ and $k_{-O}$ in Equation 4 with values for $k_{H}$, $k_{H}$, and $k_{CO}$, fixed at 1911 s$^{-1}$, 517 s$^{-1}$, and 6.0 µM s$^{-1}$, respectively (12). The fitted values for $k_O$ and $k_{-O}$ are 64s$^{-1}$, and 640s$^{-1}$, respectively. (B) The program KINSIM4.0 was used to generate simulated time courses for rHb1 rapid mixing at [CO] of 500 µM and 5 µM according to the mechanism described by Equation 3. The corresponding experimentally acquired time courses at the same concentrations were then normalized and plotted within this figure. This indicates that a more accurate description of binding is provided by Equation 3 than by Equation 1 (Figure 5B).
Figure 7. The concentration dependence of CO binding to H73L rHb1. The concentration dependence of CO binding to H73L rHb1. Each data point is an average of several replications of rapid mixing experiments at each concentration of CO, and the error bars are the variance in each data point. The solid line is a least squares fit for $k_O$ and $k_{O^{-}}$ in Equation 8 with $k_{CO}$ fixed 161 $\mu$M$^{-1}$ s$^{-1}$ (12). The values of $k_O$ and $k_{O^{-}}$ extracted from this fit were 80s$^{-1}$ and 8000s$^{-1}$, respectively.
CHAPTER 3: HUMAN NEUROGLOBIN, A HEXACOORDINATE HEMOGLOBIN
THAT REVERSIBLY BINDS OXYGEN

A paper published in the Journal of Biological Chemistry

James T. Trent III, Richard A. Watts, and Mark S. Hargrove

Abstract: Neuroglobin is a newly discovered mammalian hemoglobin that is predominately expressed in the brain (Burmester et al., 2000. Nature, 407, 520-523). Sharing less than 25% identity with other vertebrate globins, neuroglobin also has less than 30% identity to the annelid nerve myoglobin that it bears closest resemblance to among the known hemoglobins. Spectroscopic and kinetic experiments with the recombinant protein indicate that human neuroglobin is the first example of a hexacoordinate hemoglobin in vertebrates, and is similar to plant and bacterial hexacoordinate hemoglobins in several respects. The ramifications of hexacoordination, and potential physiological roles are explored in light of the determination of an O₂ affinity that precludes neuroglobin from functioning in traditional O₂ storage and transport.

Abbreviations: Hbs-hemoglobins, Mbs-myoglobins, NHb-neuroglobin, hxHbs-hexacoordinate hemoglobins, nsHbs-nonsymbiotic plant hemoglobins, trHbs-truncated hemoglobins, rHb1-rice nonsymbiotic plant hemoglobin 1, Lba-Leguminous plant hemoglobins.

---

Introduction

Hemoglobins (Hbs) are a large family of proteins traditionally associated with the storage and transport of oxygen. The reactions associated with the reversible binding of oxygen and other ligands to Hbs have served as the basis for developing methods and models still in use today by investigators interested in the dynamics of protein-ligand interactions (1-3). Yet, despite a history of extensive study, current research is revealing aspects of this venerable protein family that necessitate reevaluation of the old Hb paradigms, from biophysical and physiological perspectives alike.

Hbs are generally described as proteins that bind heme \( b \) (iron protoporphyrin IX) using a highly conserved alpha helical “globin” fold. The iron atom in a Hb heme prosthetic group normally exists in either the ferrous (\( \text{Fe}^{2+} \)) or ferric (\( \text{Fe}^{3+} \)) redox state, and has an octahedral coordination geometry with six potential binding sites. Heme pyrrole nitrogens occupy the four equatorial binding sites, and one of the axial sites is coordinated by a His side chain that is invariant in all known Hbs. This “proximal” coordinate-covalent bond stabilizes the protein structure by slowing heme dissociation, concomitantly influencing the reactivity of the heme iron on the “distal” side. Binding of the proximal His residue leaves one remaining axial coordination site on the distal side of the heme available for ligand binding. The occupation of this binding site is commonly described using the terms “pentacoordinate” and “hexacoordinate”; if the distal site is empty, the heme iron has five of six binding sites occupied and is therefore pentacoordinate.

The most familiar examples of Hbs are the two vertebrate globins found in humans and other mammals. Human “hemoglobin” is a tetramer of Hb domains found in red blood cells, where it transports oxygen from the lungs to other tissues. Human myoglobin is a monomeric Hb that stores oxygen in muscle tissue and facilitates its diffusion from the red
blood cells to mitochondria for use in oxidative phosphorylation (4). Both of these proteins use a combination of mechanisms to regulate their oxygen binding and dissociation rate constants. This creates affinity levels that are appropriate for transport under the differing physiological conditions within which each protein functions. One of these mechanisms involves a change in the strain on the bond joining the proximal His and heme iron upon oxygen binding. This change alters the reactivity of the heme iron, and forms the root of allostery in human hemoglobin (5,6). The volume and electrostatic properties of the distal heme pocket also modulate ligand specificity and binding affinity (3,7). In combination with a variety of quaternary structural arrangements and other subtle variations of the globin structure, these mechanisms produce the wide range of oxygen affinities observed in Hbs from various organisms (8,9). However, among these variations, a pentacoordinate heme iron in the deoxygenated state was held to be a common characteristic due to the perceived need for an unhindered binding site for rapid ligand association and stabilization. Recent studies are forcing a reevaluation of this perception.

The last several years have seen a large number of new heme proteins discovered in a range of organisms representing all kingdoms of life (10-15). *Homo sapiens* recently joined this growing list of organisms with the identification of a third human hemoglobin termed “neuroglobin” (NHb) because of its predominante expression in the brain (16). Early biophysical study of several new proteins revealed surprising results; some of the Hbs formed reversible, intramolecular hexacoordinate complexes (10,17,18). Instead of simple bimolecular reactions as are observed with traditional pentacoordinate Hbs, ligand binding to hexacoordinate Hbs involves a competition between the exogenous ligand and intramolecular coordination by a local amino acid side chain (19). The potential for significant modulation of kinetic behavior and a common occurrence among recently discovered Hbs and other
heme proteins suggest that hexacoordination is a new mechanism for regulating ligand binding to heme proteins.

Biophysical properties of hexacoordinate hemoglobins (hxHbs) from organisms representing three of the kingdoms have thus far been investigated. The presence of hxHbs appears ubiquitous in plants (the “non-symbiotic Hbs” or nsHbs), with most species containing two or three different nsHbs that each display unique expression patterns (20,21). The cyanobacterium *Synechocystis* contains a hxHb that is also a member of the “truncated” family of hemoglobins (17; Hvitved et al., submitted). Truncated hemoglobins (trHbs) are the smallest members of the Hb family, with an abbreviation of the classic globin fold and average molecular weights under 14Kd (22). An algal hxHb with considerable homology to the trHbs has also been identified in *Chlamydomonas* (23). No physiological functions have yet been determined for these proteins and the oxygen dissociation behavior of hxHbs suggests that they are not involved in oxygen storage or transport (24). However, there is circumstantial support for potential roles in terminal oxidase activity, nitric oxide destruction, or sensing of oxygen and other gaseous ligands (10,25).

The precedents set by earlier discoveries of hxHbs, coupled with minimal primary sequence homology to vertebrate globins and suggestive absorbance spectra from mouse neuroglobin (16), evoked the possibility that NHb might be a human hxHb. This paper describes our exploration of that possibility through the investigation of coordination character and ligand binding behavior in NHb.

**Materials and Methods**

*Cloning of Human Neuroglobin*— Online sequence databases (GeneBank, IMAGE consortium, Incyte Lifesequence Gold) were screened for EST sequences with homology to
the published DNA sequence for human neuroglobin (accession number AJ245944). An EST clone (ID 3973744) from Incyte Genomics contained the most complete copy of NHb, having base pairs 41-456 of the gene Cdna. Oligonucleotide primers were designed to add base pairs 1-40, and to incorporate NdeI and EcoRI restriction sites at the 5’ and 3’ ends of the gene respectively. The complete NHb Cdna was synthesized by PCR using these primers, and then cloned into the expression vector Pet29a (Novagen) for recombinant protein generation.

**Recombinant Protein Generation and Spectroscopy**- Human NHb was expressed by the host strain *E. coli* BL21(λDE3)-CodonPlus-RP (Stratagene) in a fermentation apparatus described previously (26). Expression media consisted of 20 L of Terrific Broth supplemented with 50 μg/ml kanamycin, 30 μg/ml chloramphenicol, 3 μg/ml hemin chloride (solubilized in 0.1M NaOH and added post-induction), and 1.0 ml antifoam (Sigma). The expression strain containing neuroglobin was cultured at 37°C to an optical density of 1.7 at 600 nm, then expression was induced by the addition of IPTG (Sigma) to a final concentration of 400 μM. Protein expression was allowed to continue for six hours, whereupon the culture was harvested through formation of a cell pellet by centrifugation (~5000 g for 10 minutes.)

Following cell lysis, neuroglobin was purified in a four-stage process involving ammonium sulfate fractionation (20% and 60%), phenyl-sepharose column chromatography (1.25M (NH₄)₂SO₄ binding, 100Mm (NH₄)₂SO₄ elution), DEAE-cellulose column chromatography (the protein was bound in the absence of NaCl, and eluted in 100Mm NaCl), and a final size exclusion chromatography step (Sephacryl S-200). The protein solution was dialyzed into 20 Mm Tris (Ph 8.0) after the phenyl-sepharose separation and before the DEAE-cellulose purification step. A protease inhibitor cocktail (Sigma) was included in the
lysis buffer (50Mm Tris, 1Mm DTT Ph 8.0) and all purification buffers except that used in
the final SEC purification step. The protein was maintained at 4°C at all times, excluding
samples undergoing kinetic or spectroscopic analysis. Protein expression was monitored
with SDS-PAGE, and the purification process was evaluated through both SDS-PAGE and
spectroscopic analysis of Soret/280 nm ratios. Absorbance spectra were collected with a
Varian Cary 50 Bio spectrophotometer. The protein sample buffer consisted of 100Mm
potassium phosphate at 20°C and a Ph of 7.0 unless otherwise specified. Reduced protein
spectra were obtained in N₂ sparged sample buffer following reduction of the protein with
sodium dithionite.

**Kinetic measurements**- Rapid mixing experiments were conducted with a Durrum
stopped flow apparatus and OLIS data collection software using previously described
methods (27,28). Oxygen dissociation rate constants were determined using the ligand
displacement reaction by mixing oxygenated samples with carbon monoxide. Direct
observation of oxygen dissociation rate constants by mixing oxygenated samples with
solutions of carbon monoxide and sodium dithionite produced values equivalent to those
calculated using the ligand displacement reaction.

The flash photolysis apparatus, and the methods used with the CO samples to
determine hexacoordination and bimolecular CO association rate constants, have been
described previously (19,27). Complementary experiments with oxygen as the exogenous
ligand were carried out as follows. By varying the ratio of N₂ and O₂ with gas flow tubes, a
stream of the desired [O₂] was generated. This gas stream was used to equilibrate the buffer
reservoir of a small Sephadex G25 desalting column. A few grains of sodium dithionite were
used to reduce a protein sample (30 µl of 1 Mm protein), which was then passed over the
equilibrated desalting column. Oxygenated samples were collected in sealed 550µl glass
cuvettes that had been flushed with the appropriate atmosphere, then used in laser flash photolysis experiments. The photolysis beam in the oxygen experiments was generated with a Candela liquid dye laser using 30 µM rhodamine 575 dye in methanol as the lasing medium. Curve fitting analysis and generation of figures were done using the program Igor Pro (Wavemetrics, Inc).

Results

Ligand binding and hexacoordination - Figure 1 is an overlay of the normalized spectra of reduced, deoxygenated sperm whale myoglobin (Mb) and human NHb. The split peak in the visible region, similar to cytochrome b5, is characteristic of a hexacoordinate heme iron and a signature of the hxHb class of proteins (18,29). Table 1 provides the absorbance maximum peak wavelengths for Mb, NHb, and the nsHb from rice (rHb1). The NHb absorbance peaks are nearly identical to those of rHb1, suggesting bis-histidyl coordination in both the ferric and deoxy ferrous forms.

The following equation describes intramolecular hexacoordination and ligand binding under conditions where exogenous ligand dissociation is slow with respect to rate constants for hexacoordination and bimolecular ligand binding (26).

\[
HbH \xrightarrow{k_{-H}} HbP \xrightarrow{k'_{L}} HbL
\]  Equation 1.

Analysis of the rate constants in Equation 1 is possible by use of a method incorporating flash photolysis, and two rate constants \(\gamma_1\) and \(\gamma_2\), as has been described in detail previously (30,19). In this method, time courses for ligand rebinding are fit by a two exponential decay to generate the two aforementioned rates, \(\gamma_1\) and \(\gamma_2\). These rate constants share the following relationships with respect to the mechanism in Equation 1.
\[
\gamma_1 + \gamma_2 = k_{-H} + k_H + k_L' [L] \quad \text{Equation 2.}
\]
\[
\gamma_1 \gamma_2 = k_{-H} k_L' [L] \quad \text{Equation 3.}
\]

A plot of \(\gamma_1 + \gamma_2\) versus \([L]\) is linear, with a slope equal to \(k_L'\) and a y-intercept of \(k_{-H} + k_H\).

The slope of the plot \(\gamma_1 \gamma_2\) versus \([L]\) provides \(k_{-H}\) when divided by \(k_L'\); \(k_H\) is determined by subtracting \(k_{-H}\) from the previous y-intercept.

Time courses for oxygen and carbon monoxide binding following flash photolysis were treated according to this method. This determined the bimolecular association rate constant for each ligand to the pentacoordinate protein, and also the rate constants for association and dissociation of the distally hexacoordinating residue. Figure 2 graphically illustrates this data analysis process. Time courses for carbon monoxide association are shown as a function of \([\text{CO}]\) in Figure 2A. Each time course was fit to two exponential terms, providing the values of \(\gamma_1\) and \(\gamma_2\) used in Equations 2 and 3. The sum of these rate constants is plotted against \([\text{CO}]\) in Figure 2C, and has a slope of 38 \(\mu\text{M}^{-1}\text{s}^{-1}\) and a y-intercept of 18,000 \(\text{s}^{-1}\). A plot of \(\gamma_1 \gamma_2\) versus \([\text{CO}]\) is shown in Figure 2D, and has a slope of 310,000 \(\mu\text{M}^{-1}\text{s}^{-2}\) and a y-intercept near zero. The rate constants calculated from these values are reported in Table 2.

Time courses for oxygen binding at different \([\text{O}_2]\) are shown in Figure 2B. The biphasic nature of these data is evident, particularly at higher \(\text{O}_2\) concentrations. Plots of \(\gamma_1 + \gamma_2\) and \(\gamma_1 \gamma_2\) versus \([\text{O}_2]\) are shown along with those for CO in Figures 2C and 2D. The slopes of the oxygen data in Figure 2C indicate that the oxygen binding bimolecular rate constant is 130 \(\mu\text{M}^{-1}\text{s}^{-1}\). The y-intercepts of Figures 2C and D are within error of one another, indicating that the same values of \(k_H\) and \(k_{-H}\) were measured with both \(\text{O}_2\) and CO as exogenous ligands.
Ligand dissociation—Previously described methods (28,31) employing ligand replacement reactions initiated by stopped flow rapid mixing were used to determine the $O_2$ and CO dissociation rate constants reported for human NHb in Table 2. The dissociation rate constant for oxygen ($k_{O_2}$) is calculated from the following relationship where $k_{obs}$ is the rate constant observed for the replacement reaction, and $k'_{O_2}$ and $k'_{CO}$ are the bimolecular rate constants for oxygen and carbon monoxide binding.

$$k_{O_2} = k_{obs} \left( 1 + \frac{k'_{O_2}}{k'_{CO}} \frac{[O_2]}{[CO]} \right)$$  

Equation 4.

Reactions performed using 1000 µM [CO] as the displacing ligand and protein samples saturated under different concentrations of oxygen (262 µM $O_2$ and 1250 µM $O_2$) provided corroborating values of 0.3 s$^{-1}$ for $k_{O_2}$. CO dissociation experiments used 2000 µM NO as the displacing ligand and protein samples in less than 50 µM CO. Under these conditions, $k_{obs}$ is equivalent to $k_{CO}$ and was measured to be 0.007 s$^{-1}$ for NHb.

Discussion

As a hexacordinate hemoglobin, human NHb is a fundamentally different protein than Mb and human Hb. Neuroglobin possesses biophysical characteristics that are more in common with the bacterial and plant hxHbs. The presence of these relatively similar proteins in so many different organisms hints of a possible common physiological function that remains to be identified. The discussion presented here provides a comparison of ligand binding in NHb and several other Hbs in the context of known and possible functions of these proteins.

Hexacoordination and ligand binding kinetics—Spectroscopic analysis of a reduced, deoxygenated Hb (Figure 1) clearly distinguishes between pentacoordinate and
hexacoordinate heme proteins. The Soret peak wavelengths in Table 1 illustrate the similarity of the spectra from rHb1 and NHb, and demonstrate the difference between the coordination states of these proteins and pentacoordinate Mb. In addition to the equilibrium spectral signature, further demonstration of hexacoordinate character is shown by the biphasic time courses for ligand rebinding following flash photolysis (Figure 2). Biphasic time courses for both CO and O2 rebinding in NHb are indicative of very rapid hexacoordination rate constants. Values of $k_H$ and $k_{-H}$ are smaller in rHb1 and the bimolecular reaction with O2 out-competes hexacoordination, resulting in single exponential rebinding (19).

The bimolecular rate constants for O2 and CO binding to NHb are fast compared to the vertebrate globins and plant nsHbs, on a level similar to those of soybean Lba and the nerve Mb from Aphrodite (32,33). However, the oxygen dissociation rate constants for Aphrodite nerve Mb and soybean Lba are much faster than that of NHb. With respect to O2 dissociation NHb is more similar to the plant nsHbs, Synechocystis Hb, and the hexacoordinate Chlamydomonas Hb, all of which have slow oxygen dissociation rate constants. This particular combination of rate constants ($k_{O2} > 100 \mu M^{-1}s^{-1}$ and $k_{O2} < 1 s^{-1}$) makes NHb unique among previously characterized hxHbs, nerve Hbs, and the other human Hbs.

**Equilibrium affinity constants**—The equilibrium affinity (Kd or half saturation concentration, $P_{50}$) is the kinetic description of physiological relevance. HxHbs differ from the classic pentacoordinate Hbs, in that the effects of competition between distal residue and ligand for the binding site must be considered when describing an equilibrium affinity constant. The ligand binds with a hxHb when it is in the pentacoordinate form (HbP), according to the reaction described by the following mechanism.
\[
Hb_H^{k_H} \xleftrightarrow[k_H]{k_L^{-1}} Hb_P + L \rightarrow Hb_L^{k_L} \quad \text{Equation 5.}
\]

In this equation, \(Hb_H\) is the hexacoordinate species, \(L\) is the exogenous ligand, \(k'_L\) is the bimolecular rate constant for ligand binding to \(Hb_P\), and \(k_L\) is the ligand dissociation rate constant. The equilibrium affinity of hxHbs is a function of \(K_H\), where \(K_H\) is the proportion of protein in the hexacoor- dinate vs. pentacordinate states in a deoxygenated sample. The fraction of protein present as \(Hb_L\) can be calculated with Equation 6, where \(K'_L = k'_L/k_L\), and \(K_H = k_H/k_H\).

\[
\gamma_{Hb_L} = \frac{K'}{1 + K_H} \left[CO\right] \quad \text{Equation 6.}
\]

Equation 6 incorporates the influence of hexacoordination on the fraction of ligand-bound protein, so the effective affinity constant is described by Equation 7 (21,34).

\[
K'_{L,H} = \frac{K'}{1 + K_H} \quad \text{Equation 7.}
\]

In other words, \(P_{50}\) is now described by Equation 7 rather than by \(K_d\), and as \(K_H\) increases the ligand affinity decreases. This equation indicates that hexacoordination has the net effect of reducing the affinity of a hemoglobin for ligands, and the value of \(K_H\) determines the degree of this reduction. As \(K_H\) for NHb is equal to 1.2, ligand affinities in this protein are effectively half that expected were only the bimolecular association and ligand dissociation rate constants considered. The affinity constants reported in Table 2 were calculated using Equation 7, and reflect the effects of hexacoordination. Our value for neuroglobin oxygen affinity is \(~1,000\) times higher than that reported for mouse NHb which was determined by equilibrium binding methods (16). However, an affinity as high as that
reported here makes equilibrium binding extremely difficult to measure, as protein concentrations would need to be in the sub-nanomolar concentration range to prevent stoichiometric binding.

**Physiological significance** - The potential importance of the role of hemoglobins in facilitated diffusion of oxygen and other ligands was first investigated by Wittenberg and Wyman in the 1960s (35,36). The term facilitated diffusion was used to describe the observation that solutions of ferrous Mb (and some other Hbs) could transport oxygen along a concentration gradient more efficiently than buffer alone. However, for hemoglobins to carry out this function, they must exist in partial saturation in the transport gradient and their kinetic rate constants must not limit the speed of diffusion (37).

Mb and human Hb have oxygen affinities and protein concentrations that fall within these parameters for oxygen transport in their respective environments. Hbs with extremes in affinity need a corresponding extreme in environment to function in facilitated diffusion. Leghemoglobins are an example of this; their high affinity for oxygen ( ~ 20 times that of Mb) enables transport of oxygen in the hypoxic root nodules of leguminous plants (32,38). NHb does not meet the kinetic and equilibrium requirements for functioning in oxygen transport by facilitated diffusion under known physiological conditions. Its oxygen affinity is so high, to serve in this role would require physiological oxygen gradients in the nanomolar range. The example set by leghemoglobins indicates that a role in the facilitated diffusion of other ligands, or under unique concentration gradients, may still be possible. However, it is more likely that NHb serves in a functional role similar to those postulated for other new Hbs such as NO or O₂ scavenging, or as a sensor for gaseous ligands like CO, NO, or O₂. These roles have been recently attributed to other hemoglobins and heme proteins for which homologous proteins have not yet been identified in mammals (39,40).
The physiological significance of hexacoordination as a ligand binding regulatory mechanism is not established with any certainty, but a comparison of the hexacoordination parameters of NHb and the other hxHbs reveals the beginning of a precedent in the rates associated with the distally coordinating residue. $K_H$ varies by less than 20 fold in all three hxHbs for which this value has been measured, and both the rate constants for association and dissociation of this residue are on the order of exogenous ligand binding in each protein. For hexacoordination to serve as a mechanism in regulating exogenous ligand binding, it would have to display exactly this sort of behavior. If $K_H$ were too large, exogenous ligands could not bind. If $k_H$ and $k_{-H}$ were too small, a rapid equilibrium between hexacoordination and ligand binding would not be established. While data are still quite limited, observations thus far support the hypothesis of hexacoordination having a functionally significant role in ligand binding regulation.

**Conclusions** - The work presented here has identified human neuroglobin as the first hexacoordinate hemoglobin characterized in vertebrates. Our study of its kinetic behavior indicates that it does not function in a role similar to the other human Hbs. As no physiological role(s) is known for other hxHbs, the function of neuroglobin in the human brain remains an intriguing subject of research.

**Acknowledgements**

We thank Dr. John S. Olson for his discussions of equilibrium and hexacoordination and for his generous mentoring of M.S.H. and R.A.W. We would also like to acknowledge the USDA for support of this work (Award No. 99-35306-7833).
References


34. Watts, R. (1999) in *Division of Biochemistry and Molecular Biology*, Australian National University, Canberra, Australia


Table 1. Absorbance maxima for NHb, Mb, and rHb1

<table>
<thead>
<tr>
<th>Protein</th>
<th>Deoxy (nm)</th>
<th>Oxy (nm)</th>
<th>CO (nm)</th>
<th>Ferric (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHb</td>
<td>424</td>
<td>414</td>
<td>417</td>
<td>413</td>
</tr>
<tr>
<td>Mb</td>
<td>434</td>
<td>418</td>
<td>423</td>
<td>409</td>
</tr>
<tr>
<td>rHb1</td>
<td>424</td>
<td>412</td>
<td>416</td>
<td>410</td>
</tr>
</tbody>
</table>

Table 2. Rate constants for ligand binding to NHb and other hemoglobins.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$k'_O2$ $\mu M^{-1}s^{-1}$</th>
<th>$k_O2$ $s^{-1}$</th>
<th>$bK_{O2}$ $\mu M^{-1}$</th>
<th>$k'_CO$ $\mu M^{-1}s^{-1}$</th>
<th>$k_CO$ $s^{-1}$</th>
<th>$bK_{CO}$ $\mu M^{-1}$</th>
<th>$k_H$ $s^{-1}$</th>
<th>$k_H$ $s^{-1}$</th>
<th>$K_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human NHb</td>
<td>130</td>
<td>0.3</td>
<td>200</td>
<td>38</td>
<td>0.007</td>
<td>2,500</td>
<td>8,200</td>
<td>9,800</td>
<td>1.2</td>
</tr>
<tr>
<td>Hb α-chain R-state&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19</td>
<td>15</td>
<td>1.3</td>
<td>2.2</td>
<td>0.009</td>
<td>245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb β-chain R state&lt;sup&gt;c&lt;/sup&gt;</td>
<td>74</td>
<td>47</td>
<td>1.6</td>
<td>5.9</td>
<td>0.01</td>
<td>590</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphrodite nerve Mb&lt;sup&gt;d&lt;/sup&gt;</td>
<td>170</td>
<td>360</td>
<td>0.47</td>
<td>21</td>
<td>0.1</td>
<td>210</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice rHb1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>68</td>
<td>0.038</td>
<td>1,400</td>
<td>8</td>
<td>0.001</td>
<td>5,150</td>
<td>1,900</td>
<td>520</td>
<td>0.3</td>
</tr>
<tr>
<td>Synechocystis Hb&lt;sup&gt;f&lt;/sup&gt;</td>
<td>240</td>
<td>0.014</td>
<td>3,100</td>
<td>90</td>
<td>0.002</td>
<td>8,200</td>
<td>930</td>
<td>4,200</td>
<td>4.5</td>
</tr>
<tr>
<td>Chlamydomonas Hb&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sperm Whale Mb&lt;sup&gt;h&lt;/sup&gt;</td>
<td>17</td>
<td>15</td>
<td>1.1</td>
<td>0.51</td>
<td>0.02</td>
<td>25.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybean Lba&lt;sup&gt;i&lt;/sup&gt;</td>
<td>130</td>
<td>5.6</td>
<td>23</td>
<td>16</td>
<td>0.008</td>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Hexacoordinate hemoglobins are in bold.

<sup>b</sup>Equilibrium values for the hxHbs were calculated with Equation 7.

<sup>c</sup>(41); <sup>d</sup>(33); <sup>e</sup>(19); <sup>f</sup>(Hvitved et al., submitted); <sup>g</sup>(10); <sup>h</sup>(42); <sup>i</sup>(43)
Figure 1. Absorbance spectra of deoxy, ferrous NHb and Mb. This figure presents an overlay of the normalized absorbance spectra of the reduced forms of sperm whale Mb (dashed line) and human NHb (solid line). Magnified five-fold, the visible spectral region illustrates the difference between the electronic states of pentacoordinate Mb and hexacoordinate NHb in their deoxy, ferrous forms. The principle Soret peak in the 400nm region of NHb is clearly blue-shifted with respect to Mb, as is the norm for known hxHbs (see Table 1).
Figure 2.  Bimolecular and hexacoordinate rate determinations.  (A) Time courses of CO rebinding to the pentacoordinate NHb protein following flash photolysis at differing [CO].  (B) Time courses of O$_2$ rebinding to NHb following flash photolysis for differing [O$_2$].  (C) The sum of the two rate constants extracted from (A) and (B) plotted against ligand concentration.  (D) A plot of the product of the two rate constants describing (A) and (B) versus ligand concentration.  The use of these plots to determine the kinetic constants in Table 2 is described in the Results section.
CHAPTER 4: A UBIQUITOUSLY EXPRESSED HUMAN HEXACOORDINATE HEMOGLOBIN

A paper published in the Journal of Biological Chemistry\(^1\)

James T. Trent III, and Mark S. Hargrove

Abstract: We have identified a new human hemoglobin that we call histoglobin, because it is expressed in a wide array of tissues. Histoglobin shares less than 30% identity with the other human hemoglobins, and the gene contains an intron in an unprecedented location. Spectroscopic and kinetic experiments with recombinant human histoglobin indicate that it is a hexacoordinate hemoglobin with significantly different ligand binding characteristics than the other human hexacoordinate hemoglobin, neuroglobin. In contrast to the very high oxygen affinities displayed by most hexacoordinate hemoglobins, the biophysical characteristics of histoglobin indicate that it could facilitate oxygen transport. The discovery of histoglobin demonstrates that humans, like plants, differentially express multiple hexacoordinate hemoglobins.

Abbreviations: HGb-histoglobin, NGb-neuroglobin, Hb-hemoglobin, Mb-myoglobin, hxHb-hexacoordinate hemoglobin

\(^1\) Journal of Biological Chemistry, (2002), 277(22): 19538-19545
Introduction

Hemoglobins (Hbs) are a class of proteins long associated with functions in respiration and the storage and transport of oxygen. Recent research has also focused upon a potential physiological role of Hb relating to its interactions with nitric oxide (NO) during conditions of oxidative stress. Several examples of physiologically important reactions involving Hb and NO now exist: 1) Bacterial and yeast flavohemoglobins scavenge NO resulting from oxidative bursts, protecting these cells from the host immune response to infection (1-3). 2) A critical hindrance to cell free Hb-based blood substitutes is the severe hypertension resulting from their scavenging of NO (4). 3) NO is scavenged by myoglobin in cardiac tissue (5,6). 4) The Hb from the parasitic worm *Ascaris* is reported to function as an NO-activated dioxygenase (7). 5) *Mycobacterium tuberculosis* contains a truncated Hb that is reported to be an NO scavenger (8). 6) The binding of NO to human Hb thiols has been implicated in vascular control (9). In addition, there is mounting evidence that Hbs are involved in cell survival during hypoxia. Human neuroglobin (NGb) has recently been linked to tissue protection during hypoxic-ischemia injury (10), and similar low-oxygen conditions induce the expression of a ubiquitous nonsymbiotic plant Hb (nsHb) (11,12).

Both NGb and nsHbs are members of a newly discovered class of proteins called “hexacoordinate” Hbs (hxHbs) that have been found in animals (13,14), protists (15), cyanobacteria (16-18), and all plants (19,20). In addition to potentially new physiological functions, hxHbs use an alternative mechanism to regulate ligand binding that involves reversible intramolecular coordination of the heme iron; prior to the discovery of this mechanism, an open coordination site was believed necessary for reversible ligand binding in Hbs (Figure 1). Despite this apparent hindrance to entering ligands, hxHbs are capable of
reversibly binding oxygen and other heme ligands with unusually high affinities (21). The molecular details of the hexacoordination mechanism are not known with certainty. Biophysical examination of the details of this reaction are necessary to distinguish whether hexacoordination is a mechanism for regulating ligand affinity, or a requirement for a novel biochemical reaction.

In spite of the prevalence of hxHbs, the physiological function(s) of these proteins is unknown. However, there is growing evidence linking hxHbs with NO scavenging and a protective role during hypoxia (10,21). If hxHbs serve to protect cells from damage during the generation of NO or other reactive oxygen species (ROS), their expression in a wide range of tissues could be expected. Plants are known to contain two or three different hxHbs expressed in a variety of tissues (11,12), but the human hxHb NGb is essentially expressed only in the brain (14). When coupled with a link to hypoxia in plant hxHbs and NGb, this suggests that humans may harbor more than one hxHb. Our inquiry into this possibility has discovered histoglobin (HGb), a hxHb ubiquitously expressed in human tissues with behavior unique in comparison to the other human Hbs.

Materials and Methods

Identification, Cloning, and Sequence Analysis- The histoglobin gene was identified utilizing ALLGENE (22) to mine the publicly available EST and genomic sequencing data on Mus musculus and Homo sapiens for predicted genes harboring a globin domain. The resulting data were reduced by selection for genes with a GeneBank identified expressed sequence tag (EST) clone. Of the remaining candidate genes, all coding for proteins containing over 250 residues were culled, as currently identified hxHbs are composed of
~200 or fewer amino acids. The final candidates were then evaluated based on sequence homology with the vertebrate hemoglobins, to eliminate those likely originating from splicing errors, etc. of the known globin genes. This process identified the putative gene DT.40262016 (ALLGENE identification number) as a potential novel mammalian hemoglobin. The IMAGE EST clone R87866 corresponding to DT.40262016 was purchased from Incyte Genomics. The human histoglobin cDNA sequence was found through sequencing of R87866. Intron determination and chromosome localization were performed using the public human genome database. The complete cDNA sequence has been posted by the NCBI annotation project (accession # XM058818). Oligonucleotide primers were designed to incorporate NdeI and EcoRI restriction sites at the 5’ and 3’ ends of the gene, respectively. The histoglobin cDNA was synthesized by PCR using these primers, and then cloned into the Novagen expression vectors pET29a (no-tag) and pET28a (HIS6x-tagged) to generate constructs for recombinant protein generation. The protein alignment (Figure 2A) was generated using the ClustalW algorithm, and cross checked against previously published sequence alignments of the human hemoglobins.

Recombinant Protein Generation and Spectroscopy- Human HGb was expressed in Z-competent (Zymo Research) Escherichia coli BL21(λDE3)-CodonPlus-RP cells (Stratagene), using both a previously described fermentation apparatus (23), and 2L culture flasks. Recombinant expression cultures were grown at 37°C for a period of 14 to 16 hours post-inoculation, in 2xYT nutrient medium supplemented with 50 μg/ml kanamycin. These cultures were harvested by centrifugation, and the cells lysed with two passages through an Avestin EmulsiFlex-C5 homogenizer at 25,000 psi.
The non-tagged HGb was purified from lysate using a four-stage process composed of ammonium sulfate fractionation (40% and 70%), phenyl-sepharose column chromatography (1.5M [NH₄]₂SO₄ binding, 0.75M [NH₄]₂SO₄ elution), DEAE-cellulose column chromatography (0mM NaCl binding, 50mM NaCl elution), and a final size exclusion chromatography step (Sephacryl S-200). The protein solution was dialyzed into 20mM Tris (pH 8.0) after the phenyl-sepharose separation and before the DEAE-cellulose purification step. The HIS6x-tagged protein was purified using an affinity column (Ni-NTA) followed by a desalting column (G-25). Purification was at 4°C, and a protease inhibitor cocktail (Sigma) was included in the lysis buffers (50mM Tris, 1mM DTT pH 8.0: non-tagged HGb, 50mM KPO₄, 5mM imidazole pH 7.5: HIS6X-tagged HGb) and other purification buffers, except that used in the final size exclusion or desalting step (100mM KPO₄ pH 7.0).

All protein experiments were initially conducted in quadruplicate, utilizing HIS6x-tagged and non-tagged, green and red HGb. All species of histoglobin exhibited functionally identical behavior, excluding the obvious spectral difference between the green and red protein samples. In light of this, further replications were conducted using the red versions of the protein, with the exception of those experiments explicitly describing the green protein. Reduced protein spectra were obtained in N₂ sparged sample buffer following reduction of the protein with sodium dithionite.

**RNA Hybridization Assay**- A human adult normal tissue RNA Dot Blot I (BioChain Institute) was screened using ³²P labeled probes generated with random hexamer primers and HGb or NGb cDNA. Hybridization buffer consisted of PerfectHYB Plus (Sigma) with 0.1mg/mL sheared, denatured Salmon testis DNA as a blocking reagent. The membrane was
pre-hybridized for 1 hour at 60°C, and hybridization was allowed to proceed for 12 hours at 60°C. The membrane was washed once (2X SSC, 0.1% SDS) for 5 minutes at 25°C, followed by two washes (0.5X SSC, 0.1% SDS) for 20 minutes at 60°C and a final wash (0.1X SSC, 0.1% SDS) for 10 minutes at 60°C. The membrane was then sealed in saran wrap and exposed to a phosphoimager for 30 hours. The membrane was first probed with HGb, then stripped and the identical procedure outlined above repeated using the probe generated from NGb. The NGb hybridized membrane required 48 hours exposure to the phosphoimager.

**Oxygen Sensitivity/Heme exchange experiments**- The data illustrated in Figure 4B were obtained from 1L cultures grown in 2xYT media at 37°C for 12 hours. Each culture was inoculated simultaneously, then differentially aerated by varying agitation speeds (rpm). After 12 hours these cultures were harvested, and the HGb protein purified. Protein samples were reduced with sodium dithionite, and a visible spectrum taken. The 0% saturation data illustrated in Figure 4C was obtained using a sealed 1L 2xYT culture where the media and flask were sparged with N₂ both prior to, and immediately following inoculation. The 100% saturation data illustrated in Figure 4C was obtained using a 1L 2xYT culture aerated with a 1L/min flow of pure O₂. Both cultures were agitated at 150 rpm. The heme cofactor was removed from samples of both red and green HGb using the methyl ethyl ketone method to produce the corresponding “green” or “red” apoprotein (24). Titrating the apoprotein sample with hemin chloride solubilized in 0.1M NaOH generated reconstituted holoproteins.

**Kinetic measurements**- All kinetic experiments were performed at 20°C, and protein samples buffered in 100mM potassium phosphate at pH 7.0, unless otherwise specified. Rapid mixing experiments were conducted using previously described methods (25,26).
Oxygen dissociation rate constants were measured using both the ligand displacement reaction (mixing oxygenated samples with carbon monoxide) and direct observation of oxygen dissociation rate constants (mixing oxygenated samples with solutions of carbon monoxide and sodium dithionite). The flash photolysis apparatus and the methods used to measure the hexacoordination and bimolecular association rate constants have been described previously (18, 27, 28). The program Igor Pro (Wavemetrics, Inc) was used for curve fitting and generation of figures.

Results

**HGb Identification and Tissue Expression** - The search described in the materials and methods section yielded a candidate gene corresponding to the EST sequence identified by GeneBank as R87866. Cloning and subsequent sequencing of R87866 resulted in a 573bp cDNA coding for the 190 residue protein we term histoglobin (HGb). The primary sequence identity shared by HGb with the other principal human hemoglobins is less than 30%, as is illustrated in the alignment shown in Figure 2A. Comparative sequence analysis of the HGb cDNA with the NCBI human genomic database indicates the gene is located at chromosome 17q25 on the minus strand. The HGb gene structure concomitantly deduced during this analysis indicates there are four exons interrupted by three introns, as is diagrammed in Figure 2B.

DNA hybridization results for HGb and NGb are found in Figure 3. The tissue types corresponding to each dot in this figure are listed in Table 1. Figure 3A indicates that HGb is expressed to a varying degree in all of the tissues represented in Table 1. The expression
results for NGb (Figure 3B) are in agreement with previously published data showing this protein is predominantly expressed only in the tissues of the brain (14).

**Recombinant Protein Generation and Analysis**- Small scale (<1L) cultures generating recombinant HGb were grown in shake flasks to evaluate expression levels. The cell pellets from these cultures, as well as the protein purified from them, were red in appearance. Scaling up recombinant protein production by use of the fermentation apparatus resulted in both cell pellets and HGb protein with a deep pine green color. The visible spectra of the red and green proteins are shown in Figure 4A. The green protein derives its color from an absorbance band at 630nm that is lacking in red HGb, and is irreversible in the purified protein. The principal difference between the cultures grown in the fermentation apparatus and those grown in shake flasks is the level of oxygen, as the fermentation apparatus is aerated using pure oxygen. A series of cultures were grown under conditions that varied the level of oxygen, and the relative color of the resulting proteins was assessed by spectroscopy. The data from these growths are plotted in Figures 4B and 4C. The data in Figure 4C demonstrates that it is oxygen concentration and not some other aspect of culture agitation speed that influences the generation of green HGb. A correlation between the abundance of oxygen during recombinant protein expression and the amount of green protein generated is evident.

To assess whether the difference in color originates in the heme cofactor or the globin, apoHGb was generated by removing heme from the red and green proteins. Upon removal, heme from the green protein was green and that from the red protein, red. Reconstitution of the apoprotein generated from green HGb with heme produced holoprotein having a visible spectrum identical to red HGb, as is shown in Figure 4A. This
indicates that the oxygen-level dependent modification giving rise to the green color is associated with the heme co-factor. In additional support of this is the identical size measured for both red and green proteins using SDS-PAGE (data not shown). However, this modification of the heme cofactor had no observable impact on either protein purification or kinetic characteristics of the red and green proteins.

**Hexacoordination and Flash Photolysis Kinetics**—Figure 5 presents an overlay of the absorbance spectra of reduced, deoxygenated sperm whale myoglobin (Mb), human NGb, and HGb. The split peak in the visible region is characteristic of a hexacoordinate heme iron and a signature of the hxHb class of proteins (29,30). The HGb absorbance peaks are nearly identical to those of NGb, suggesting *bis*-histidyl coordination in the deoxygenated, ferrous form.

Time courses for bimolecular ligand recombination with pentacoordinate Hbs following flash photolysis are described by a single exponential decay. However, the time courses for ligand binding to hxHbs such as NGb are not monoexponential, because intramolecular coordination competes with rebinding of the exogenous ligand as described by Equation 1. In this equation, the subscripts H, P, and L refer to the hexacoordinate, pentacoordinate, and ligand-bound forms of the Hb, respectively.

\[
Hb_{H}^{k_{H}} \overset{k_{H}}{\rightleftharpoons} Hb_{P}^{k_{P}} \rightarrow Hb_{L}^{k_{L}[L]} \quad \text{Equation 1}
\]

Analysis of ligand binding under these circumstances is possible by using a procedure described in detail previously (27,28). Time courses for oxygen and carbon monoxide binding to HGb following flash photolysis were initially analyzed using this method. Yet, unlike the obviously bi-exponential behavior exhibited by NGb (28), the ligand binding time
courses for HGb were well described by a single exponential decay. Figures 6A and 7A are the overlaid residuals from single and bi-exponential fits to the CO and O₂ rebinding time courses shown in Figure 6B and in Figure 7B. These residuals are indistinguishable from one another, and indicate that fitting these data to a bi-exponential decay is not warranted. To assess the bimolecular association rate constant for each ligand, rate constants extracted from single exponential fits were plotted against concentration as shown in Figure 6C (CO) and Figure 7C (O₂). The slopes of the linear fitted curves to these data are reported as the HGb bimolecular rate constants in Table 2.

The intercept of these linear fits should be zero if the data reflect either a simple bimolecular binding reaction (with a slow dissociation rate constant), or a reaction in which the bimolecular rate constant is substantially greater than all other binding events \(k'_L[L] \gg k_H + k_{H^+}\) (27). The linear fit to the oxygen binding data has an intercept within error of zero, indicating the reaction measured reflects only the oxygen rebinding event. However, the linear fit to the CO data in Figure 6C has a non-zero intercept at 440 s⁻¹. This suggests that the time courses giving rise to these data measure a reaction more complex than simple bimolecular rebinding of CO. As spectroscopic data indicate HGb is hexacoordinate at equilibrium, the rate constants associated with the hexacoordination reaction are a likely source of the additional complexity observed.

**Stopped Flow Rapid Mixing Kinetics-** To investigate the possibility that HGb possesses a hexacoordination dissociation rate constant \(k_{H^+}\) too slow to be measured with the flash photolysis method, CO binding was examined using rapid mixing to initiate the reaction. The time courses for CO binding to the ferrous, deoxy protein at several different concentrations of CO are shown in Figure 8A. Although [CO] ranges between 25 and
500µM, the appearance of the ligand binding time courses does not show a 20-fold change in reaction half-time. This phenomenon has been previously described in hxHbs, and is the typical behavior of these proteins (23). These time courses require fitting to a three exponential decay curve to be accurately described. The fastest of the rate components was assigned as the hexacoordination dissociation rate constant ($k_{H}$) in accordance with the mechanism described in Equation 1. The slower rate components of these reactions have been previously discussed in the context of a model for ligand binding to hxHbs (23). The fastest rate component extracted from the CO binding time courses are plotted against [CO] in Figure 8B. As these values reach an asymptote at approximately 5 s$^{-1}$, this value is reported for $k_{H}$ in Table 2. This interpretation of data indicates the non-zero intercept in Figure 6C arises from time courses with a hexacoordination contribution predominately composed of the association reaction, and this is reflected in the reported value of 430 s$^{-1}$ for $k_{H}$ in Table 2.

The HGb O$_2$ and CO dissociation rate constants reported in Table 2 were measured with ligand replacement reactions initiated by stopped flow rapid mixing (25,31). Reactions using 1000µM [CO] as the displacing ligand, and protein samples saturated under different concentrations of oxygen (262µM O$_2$ and 1250µM O$_2$) provide corroborating values of 0.35 s$^{-1}$ for $k_{O2}$. CO dissociation experiments used 2000µM NO as the displacing ligand and protein samples in less than 50µM CO. Under these conditions, $k_{obs}$ is equivalent to $k_{CO}$ and was measured to be 0.003 s$^{-1}$ for HGb.
Discussion

Human Hb and Mb are proteins whose physiological roles in oxygen transport and respiration are among the most clearly defined and well understood. Yet, hxHbs have thus far demonstrated oxygen affinities that preclude their functioning within these roles (12,15,18,21,29). From its biophysical behavior to the tissues within which it is expressed, HGb exhibits fundamental differences from the other human Hbs. The discussion that follows examines these differences in the context of potential physiological significance.

*Primary structure comparison*- Comparison of the nucleic acid and protein sequences of HGb with the other human Hbs highlights both the similarities and differences that characterize this protein. A primary sequence alignment of the human Hbs (Figure 2A) illustrates that HGb shares many of the elements that are common to the vertebrate globins, including the invariant proximal His$^{F8}$ (the eighth amino acid along the F helix with reference to the structure of myoglobin) Phe$^{CD1}$, and the distal His$^{E7}$. The amino acids comprising the heme pocket in HGb appear to have more in common with pentacoordinate Mb than with Ngb. For example, the E6 residue in the distal heme pocket of HGb and Mb is a Lys, in contrast to the Asp residue occupying this position in Ngb; the F9 position adjacent to the proximal His in Ngb contains a polar residue, as opposed to the aliphatic sidechains in this position in Mb and HGb. The similarities in the heme pocket primary structure between Mb and HGb are intriguing, particularly in the context of the nearly identical O$_2$ and CO equilibrium affinity constants for these proteins (Table 2).

Most Hb genes contain two conserved introns at positions B12-2 (between nucleotides 2 and 3 of the codon for amino acid B12) and G7-0. A third intron at position E11-0 is conserved in all plant Hbs and also found in Ngb (14,32,33). As illustrated in
Figure 2B, HGb contains three introns, as do all other hxHbs. However, while two of these introns are located at the conserved B12-2 and G7-0 positions, the location of the third intron at H36-2 is unprecedented. The three intron gene structure found in hxHbs is believed to indicate an earlier evolutionary origin than pentacoordinate mammalian Hbs (33). Perhaps HGb represents an intermediate evolutionary step between the more recently evolved pentacoordinate Hbs and other hxHbs such as NGb and the plant nsHbs.

As is the case with NGb, there are putative HGb genes in both the rat and mouse that are highly homologous to the human version. The putative mouse HGb gene arises from an EST sequence (accession #AK019410), and the rat homologue was noted in a very recent proteomic study where it was called Stellate Cell Activation associated Protein (STAP) (34). Considering the numerous genes already bearing this acronym (35-37), the characterization of HGb as a hxHb, and its expression in many tissues besides hepatic stellate cells, it seemed logical to continue our reference to this gene as HGb. The human HGb primary structure shares over 90% identity with its homologues in the rat and mouse. This high degree of conservation implies that the function of these proteins is rigidly dependent upon the specific functional properties conveyed by these particular structures.

**Hexacoordination and ligand binding**- A key element in learning the physiological role(s) held by HGb, as well as other hxHbs, is biophysical study of the attributes that define the behavior of these proteins. The spectroscopic analysis of reduced, deoxygenated HGb (Figure 5) clearly distinguishes it from pentacoordinate Mb. The Soret peak wavelengths of HGb are nearly identical to those of NGb, and these spectra illustrate the difference between the coordination states of these hxHbs and pentacoordinate Mb.
In addition to the equilibrium spectral signature, another manifestation of hexacoordinate character is biphasic time courses for ligand rebinding following flash photolysis (27,28). However, the appearance of these biphasic time courses depends upon the relationship between the rate constants of hexacoordination and bimolecular ligand-binding. If hexacoordination is out-competed by bimolecular ligand rebinding, then single exponential time courses will be observed (27). The time courses illustrated in Figure 6B and Figure 7B exhibit this single exponential form, and indicate that the rates associated with hexacoordination in HGb must be considerably slower than those of NGb (28). However, the non-zero intercept in Figure 6C suggests that CO rebinding rate constants at the lowest concentrations of CO measured were of a similar order as the HGb rate constants for hexacoordination. A more quantitative assessment of these values in HGb was possible if this intercept could be correlated with data from an independent method of measuring hexacoordination rate constants. These data were obtained using stopped-flow rapid mixing to ascertain the magnitude of the hexacoordination dissociation rate constant (Figure 8).

The intercept from Figure 6C ($-k_{II}$) and the approximate asymptote value from Figure 8B ($-k_{II}$) can be correlated with Equation 2, which describes the expected rate constant for ligand binding initiated by rapid mixing according to the mechanism described in Equation 1 (23).

$$k_{obs,II} = \frac{k_{-II}k'_L[L]}{k_{-II} + k_{II} + k'_L[L]}$$

Equation 2

A simulation of the expected rate constants was created using Equation 2 and the ligand binding and hexacoordination rate constants reported for HGb in Table 2, then overlaid on the observed data in Figure 8B. The correspondence of this simulated curve with the
observed values supports assignment of $k_{-H}$ to the fastest phase of ligand binding observed following rapid mixing.

The rate constants associated with hexacoordination in HGb are considerably smaller than those observed thus far in other hxHbs (18,27,28). The effect of hexacoordination on equilibrium affinity constants can be calculated using the following equation, where $K_{L,Pent}$ is the equilibrium constant for ligand association to the pentacoordinate form of the protein ($k'_L/k_L$), and $K_H$ is the equilibrium constant for hexacoordination ($k_H/k_{-H}$) (28).

$$K_{L,\text{effective}} = \frac{K_{L,Pent}}{1 + K_H}$$  

Equation 3

The $K_H$ reported for HGb in Table 2 is the largest hexacoordination equilibrium constant yet observed in a hxHb, and dramatically differs from the $K_H$ of NGb. With the contribution of hexacoordination factored in, the equilibrium affinity constants for HGb are very similar to those of Mb (Table 2).

**Formation of green HGb** Green heme proteins are not unprecedented, and their color can arise through several different mechanisms: 1) Myeloperoxidase contains a conventional heme co-factor that is covalently attached to the protein by two methoxy esters, and a methionine derived sulfonium linkage. Coupled with a non-planar bend in the heme, these linkages are believed to be the origin of its green color (38). 2) Biliverdin and verdoheme are green heme derivatives produced by heme oxygenase during the degradation of heme (39). 3) Sulfhemoglobin (sulfHb), a green protein associated with certain blood pathologies, owes its color to the incorporation of sulfur into the porphyrin ring, forming sulfhemin (40).
The green color in HGb is due to a heme modification, as addition of iron protoporphyrin IX to (previously green) apoHGb results in red protein. Additionally, red HGb is stable and does not degrade to the green protein. This suggests that green HGb is not a result of mechanisms 1) and 2) described above. However, it is possible that green HGb is a sulfHb. In support of this view is the fact that formation of sulfHbs is known to be dependent on the availability of oxygen (40), as we have shown to be the case with formation of green HGb. And like sulfHb, the absorbance spectrum of green HGb is very similar to that of the red protein (41), containing only the additional peak around 630nm. However, sulfmyoglobin is associated with 2500- and 10-fold reductions in O₂ and CO binding, respectively (42,43). In marked contrast to this, green HGb appears to differ little from red HGb in ligand binding behavior. It is possible that the reaction time courses we observed for green HGb are attributable to the presence of a fraction of the red protein. Yet, if this were the case smaller absorbance change amplitudes would be expected for the CO and O₂ ligand binding reactions with the green protein compared to the red, and this difference was not observed (data not shown).

The mechanism for generation of sulfHbs involves a ferryl heme iron and sulfide (41). Given the conditions under which it is generated, if green HGb is a sulfHb it must either have a more stable ferryl oxidation state compared to other Hbs, or harbor a readily accessible sulfur atom. In support of the latter is a cysteine located in the distal heme pocket (position E9) that might facilitate sulfheme formation. With regard to the former, a ferryl heme iron is a component of peroxidase compound I and HGb has been attributed peroxidase activity (34). However, the level of activity is very low compared to known peroxidases (44). In fact, the level of peroxidase activity attributed to HGb is similar to the "pseudo"
peroxidase activity of other Hbs including myoglobin and soybean leghemoglobin (45,46). This suggests the primary role of HGb is not that of a peroxidase.

**Physiological significance** - The transport and facilitated diffusion of oxygen and other ligands by hemoglobins has been a subject of investigation since the 1960s (47,48). The relationship between Hb kinetic rate constants and the environment the protein functions within is fairly well established for these physiological roles (49-51). Although other hxHbs characterized thus far possess oxygen affinities that are too high for these functions, the oxygen affinity of HGb is of the same order as Mb and should allow it to serve in a similar role. It is therefore possible that HGb supports the facilitated diffusion of oxygen in those tissues that do not express Mb. In this scenario, hexacoordination would serve simply to decrease oxygen affinity, thereby allowing transport to a higher affinity oxidase.

Another hypothesis is that hxHbs (including HGb) are involved in a general mechanism for scavenging NO and/or other ROS in both plants and animals. Plant hxHbs and NGb are both up-regulated by hypoxia (10,19,20). Expression of plant hxHbs is also stimulated by conditions that activate the plant disease resistance pathway (52-54), which generates NO and other ROS (55,56). Reperfusion injury following ischemia in animals has been associated with NO (57); it is intriguing that both plants and animals express biochemically similar hxHbs in response to this type of challenge.

The expression of HGb has not yet been linked with hypoxia. Nevertheless, this possibility is interesting in the context of the oxygen-dependant modification of the heme co-factor. While this modification may be an artifact of recombinant expression, it has not been observed during the generation of other recombinant Hbs in this laboratory. It has been
previously proposed that hxHbs may play a role in sensing gaseous ligands (29,58). If HGb
serves this function, the heme modification may be a component of the sensory mechanism.

**Conclusions**- The study of HGb presented here identifies a new human gene that
encodes a member of a biophysically defined class of proteins called hexacoordinate
hemoglobins. These proteins are found in most organisms, and possess a regulatory ligand
binding mechanism that fundamentally differs from traditional pentacoordinate Hbs. In HGb
this mechanism results in exogenous ligand equilibrium affinity constants that are very
similar to those of Mb. And while considerable uncertainty remains as to the physiological
role(s) served by HGb or the other hxHbs, mounting evidence suggests a potential protective
function during conditions of oxidative stress.

**Acknowledgements**

We would like to thank Drs. Robert Thornburg and Clay Carter for providing their expertise
and assistance with the hybridization experiments, and Dr. Cyril Appleby for helpful
suggestions and a careful reading of this manuscript. This work was funded in part through
United States Department of Agriculture Grant No. 99-35306-7833.
References


21. Hargrove, M., Brucker, E., Stec, B., Sarath, G., Arredondo-Peter, R., Klucas, R.,
22. ALLGENES. (2001), version 2.2  The Computational Biology and Informatics
    Laboratory. Center for Bioinformatics, University of Pennsylvania


Reactive oxygen species in biological systems (Gilbert, D. L., and Colton, C. A.,
Eds.), Kluwer Academic, New York

Porphyrrins and metalloporphyrins (Smith, K. M., Ed.), Elsevier, Amsterdam

3377

41. Berzofsky, J. A., Peisach, J., and Horecker, B. L. (1972) J. Biol. Chem. 247, 3783-
3791

7372


44. Dunford, B. H. (1999) Model Peroxidases from Yeast and Horseradish, Cloned


Acta. 614, 303-312


Table 1. Tissue sources for each RNA dot in Figure 2.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>atrium (left)</td>
<td>frontal lobe</td>
<td>parotid</td>
<td>throat</td>
<td>kidney</td>
<td>thyroid</td>
</tr>
<tr>
<td>B</td>
<td>atrium (right)</td>
<td>temporal lobe</td>
<td>esophagus</td>
<td>trachea bronchial</td>
<td>bladder</td>
<td>pancreas</td>
</tr>
<tr>
<td>C</td>
<td>ventricle (left)</td>
<td>occipital lobe</td>
<td>stomach</td>
<td>lung (left)</td>
<td>prostate</td>
<td>adrenal</td>
</tr>
<tr>
<td>D</td>
<td>ventricle (right)</td>
<td>parietal lobe</td>
<td>small intestine</td>
<td>lung (right)</td>
<td>testis</td>
<td>tonsil</td>
</tr>
<tr>
<td>E</td>
<td>septum interventricle</td>
<td>thalamus</td>
<td>colon</td>
<td>diaphragm</td>
<td>uterus</td>
<td>thymus</td>
</tr>
<tr>
<td>F</td>
<td>pericardium</td>
<td>pons</td>
<td>rectum</td>
<td>skeletal muscle</td>
<td>breast</td>
<td>spleen</td>
</tr>
<tr>
<td>G</td>
<td>human DNA</td>
<td>cerebellum</td>
<td>liver</td>
<td>tongue</td>
<td>ovary</td>
<td>lymph node</td>
</tr>
<tr>
<td>H</td>
<td>plasmid DNA</td>
<td>spinal cord</td>
<td>gallbladder</td>
<td>adipose tissue</td>
<td>placenta</td>
<td>appendix</td>
</tr>
</tbody>
</table>
Table 2. Rate constants associated with ligand binding to the human hemoglobins.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$k'_{o2}$ $\mu$M$^{-1}$s$^{-1}$</th>
<th>$k_{O2}$ s$^{-1}$</th>
<th>$^{b}K_{O2}$ $\mu$M$^{-1}$</th>
<th>$k'_{CO}$ $\mu$M$^{-1}$s$^{-1}$</th>
<th>$k_{CO}$ s$^{-1}$</th>
<th>$^{b}K_{CO}$ $\mu$M$^{-1}$</th>
<th>$k_{H}$ s$^{-1}$</th>
<th>$k_{H}$ s$^{-1}$</th>
<th>$K_{H}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histoglobin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroglobin$^{c}$</td>
<td>130</td>
<td>0.3</td>
<td>200</td>
<td>38</td>
<td>0.007</td>
<td>2,500</td>
<td>8,200</td>
<td>9,800</td>
<td>1.2</td>
</tr>
<tr>
<td>Hb $\alpha$-chain R-state$^{d}$</td>
<td>19</td>
<td>15</td>
<td>2.2</td>
<td>0.009</td>
<td>245</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Hb $\beta$-chain R-state$^{d}$</td>
<td>74</td>
<td>47</td>
<td>5.9</td>
<td>0.01</td>
<td>590</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Myoglobin$^{e}$</td>
<td>17</td>
<td>15</td>
<td>0.51</td>
<td>0.02</td>
<td>25.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

$^{a}$Hexacoordinate hemoglobins are in bold.

$^{b}$These values were calculated using Equation 3.

$^{c}$(24); $^{d}$(59); $^{e}$(60)
Figure 1. Heme coordination states. A) A pentacoordinate heme. B) A hexacoordinate heme. Two His side chains coordinate the heme iron in hexacoordinate hemoglobins.
Figure 2. Sequence alignment of the major human Hbs and the Hgb gene structure. A) A primary sequence alignment of the five principal adult human hemoglobins. Residues sharing identity are highlighted in black, and those of high homology are shaded in gray. Sequence positions E7, F8, and CD1 (in reference to Mb) are annotated. B) A schematic of the gene structure of HGb.
Figure 3. Hybridization assays for tissue expression. Phosphoimages of hybridization results for A) histoglobin and B) neuroglobin. The tissues corresponding to each dot in A and B are listed in Table 1.
Figure 4. Red and green HGb, and the correlation of protein color with oxygen during expression. A) Overlay of the normalized visible spectra of green HGb (solid line), red HGb (dashed line) and reconstituted green apoprotein (dotted line). The green protein has a prominent absorbance peak centered on 630nm that is not present in the red protein. The reconstituted green apoprotein and red HGb have identical spectra, indicating the 630nm absorbance peak originates from the heme co-factor. All spectra are of the ferrous, CO bound protein forms. B) A plot of the ratio of the 630nm/543nm absorbance peaks vs. shake flask culture agitation speed. This demonstrates a correlation between the level of oxygen saturation during the culture growth and the proportion of green protein to red protein that is generated. C) A bar graph of the ratio described in B for protein grown in cultures aerated with either N₂ or 100% O₂ at identical agitation speeds. This demonstrates that it is oxygen concentration and not some other aspect of culture agitation speed that influences the generation of green HGb.
Figure 5. Visible absorbance spectra of ferrous, deoxy sperm whale Mb, human NGb, and HGb. This overlay of the normalized absorbance spectra from the ferrous, deoxygenated forms of Mb, NGb and HGb illustrates the hexacoordinate character of HGb. The pentacoordinate Mb has a single, broad absorbance peak in the visible region, whereas HGb has two peaks very similar to the hexacoordinate NGb. Split peaks in this spectral region in the absence of exogenous ligands is a signature of the hxHb class of proteins.
Figure 6. CO flash photolysis ligand binding kinetics. A) The residuals from single exponential (circles) and double exponential (crosses) fitted curves to the low [CO] time course (top residuals) and the high [CO] time course (bottom residuals) in B. The residuals are indistinguishable from one another. B) Time courses for CO rebinding to HGb following flash photolysis at 1000µM [CO] and 100µM [CO]. C) CO rebinding rate constants are plotted vs. CO concentration. A linear fit to these data yields a line with a slope of 5.6 µM⁻¹s⁻¹ and a y-intercept of 440 s⁻¹. The y-intercept value reflects a contribution to rebinding from the hexacoordinate component of the reaction.
Figure 7. O₂ flash photolysis ligand binding kinetics. A) The residuals from single exponential (circles) and double exponential (crosses) fitted curves to the low [O₂] time course (top residuals) and the high [O₂] time course (bottom residuals) in B. The residuals are indistinguishable from one another. B) Time courses for O₂ rebinding to HGb following flash photolysis at 1250μM [O₂] and 262μM [O₂]. C) O₂ rebinding rate constants are plotted vs. O₂ concentration. A linear fit to these data yields a line with a slope of 30 μM⁻¹ s⁻¹ and a y-intercept within error of zero. The y-intercept value indicates that O₂ rebinding is sufficiently fast to out-compete the hexacoordinate component of the reaction.
Figure 8. Rapid mixing ligand binding kinetics. A) The time courses for CO binding to the ferrous, deoxy form of HGb following rapid mixing are shown for several different concentrations of CO. The time scale for these reactions indicates they measure a component of the binding reaction other than just the bimolecular event. B) A plot of the fastest rate components extracted from three-exponential fits to the time courses shown in A vs. CO concentration. A curve simulating $k_{obs}$ generated using Equation 2 with the ligand binding and hexacoordination rate constants reported for HGb in Table 2 is overlaid on the observed values in B. This plot demonstrates that the reported hexacoordination rate constants are consistent with observed ligand binding data.
CHAPTER 5: CRYSTALLOGRAPHIC ANALYSIS OF SYNECHOCYSTIS CYANOGLOBIN REVEALS THE STRUCTURAL CHANGES ACCOMPANYING LIGAND BINDING IN A HEXACOORDINATE HEMOGLOBIN

A paper published in the *Journal of Molecular Biology*\(^1\)

James T. Trent III, Suman Kundu, Julie A. Hoy, and Mark S. Hargrove

**Abstract:** The crystal structures of cyanide and azide-bound forms of the truncated hemoglobin from *Synechocystis* are presented at ~1.8 Å resolution. These structures enable the first direct comparison of the exogenous and endogenous ligand-bound forms of a hexacoordinate hemoglobin. This structural analysis reveals a significant protein conformational shift without precedent in hemoglobins. FTIR spectroscopy experiments show a behavior in solution that is predicted by the crystal structure. Features of the structural transition between the two protein conformations are discussed as they relate to the complex ligand-binding behavior observed in hexacoordinate hemoglobins, the regulation of the globin fold, and the potential physiological function of this class of proteins.

\(^1\) *Journal of Molecular Biology*, (2004), 341(4): 1097-1108
Introduction

Numerous studies over the past several years have revealed hemoglobins (Hbs) in nearly all organisms, and uncovered examples of function, structure and kinetic behavior previously unsuspected in this super-family of proteins. 1; 2; 3; 4; 5 The expansion in both number and diversity of its members has given rise to a Hb nomenclature that divides this super-family into several subfamilies and classes.

One of these families comprises the “truncated” hemoglobins (trHbs), whose members include cyanobacteria, eubacteria and plants. The trHbs are typically 20-30 amino acid residues shorter than vertebrate Hbs, with an abbreviation of the classical “3-on-3” α-helical globin fold. 1 Among the distinguishing features of this “2-on-2” α-helical-sandwich fold are a hydrophobic cavity that links solvent to the ligand-binding site, and a hydrogen-bonding network connecting the bound ligand with a conserved Tyr side-chain in the distal heme pocket. 6; 7

The hexacoordinate Hbs (hxHbs) are another class, and differ from other Hbs in that the ligand-binding sites of their heme cofactors are coordinated by an intramolecular amino acid side-chain in the absence of exogenous ligands (Figure 1). This intramolecular coordination is reversible however, and hxHbs bind most ferrous and ferric Hb ligands with relatively high affinities. 8; 9 Examples of these proteins are found throughout the other Hb families, and representatives exist in nearly all plants and animals. 3; 4

Physiological functions have yet to be assigned with certainty to most members of the trHb family and hxHb class of proteins. However, the unusual biophysical traits that define these proteins may well be a requisite component of their functional role; one likely differing from the traditional oxygen transport service most associated with Hbs. 3; 10 A detailed
understanding of these biophysical phenomena are crucial to ascertaining what roles these proteins may be performing. Yet, bringing clarity to the difficult kinetic picture found in hexHbs, 8; 9; 11; 12; 13; 14; 15 or learning the impact of the hydrophobic cavities seen in trHbs, 6; 16; 17; 18 requires the ability to visualize the ligand-binding reactions at a molecular level.

Towards this goal, we present the x-ray crystallographic structures of liganded cyanoglobin from *Synechocystis* sp. PCC6803 (SynHb), a member of both the trHb family and hexacoordinate Hb class. 19; 20 The cyanomet (CN-SynHb) and azomet (AZ-SynHb) structures described herein, coupled with previous structures of SynHb 7; 21, allow the first direct comparison between the endogenous and exogenous ligand-bound states of a hexHb.

**Results**

Both the cyanomet and azomet forms of SynHb crystallize in tetragonal space group I4₁ with similar crystallographic constants (see Table 1). The CN-SynHb structure served as a molecular replacement model for the azide-bound structure, yielding an initial $R_{free}$ of ~30% that indicates the two molecules are quite similar. Following refinement, the AZ-SynHb and CN-SynHb models are fundamentally distinguishable only in their ligands, number of solvent molecules, and a few alternate side-chain rotamer conformations. These differences are such that the structures may be considered to represent two “equivalent” exogenously coordinated protein substates. Given such structural similarity, we chose to focus analysis on the CN-SynHb structure, to improve clarity and conserve space.

Two naming conventions are used in this work to identify a particular amino acid side-chain of interest. The most familiar convention designates an amino acid by its number in the peptide chain beginning with the N-terminus, i.e. Ser^3 refers to the third residue (a
serine) from the N-terminus. The other convention identifies a residue by its homologous position in the standard sperm whale myoglobin (swMb) structure. A particular amino acid is described by its corresponding helix and position within that helix, i.e. E7 refers to the residue occupying a position structurally homologous to the seventh residue in the E helix of swMb.

Because globins vary in primary sequence length, the latter naming convention is more facile when discussing residue functionality within the context of more than one globin. Structural differences between the classic globin fold and trHbs however, mean that occasionnally the former naming convention will be used to ensure the unambiguous identification of residues under discussion.

**Ligand-bound SynHb compared to other trHbs.** As anticipated from the molecular replacement solution, CN-SynHb displays an overall fold consistent with that of the cyanide-bound trHb from *C. eugametos* (CtrHb). The 2-on-2 helical architecture that defines the trHbs is exhibited by the structures of SynHb and CtrHb (Figures 2A and 2B). A distinguishing feature of the trHb fold is a continuous or nearly contiguous "tunnel" linking solvent and the heme pocket, which is observed in surface representations as typified in Figure 2B by CtrHb.\(^6;18\) Intriguingly, this tunnel is not observed in the crystal structure of the endogenously coordinated SynHb (Hx-SynHb).\(^21\) Yet, as is apparent in Figure 2C, a tunnel cavity is present in CN-SynHb. Although the overall shape of the tunnel is similar in these proteins, CN-SynHb shows an additional cavity in the proximal heme pocket that is not witnessed in CtrHb.

The trHb “ligand” tunnel has been suggested to serve as an efficient path between solvent and the binding site in the heme pocket, enabling the fast bimolecular ligand-binding
rate constants measured in these proteins. As this tunnel is only observed after a ligand is bound in SynHb however, the conformational change associated with its formation must occur before ligand binding to allow such a function. Alternatively, ligands may enter through other contacts with the heme pocket that exist in the Hx-SynHb structure, or perhaps the pathway for ligand entry is simply not obvious from this structure. Another possibility consistent with tunnel formation post-ligand binding in SynHb is that it actually serves to facilitate ligand escape from the heme pocket.

To explore possible routes by which ligands gain access to the heme pocket, solvent accessible surfaces were calculated for both the Hx-SynHb and CN-SynHb structures. The results shown in Figure 2D and E indicate that the Hx-SynHb distal heme pocket displays greater solvent accessibility from the propionyl face of the heme than that of CN-SynHb. These results are surprising given the collapse of the distal pocket necessary for endogenous hexacoordination, and suggest a direct route for ligand binding between solvent and the distal heme pocket.

The amino acid residues surrounding the heme cofactor in CN-SynHb are shown in Figure 3, and are similarly positioned as those of CtrHb. Unlike CtrHb however, the electron density for the heme pockets of CN-SynHb (Figure 3A) reveals a clear covalent linkage between His and the heme cofactor that is 2.0 Å in length. Moreover, despite the near equivalent spatial arrangement of residues illustrated in Figure 3B, CN-SynHb and CtrHb appear to offer different distal heme pocket interactions with the bound ligand. In CN-SynHb only the B10 residue makes a close contact with the cyanide, whereas both B10 and E7 side-chains are involved in close contacts in CtrHb. This may explain the marked difference in the geometry of the ligands with respect to the heme plane observed in the two
structures. The orbital structure of cyanide should result in a N-C-Fe angle of 180°, yet many
macromolecular structures have found deviations from this angle. The additional close
contact with the ligand in CtrHb may strain the coordinate-covalent bond with the heme, and
thus prevent cyanide from adopting the perpendicular position relative to the heme plane that
is observed in CN-SynHb.

*Global changes in structure induced by ligand binding to SynHb.* The side-chain of
His$^{46}$ in SynHb is the principal source of endogenous hexacoordination in the protein. Typically, the endogenous ligand in hxHbs occupies the E7 helix position, but His$^{46}$ aligns to
the position at E10 in SynHb. Structures of CN-SynHb and Hx-SynHb are superimposed in
Figure 4A, and this alignment is guided by C$\alpha$ atoms rather than the heme position. This
facilitates a direct comparison of the globin tertiary fold changes between the two structures,
without introducing the artifactual residue displacements observed when there are even
minor differences between heme plane orientations. The color gradient of the CN-SynHb
ribbon structure ranges from blue ($< 1.5$ Å) to red ($> 3.0$ Å) in correlation with its rms
deviation from Hx-SynHb.

This overlay reveals that tertiary structural differences between endogenous and
exogenous ligand-bound SynHb are principally composed of two regions in the protein.
When cyanide bound, the A and upper B helices in SynHb are shifted “inward” towards the
heme as compared with the Hx-SynHb structure, and the A helix in particular is in closer
proximity to the E helix. The second region consists of the E helix and EF or “pre-F” loop,
which swings out and away from the heme in CN-SynHb relative to the endogenously
coordinate protein. The remaining domains of the globin show much less variance between
the ligand states, and a previously described covalent linkage of the heme and His$^{117}$ is
present in both structures.\textsuperscript{23, 24} However, the cyanide-bound heme cofactor is tilted several degrees from the heme plane orientation found in Hx-SynHb, and further recessed within a distal pocket enlarged by the E helix shift.

The magnitude of the transition between endogenous and exogenous ligand-bound conformations in SynHb was evaluated in the context of other Hbs for which liganded and non-liganded structures have been solved. The rms deviation between bound and unbound ligand states for several relevant Hbs were calculated using the same C\textalpha{} guided backbone alignment used with SynHb. These rms deviation values are presented in Figure 4B-G, with the SynHb deviations plotted in black and overlaid by the comparative Hb values in red. The vertebrate hemoglobin swMb shows little structural change upon ligand binding relative to SynHb in Figure 4B.\textsuperscript{22, 25} A comparison with the bacterial hemoglobin from \textit{Vitreoscilla} in Figure 4C indicates that exogenous ligand binding in SynHb involves significantly greater structural rearrangement.\textsuperscript{26}

The Hbs compared with SynHb in Figure 4D-G are those proteins for which both ligand-bound and hemichrome structures currently have been solved, although it must be noted that none of these globins are considered “true” members of the hxHb class. True hxHbs are those whose endogenous-ligand heme coordination does not arise because of mutagenesis or protein denaturation processes. In Figure 4D the rms deviation between liganded swMb and a double mutant of swMb exhibiting bis-histidyl coordination\textsuperscript{27} is hardly distinguishable from the comparison of Figure 4A. The proteins used for comparison to SynHb in Figures 4E & 4F are tetrameric globins, with quaternary structures made up of hemichrome and liganded monomer subunits.\textsuperscript{28, 29} The rms deviation calculations involved only the hemichrome monomer subunit (and its ligand-bound form). The \( \beta \) globin from the
Antarctic fish *Trematomus newnesi* (HbTnβ) displays structural differences between the hemichrome and exogenous ligand-bound states that are of a similar order as that seen in the previous comparisons (Figure 4E). However, the horse hemoglobin α-chain (eHb) exhibits a more pronounced difference in hemichrome and exogenous ligand-bound conformations (Figure 4F). The transition from aquomet to hemichrome state occurs upon lowering the pH to 5.4, and the two rms deviation spikes in Figure 4F involve the histidines at positions CD3 and D1 changing from β-turn to helix. 29

Only the comparison in Figure 4G shows a hemichrome protein/liganded protein rms deviation that is as significant as is seen in SynHb. The sea cucumber *Caudina arenicola* (scHb) comparison is included because it was the first structure of a hemichrome having a low-spin heme iron coordinated by both proximal and distal histidines. A ligand-bound structure of the Hb-C chain has yet to be solved however, so the rms deviations plotted in Figure 4G were calculated using the cyanomet Hb-D chain. 30 Although the sea cucumber Hb-C and Hb-D chains are highly homologous, sharing 65% identical amino acid residues, they do not provide an “equivalent” basis for comparison of the hemichrome and exogenous ligand-bound states, as do the other Hb examples.

*The different conformations in SynHb center on a “hinge” in the EF-loop.* The displacement of His46 from a coordinate position at the heme iron, to a position outside the heme pocket involves more than just a change in rotamer conformations. The entire E helix swings out from the heme pocket in CN-SynHb relative to the endogenous ligand-bound structure (Figure 4A). The residues comprising the “hinge” upon which the helix swivels are
explored in Figure 5A. Those residues from the cyanide bound structure are in red, whereas those of Hx-SynHb are black.

The EF-loop hinge is composed of two principal structural differences between the endogenous and exogenous ligand-bound conformations. These differences are in the heme-D-propionate orientation, and in the presence of an additional three-residue turn in the F helix of the Hx-SynHb structure. Both of these structural phenomena are interdependent however, and a series of hydrogen-bond network substates intermediate between the two extremes illustrated are expected to form and break during the transitioning swivel of the E helix.

In the endogenous ligand-bound conformation, the heme-D-propionate is oriented orthogonally to the heme plane on the distal side by hydrogen-bonding interactions with Lys$^{42}$ and Tyr$^{65}$. This propionate interacts electrostatically with Arg$^{64}$ and Tyr$^{61}$ when an exogenous ligand occupies the heme binding site. The distal histidine (His$^{46}$) also interacts with Tyr$^{61}$ when displaced from the heme pocket. The orientation of the heme-D-propionate is parallel with the heme plane, and proximally located in the exogenous ligand-bound conformation.

When His$^{46}$ coordinates the heme iron, the F helix harbors a three-residue helical extension that is stabilized by the hydrogen-bonding pair Arg$^{64}$ and Asp$^{62}$. Additional stability is garnered from the interaction of Tyr$^{65}$ with the heme-D-propionate. Either because of, or before exogenous ligand binding, the bond pair between Arg$^{64}$ and Asp$^{62}$ is broken and Arg$^{64}$ subsequently interacts with the heme-D-propionate.

**Structural changes in the heme pocket.** Figure 5B shows the effects of structural rearrangement in the heme pocket induced by ligand binding. The helical motions described
in Figure 4 result in a dramatic repositioning of the amino acid side-chains that interact with the bound ligand during transition from Hx-SynHb to the CN-bound conformation. The most obvious difference between the endogenous and exogenous ligand-bound distal heme pocket structures is what coordinates the heme iron. The ε nitrogen of His$^{46}$ binds the heme iron in Hx-SynHb, whereas this amino acid is displaced from the heme pocket towards the solvent when the protein occupies an exogenous ligand-bound state. The glutamines at the E7 and E11 positions do not interact either with each other or with the B10 Tyr$^{22}$ when His$^{46}$ coordinates the heme iron. These three residues are observed to have shifted positions towards the heme when an exogenous ligand occupies the binding site, allowing formation of a hydrogen-bonding network.

**FTIR spectroscopic investigation of heme pocket interactions with the bound ligand.** Amino acids occupying the E7 position commonly play a role in dictating the affinity of the ligand bound at the heme iron.$^{31; 32; 33}$ In most hxHbs this residue is a histidine and has the potential to influence ligand affinities through hydrogen bonding. And in hxHbs characterized to date this histidine has also been also the endogenous hexacoordinating ligand.$^{11; 15; 20}$ The hexacoordinating side-chain in SynHb is the histidine at the E10 position however, and in the crystallographic structure of the exogenous ligand-bound conformation, His$^{46}$ is clearly displaced from the heme pocket to the solvent. This implies that the distal histidine in SynHb does not function in the stabilization of a bound exogenous ligand.

FTIR spectroscopy of CO-bound Hbs is a technique capable of assessing the solution-state electrostatic environment around the heme ligand,$^{34}$ and thus provides a means for corroborating the crystallographic and the solution distal pocket structures. The C-O stretching frequencies of CO-bound Hbs are dictated by electrostatic potentials exerted by
amino acid residues in close proximity to the bound ligand. A positive potential (mostly in the form of a hydrogen bond) will lower the C-O stretching frequencies, whereas a negative potential generates the opposite effect. The CO-bound H46L mutant protein was examined along with wild-type CO-SynHb using this technique, and the results are presented in Figure 6.

The two spectra in Figure 6A are controls using soybean leghemoglobin (wtLba) and its H61L mutant protein. The overlay of CO-FTIR spectra in Figure 6A clearly illustrates the electrostatic interaction between His\(^{61}\) and the bound ligand. As the substitution of a non-polar amino acid for His\(^{61}\) results in a higher stretching frequency, this indicates not only that His\(^{61}\) interacts with bound CO in wtLba, but also tells us this interaction involves a positive potential. The CO-FTIR spectra for the wild-type SynHb and its H46L mutant protein are overlaid in Figure 6B. The C-O stretching frequencies for wild-type and H46L SynHb are the same, indicating that the E10 distal histidine (His\(^{46}\)) does not electrostatically interact with the bound CO ligand. These results support the distal pocket structure assigned in ligand-bound SynHb, which shows that His\(^{46}\) moves away from the bound ligand and is absent from the heme pocket.

**The different conformations in SynHb center on a “hinge” in the EF-loop.** The displacement of His\(^{46}\) from a coordinate position at the heme iron, to a position outside the heme pocket involves more than just a change in rotamer conformations. The entire E helix swings out from the heme pocket in CN-SynHb relative to the endogenous ligand-bound structure (Figure 4A). The residues comprising the “hinge” upon which the helix swivels are
explored in Figure 5A. Those residues from the cyanide bound structure are in red, whereas those of Hx-SynHb are black.

The EF-loop hinge is composed of two principal structural differences between the endogenous and exogenous ligand-bound conformations. These differences are in the heme-D-propionate orientation, and in the presence of an additional three-residue turn in the F helix of the Hx-SynHb structure. Both of these structural phenomena are interdependent however, and a series of hydrogen-bond network substates intermediate between the two extremes illustrated are expected to form and break during the transitioning swivel of the E helix.

In the endogenous ligand-bound conformation, the heme-D-propionate is oriented orthogonally to the heme plane on the distal side by hydrogen-bonding interactions with Lys$^{42}$ and Tyr$^{65}$. This propionate interacts electrostatically with Arg$^{64}$ and Tyr$^{61}$ when an exogenous ligand occupies the heme binding site. The distal histidine (His$^{46}$) also interacts with Tyr$^{61}$ when displaced from the heme pocket. The orientation of the heme-D-propionate is parallel with the heme plane, and proximally located in the exogenous ligand-bound conformation.

When His$^{46}$ coordinates the heme iron, the F helix harbors a three-residue helical extension that is stabilized by the hydrogen-bonding pair Arg$^{64}$ and Asp$^{62}$. Additional stability is garnered from the interaction of Tyr$^{65}$ with the heme-D-propionate. Either because of, or before exogenous ligand binding, the bond pair between Arg$^{64}$ and Asp$^{62}$ is broken and Arg$^{64}$ subsequently interacts with the heme-D-propionate.
Discussion

The cyanide and azide-bound structures of SynHb presented in this work provide the first instance of a hxHb in which both exogenous and endogenous ligand-bound states have been solved. This allows investigations of ligand binding to be placed in a structural context, with the goal of understanding factors that regulate the interplay between intramolecular hexacoordination and exogenous-ligand stabilization. The relevance to these processes of variances between the two SynHb conformations is explored in the following discussion.

*Intramolecular hexacoordination results from control of globin flexibility.* Among all the protein structures currently available, no Hb demonstrates a tertiary conformational change upon reversible ligand-binding that is of the magnitude observed in the hxHb from *Synechocystis*. This implies a greater flexibility in the globin fold of SynHb relative to other known Hb structures. As SynHb is both classified as a hxHb and also a member of the trHb family, this flexible globin fold may be common to one or both protein categories. This flexibility must be controlled however, else the globin roles in shielding the heme iron from auto-oxidation, and in modulation of ligand specificity and affinity would not be possible.

Evidence of the controlled nature of the globin flexibility in SynHb is found in the structural overlay of Figure 4A. Areas of highest rms deviation between the structures are confined to two specific regions of the protein. The rest of the protein shows little deviation between different ligand states, even in loop and termini regions that might be expected to display considerable variation. This low rms deviation is particularly significant because the two structures are in different space groups, so crystal contacts cannot be implicated in constraining loops or termini. This argues against a globin fold that is highly flexible in
general, instead favoring a globin fold capable of a controlled transition between two significantly divergent conformations.

Composition of the globin primary structure is clearly a component regulating the rigidity of the protein fold. Control of globin flexibility must involve the heme cofactor however, as the different ligand states giving rise to the observed structural transition first manifest at the heme iron. The binding state of the heme can be communicated through interactions between the heme and globin, interactions of the ligand and globin, or a combination of both.

Interactions between the globin and heme cofactor in SynHb involve hydrogen bonds with the heme-D-propionate, coordinate-covalent bonds with the heme iron, and a covalent bond with the heme-B-vinyl group. Possible functions of the covalent links—His and heme-B-vinyl, proximal histidine (His) and heme iron—in regulating the structural transition include anchoring a region of the globin relative to the heme. A “loose” and flexible heme does not allow the large, specifically localized conformational shift possible with a fixed reference between heme and globin. A flexible heme may simply reorient itself within the heme pocket, negating the impact of a highly localized and specific conformation shift of the protein matrix. Additionally, having two fixed points of reference between the globin and heme (as constituted by the two covalent bonds) further constrains heme orientation relative to conformational changes by preventing extensive rotation about either protein-heme bond.

The hydrogen-bonding networks associated with the heme-D-propionate in Figure 5A are likely candidates for controlling and facilitating the E helix structural movement associated with ligand binding. The additional turn in the F-helix that is present in Hx-
SynHb may provide the rigidity to “lock” the E helix in that conformation. This turn is stabilized during endogenous coordination by hydrogen-bonding between the Tyr$^{65}$ side-chain and the heme-D-propionate. This propionate hydrogen bonds with different amino acid residues (Arg$^{64}$ and Tyr$^{61}$) in the exogenous ligand-bound structure, with Tyr$^{65}$ and the helix turn it was associated with now relaxed in a more flexible loop. It is not clear if ligand binding induces this change, or if the change enables ligand binding. Regardless which comes first, the interactions between the heme-D-propionate and globin in SynHb appear to function as a “hinge” that correlates the position of the E helix with the binding state of the heme cofactor.

**The ligand tunnel in SynHb.** A tunnel is not present in the crystal structure of Hx-SynHb$^{36}$, but forms upon binding of exogenous ligands. The role and importance of this tunnel has been considered extensively in other trHbs as it could serve as a route for ligand entry or exit from the heme pocket.$^{6, 18}$ In evaluating the role of this tunnel in SynHb, we again face the issue of which comes first: ligand entry or conformational change. As exogenous ligands cannot bind to a heme iron already occupied by a hexacoordinating distal histidine, it seems logical that conformational change opening the tunnel should precede ligand entry. In this scenario, the tunnel could serve as both entry and exit routes for the ligand. However, observation of this tunnel is very much dependent on the methods used for calculation of solvent accessibility. Thus, the significance of this observation is uncertain until further empirical evidence is acquired in support of it.

**A structural interpretation of heterogeneous ligand binding in SynHb.** The kinetic behavior associated with ligand binding in SynHb is complex$^{20, 37}$, which is true for the other hxHbs as well.$^{8, 9, 38, 39, 40}$ The origins of this complexity are not understood, and are a source
of controversy and debate within the field. The amino acid side-chain(s) serving as an endogenous ligand is clearly a component of the observed complexity, but is not the only source of the phenomenon underlying the ligand-binding behavior of hxHbs.

The transition between endogenous and exogenous ligand-bound conformations in SynHb involves more than the association and dissociation of His\textsuperscript{46} from the heme iron. If structural transitions involving a similar underlying mechanism are common to the hxHb protein class, then they may be the origin of the complex behavior observed during ligand binding in these proteins. The movement of the E helix requires the breaking and formation of the hydrogen bonds associated with the previously discussed “hinge.” The energetics and stability of the hinge as it transitions from one conformation to the other may give rise to a protein population with transient fractions harboring different ligand-binding characteristics. Movement of the B helix and amino acids of the distal heme pocket into their exogenous ligand conformation, coupled with the actual electrostatic interaction between distal pocket amino acid side-chains and the ligand, are other potential contributors to the observed complexity in SynHb (and possibly the other hxHbs) ligand-binding kinetics.

**Implications for physiological function.** Physiological function has yet to be assigned for any hxHb with any degree of certainty, and thus the contribution of the SynHb transition between exogenous and endogenous-ligand conformations to its biologic role can only be speculated. However, the nature of this structural behavior suggests SynHb (and possibly the other hxHbs as well) may have a function other than the gas transport role traditionally associated with hemoglobins.

The comparative analysis presented in Figure 4B-G shows that exogenous ligand-binding in SynHb affects the globin structure far more than is required by other Hbs to
accommodate the bound ligand. For this reason, SynHb would be inefficient relative to other Hbs in a role involving just gas transport, and thus should be selected against. Considering that a much greater range of affinities has been observed in hemoglobins with a pentacoordinate scaffold, we must ask, “Did reversible intramolecular hexacoordination really evolve only to modulate ligand affinity?” Both the specificity of the regions in the globin that alter between ligand conformations, and the underlying structural mechanisms regulating this transition argue for a functional relevance. The magnitude of the structural difference in SynHb following exogenous ligand-binding may be associated with a signaling mechanism or part of some other change in a protein-protein interaction, as has been suggested for the human hXHb neuroglobin.\textsuperscript{41, 42} The conformational “switches” that comprise the structural difference between active and inactive G\textsubscript{\alpha} protein are of a similar scale as that reported for the SynHb EF "hinge" here.\textsuperscript{43} Alternatively, the hydrogen-bonding networks and the movement of the distal histidine from solvent to the heme pocket may facilitate electron and/or proton transport. An enzymatic role for SynHb performing oxidation/reduction chemistry is certainly possible.

**Materials and Methods**

*Protein production*- The methods used for the expression and purification of SynHb (and its mutants) have been previously described in detail.\textsuperscript{9, 20} In brief, a pET-28 (Novagen) vector containing the SynHb cDNA was over-expressed in *E. coli* BL21(DE3) cells. Following lysis, a purification protocol that included ammonium sulphate fractionation, hydrophobic (phenyl sepharose), anion exchange (DEAE sepharose) and size exclusion (Sephacryl S-200) chromatography resulted in protein whose purity was spectrally measured
(ratio of absorbencies at the Soret peak and 280nm, or “r/z”) as > 5.0. This protein was stored at -80°C in 10mM potassium phosphate buffer, pH 7.0 until required for experiments.

**FTIR spectroscopy**- Samples of each CO-bound protein were prepared by reducing a 2-3mM, 20 μl aliquot of the ferric hemoglobin with sodium dithionite under an atmosphere of 100% carbon monoxide. A gas-tight syringe (purged with N\textsubscript{2}) was then used to add each sample to a BioCell IR cuvette (BioTools, Inc.) as a uniform film, free of bubbles. This cuvette was then placed in a Nexus 470 FTIR spectrometer (Nicolet Instrument Corp., Wisconsin) and allowed to equilibrate (under N\textsubscript{2}) for five minutes. Spectra were recorded with one cm\textsuperscript{-1} resolution over the region 1800-2100 cm\textsuperscript{-1} using Nicolet's OMNIC software for data collection and processing. The final CO-FTIR spectra were corrected for buffer and protein background by subtracting the corresponding deoxy-Hb spectra.

**Crystallization and structure solution**- Before use in a crystal plating experiment, each aliquot of protein was oxidized by addition of a slight molar excess of potassium ferricyanide. This potassium ferricyanide was removed by passage of the sample through a G-25 size exclusion column equilibrated with a buffer of 10mM potassium phosphate, pH 7.0. Recombinant SynHb was crystallized by vapor diffusion techniques, using protein at a concentration of 30mg/mL. For the azide-bound SynHb crystal, a 2 μL protein drop was equilibrated against an equal volume of reservoir buffer consisting of 1.5M sodium citrate (pH 6.0), 20% sucrose, and 50mM sodium azide. The cyanide-bound SynHb crystallization conditions were 1.5M sodium citrate (pH 6.0), 0.2 M ammonium citrate, 20% sucrose, 3% dioxane and 25mM sodium cyanide. Protein crystals grew over the course of several weeks at 4°C, and were not stable at room temperature.
Diffraction data were collected at 100K on a Rigaku/MSC home source generator, and processed using d*TREK. CNS was used to solve the cyanide-bound (CN-SynHb) structure by molecular replacement with the trHb from *Chlamydomonas eugametos* (Protein Data Bank (PDB) code 1DLY). A crystallographic R-value of 0.47 was obtained for the initial molecular replacement solution. Refinement exploited both CNS and CCP4 program suites, and manual model rebuilding was done using the software O and Deep View Swiss PDB viewer. The azide-bound (AZ-SynHb) structure was both solved and refined with CCP4, using the CN-SynHb structure for molecular replacement. The final models have been deposited with the Protein Data Bank, and assigned PDB codes 1S6A (AZ-SynHb) and 1S69 (CN-SynHb). Both models contain 123 amino acids, with 148 and 142 waters in the cyanide-bound and azide-bound structures respectively. The crystal cell constants, data processing and refinement statistics are reported in Table 1.

**Figure creation** Structure alignments and rms (root mean square) deviation calculations were done using LSQKAB(supergompose) from the CCP4 suite of programs. Graphs were produced with Igor Pro (Wavemetrics), and structural illustrations were generated using Deep View Swiss PDB viewer. Solvent accessible surfaces were calculated with a 1.4 Å probe using Swiss PDB viewer and SURFNET. Residues are considered "close contacts" when they are within 3.0 Å of a potential hydrogen-bonding partner.

**Acknowledgments**

This work was made possible by the National Science Foundation (Award No. MCB-0077890), the USDA (Award No. 99-35306-7833), and the Iowa State University Plant Sciences Institute. We thank George Blouin and John S. Olson for assistance with the FTIR experiments.
References


<table>
<thead>
<tr>
<th>Crystal data</th>
<th>SynHb Cyanide</th>
<th>SynHb Azide</th>
</tr>
</thead>
<tbody>
<tr>
<td>space group</td>
<td>I4$_{1}$</td>
<td>I4$_{1}$</td>
</tr>
<tr>
<td>a=b, Å</td>
<td>84.06</td>
<td>83.35</td>
</tr>
<tr>
<td>c, Å</td>
<td>64.62</td>
<td>65.18</td>
</tr>
<tr>
<td>Data processing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>resolution, Å</td>
<td>59.8-1.81</td>
<td>58.72-1.82</td>
</tr>
<tr>
<td>total reflections</td>
<td>182487</td>
<td>58830</td>
</tr>
<tr>
<td>unique reflections</td>
<td>33974</td>
<td>33105</td>
</tr>
<tr>
<td>completeness, %</td>
<td>97.9(83.8)</td>
<td>96.8(78.2)</td>
</tr>
<tr>
<td>$R_{merge}$, %</td>
<td>0.055(0.274)</td>
<td>0.048(0.226)</td>
</tr>
<tr>
<td>average I/σI</td>
<td>15.3(3.0)</td>
<td>9.9(3.0)</td>
</tr>
<tr>
<td>mosaicity</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_{crystal}/R_{free}$, %</td>
<td>17.2/19.9</td>
<td>19.1/21.5</td>
</tr>
<tr>
<td>rms bond lengths, Å</td>
<td>0.023</td>
<td>0.021</td>
</tr>
<tr>
<td>rms bond angles, °</td>
<td>1.81</td>
<td>1.69</td>
</tr>
<tr>
<td>avg. temp. factor, Å$^2$</td>
<td>28.2</td>
<td>23.8</td>
</tr>
<tr>
<td>water molecules</td>
<td>148</td>
<td>142</td>
</tr>
</tbody>
</table>

*aThe numbers in parenthesis are the outer shell statistics, 1.89-1.81 Å and 1.90-1.82 Å for CN-SynHb and AZ-SynHb respectively.
Figure 1. Distinguishing "pentacoordinate" and "hexacoordinate" Hbs. A) In the absence of an exogenous ligand, a pentacoordinate Hb will have an open binding site at the heme iron as is seen with the trHb from *Vitreoscilla* (PDB code 1VHB). B) This binding site is occupied by an endogenous ligand in hexacoordinate Hbs under the same conditions as is illustrated with the trHb from *Synechocystis* (PDB code 1RTX).
Figure 2. The trHb fold and "ligand" tunnel in *Synechocystis* cyanoglobin. Each stereoview of the structural models below has the peptide backbone regions labeled according to the corresponding swMb helices. **A)** The CN-SynHb structure clearly exhibits the 2-on-2 helical fold common to the trHbs. **B)** The solvent tunnel in cyanide-bound CtrHb. Note the nearly continuous path in the molecular surface linking solvent and heme pocket. **C)** A solvent tunnel can be observed in the CN-SynHb model. This tunnel differs from that of CtrHb however, in that it does not directly connect the distal pocket to the solvent. In CN-SynHb there also appears to be an additional cavity present on the proximal side of the heme that is lacking in CtrHb. **D)** The Hx-SynHb structure with a cavity surface showing the solvent-accessible volume near the distal heme binding site. Although the solvent tunnel observed in other trHbs is not witnessed in Hx-SynHb, the heme iron is still accessible. **E)** The solvent accessibility of the heme in CN-SynHb is indicated by the cavity molecular surface. In spite of manifesting a solvent tunnel akin to other trHbs, CN-SynHb does not allow ligand entry from the solvent on the propionyl-side of the heme.
Figure 3. The heme pocket environment of CN-SynHb. A) An overlay of the CN-SynHb model heme pocket and its electron density. The covalent bond between the heme cofactor and His$^{117}$ is evident as electron density linking them. B) An overlay of cyanide-bound CtrHb (PDB code 1DLY, in black) and CN-SynHb (in red) illustrating residues involved in close contacts with the ligand. The only close contact formed with the ligand of CN-SynHb involves the Tyr$^{22}$ at the B10 position, whereas in CtrHb close contacts with the ligand are found between both the Gln at the E7 position, and the B10 Tyr. This results in a ligand geometry for cyanide in CtrHb that is not orthogonal to the heme plane, contrary to what is observed in CN-SynHb and would be predicted based upon the linear nature of the ligand molecule.
**Figure 4. The conformational changes associated with ligand binding in SynHb and other Hbs.** A) An overlay of the Hx-SynHb (gray) and CN-SynHb (colored) crystal structures. The cyanide-bound structure is colored by rms deviations between the Cα atoms of the two structures, with blue at low deviation values and transitioning to red at high values. In B) - G) the black bars represent the rms deviations (Å) between residues of CN-SynHb and Hx-SynHb that were structurally aligned by superposition of the Cα atoms. The red bars show the rms deviations (Å) between a ligand-bound and ferric structure for each of several comparison Hbs, generated using the same Cα alignment method. B) Cyanomet sperm whale Mb (PDB code 1EBC) vs. ferric swMb (PDB code 1JP6). C) The trHb from *Vitreoscilla* in ferric (PDB code 1VHB) and azomet (PDB code 2VHB) forms. D) Cyanomet swMb vs. the bis-histidyl hexacoordinate Mb mutant H64L/V68H (PDB code 1MNI). E) The carbonmonoxy and hemichrome β subunit from the Antarctic fish *Trematomus newnesi* (PDB code 1T1N) superimposed. F) The aquomet (PDB code 1NS6) and hemichrome (PDB code 1NS9) α subunit from equine Hb. G) Cyanomet Hb-D (PDB code 1HLM) vs. hemichrome Hb-C (PDB code 1HLB) from the sea cucumber *C. arenicola*. It should be noted that the superposition of the sea cucumber Hbs involves two isomers of differing primary sequence. Aside from the comparison shown in G), it is clear that the structural transition associated with ligand binding in SynHb involves significantly greater protein rearrangement than evidenced in other Hbs.
Figure 5. Key residue rearrangements during ligand binding in SynHb. A) A stereo view overlay showing the residues involved in hydrogen bonds with the heme propionyl group. In Hx-SynHb (black) the propionates are oriented “vertically,” and towards the distal side, whereas the CN-SynHb (red) heme has “horizontally” oriented propionates localized on the proximal side. A different pair of residues in the two protein forms stabilizes these alternate propionate conformations. A three-residue helical turn stabilized by hydrogen-bonding between Asp$_{62}$ and Arg$_{64}$ in Hx-SynHb is not present in the CN-SynHb structure. Instead, a loop is formed where Arg$_{64}$ participates in a hydrogen-bonding network with the heme propionate, which also involves Tyr$_{61}$ and His$_{46}$. In Hx-SynHb, this same heme propionyl group electrostatically interacts with the side-chains of amino acids Lys$_{42}$ and Tyr$_{65}$. B) A stereo view overlay of the distal heme environment, with residues of Hx-SynHb illustrated in black and those of CN-SynHb shown in red. The heme cofactors were used to generate this structural alignment to emphasize the difference between the ligand-binding environments found in the two SynHb states. In the transition from endogenous to cyanide-bound SynHb, His$_{46}$ dissociates from the heme iron and faces away from the heme pocket towards the solvent. The residues Tyr$_{22}$, Gln$_{43}$ and Gln$_{47}$ all move inward relative to the bound ligand. The proximal pocket however, does not undergo a similarly marked rearrangement.
Figure 6. CO-FTIR measurement of the ligand electrostatic environment. A) The IR spectrum of CO-bound wild-type leghemoglobin (solid line) shows a C-O stretch peak that is centered around 1948 cm\(^{-1}\). The C-O peak of a mutant of Lba in which the distal histidine is removed (dashed line) is clearly shifted to a higher wavenumber (1968 cm\(^{-1}\)), indicating loss of positive electrostatic potential. B) The IR spectra for CO-bound wild-type SynHb (solid line) and its H46L mutant protein (dashed line) do not display such pronounced peak shift. This suggests that in ligand-bound SynHb there is little interaction between the distal histidine and the bound ligand, which is consistent with the crystallographic structure of CN-SynHb.
CHAPTER 6: GENERAL CONCLUSIONS

Hexacoordinate hemoglobins are a highly diverse group of proteins found near ubiquitously amongst living organisms. They consist of a recently described branch of the well-studied hemoglobin class of proteins. Yet, despite the wealth of knowledge regarding hemoglobins in general, hxHbs require new methods to assess and describe their novel ligand binding behavior. This conclusion appears inescapable following the work presented in Chapter 2. There we observed the significant differences in ligand binding not only between pentacoordinate and hexacoordinate proteins, but between individual hexacoordinate hemoglobins as well. That this behavior is not a sole artifact of the coordinating distal histidine was also shown, from which we might conclude that “hexacoordination” represents an evolved function for ligand binding regulation differing from the well-known pentacoordinate hemoglobins.

That hxHbs are near ubiquitous in nature was a conclusion to be drawn only following the work of Chapters 3 and 4. Although established amongst some bacteria and thought to be present in all plants, it was the identification of human neuroglobin as a hxHb (Chapter 3) that expanded this family of proteins to vertebrates and man. With the discovery of histoglobin (Chapter 4) it also became apparent that expression of multiple, differing hxHbs within the same organism was common to both plants and animals. A not unfounded conclusion to be drawn from their near universal presence in living things is that hxHbs must have some fundamentally important role in life. Additional support for this conjecture is found in the incredible sequence homology (over 90% identity) across diverse species.
displayed by both neuroglobin and histoglobin. It suggests proteins exquisitely evolved to their function, with random mutation of their genes not tolerated by the host organisms.

The study presented in Chapter 5 continues to strengthen the arguments for a well-tuned, highly refined functionality to have evolved in the hxHb family. The comparative analysis allowed for the first time by the liganded hxHb structure shown in this work, dramatically reveals a rigid and highly specific structural difference between the bound/unbound states of the protein. This structural change upon ligand binding is further demonstrated to be unprecedented among the entire globin family. It might be that the previously discussed high homology among these proteins is necessary solely to retain the precise nature of this helical swing. From this we might conclude that this structural shift is an element of, or a requirement for its physiological function.

The details of the studies described in the previous chapters allow us to draw some fascinating conclusions of a broad and general nature, however, like most (if not all) work in science it also serves to highlight what future experiments might prove interesting and useful. For instance, although the descriptive mechanism (or model) presented in Chapter 2 clearly serves better than its predecessors, it still fails to account for a considerable fraction of total ligand binding. Optimally we would like to develop a model capable of accurately describing all binding behavior for these proteins. Key to both development and usefulness of this model is the ability to directly tie kinetic descriptors in the mechanism to actual structural features of the molecule. Use of natural variant/site-directed mutant proteins to map structural and kinetic behavior to one another will prove invaluable in discerning “how” the hxHbs perform their function. Additional work examining possible post-translational modifications to these proteins, and their impact on both structural and kinetic behavior need
to be done. Neuroglobin possesses a putative palmityylation site and histoglobin displayed its peculiar “green” response to oxygen levels, with both potentially undergoing various in vivo side chain modifications that may have relevance to their functionality.

The physiological role(s) or function(s) of hxHbs is obviously of great interest, and though the studies above help to define what those roles may be, they do not establish what they actually are. Considerable new insights may be drawn from the kinetic data alone when it is possible to view them from a vantage point of known physiological function. Future work that might prove favorable in assigning functional roles to hxHbs includes identification of protein binding partners (possibly relating hxHbs to a known metabolic pathway) and physiological studies involving hxHb-knock out organisms.
ACKNOWLEDGEMENTS

None of this dissertation would have been possible without the assistance and guidance of Mark Hargrove, a man clearly capable of considerable tolerance and leeway with respect to differing behavioral aspects of those he chooses to work with. Suman Kundu was an exceptional lab mate, whose assistance and innumerable conversations enriched my graduate experience. Trish Stauble proved to be a stalwart ally whose administratorial expertise is responsible in no small part for my graduation. My committee members, Amy Andreotti, Gloria Culver, Reuben Peters and Jacob Petrich also deserve mention for exhibiting some of the behavioral tolerance ascribed above to Mark Hargrove. And last, but not least, I'd like to dedicate this humble dissertation to my mother (June Sharon Trent) and family. Everyone else can go fuck themselves. (Just kidding.)