2007

The role of Tyr61 in synechocystis hemoglobin in regulating E-helix movement

Anand Venugopal

Iowa State University

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

Part of the Biophysics Commons

Recommended Citation

https://lib.dr.iastate.edu/rtd/15089

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The role of Tyr61 in synechocystis hemoglobin in regulating E-helix movement

By

Anand Venugopal

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

MASTER OF SCIENCE

Major: Biophysics

Program of Study Committee:
Mark Hargrove, Major Professor
Michael Shogren-Knaak
Hans Stauffer

Iowa State University
Ames, Iowa
2007

Copyright © Anand Venugopal, 2007. All rights reserved.
# Table of Contents

List of Figures                                           iv  
Abstract                                               v  
Chapter 1: Introduction                                 1  
  General Background                                    1  
  Hexacoordination                                      2  
  Truncated Hemoglobins                                 4  
Chapter 2: Materials and Methods                        7  
  Kinetics Studies: general methodology                 7  
    Flash Photolysis                                    7  
    Rapid Mixing                                         9  
  NMR Studies                                           9  
  Unlabelled Protein Expression and Purification       10  
  $^{15}$N Labeled Protein Expression and Purification  12  
  Flash Photolysis                                      13  
  Rapid Mixing                                          13  
  NMR Data Acquisition                                  14  
  Data Analysis and Figures                             14
Chapter 3: Results from Kinetics Studies 15

Flash Photolysis 15
Rapid Mixing 16

Discussion 19
Implications 20
Future Studies 22

Chapter 4: Results from NMR Studies 23

Discussion 24
Future Studies 26

Chapter 5: Conclusions 28

References 29
List of Figures

Figure 1. Hexacoordination and Pentacoordination 2
Figure 2. CN-synHb and hx-synHb 6
Figure 3. SynHb Absorption Spectra 8
Figure 4. Flash Photolysis Traces 15
Figure 5. Flash Photolysis $k_{\text{obs}}$ vs. [CO] 16
Figure 6. Rapid Mixing Traces 17
Figure 7. Flash Photolysis $k_{\text{obs}}$ vs. [CO] 18
Figure 8. SynHb Wild Type HSQC Spectra 23
Hexacoordinate hemoglobins have been discovered that have yet to have a functional role confidently assigned to them. The eubacteria *synechocystis* sp. PCC 6803 has been found to have a truncated hemoglobin which is hexacoordinated. The structures of the hexacoordinated and cyanide bound forms of *synechocystis* hemoglobin have been solved, and a large shift in the E-helix has been seen upon ligand binding. Also, the loop connecting the E and F helix contains residues Tyr61 and Asp62 that are involved in hydrogen bonding networks in the cyanide-bound and hexacoordinate states respectively. A mutation was made to the Y61 residue to make it leucine, and this mutant protein was studied to investigate its role in regulating intramolecular histidine coordination to the heme iron. CO binding studies of flash photolysis and rapid mixing were performed to extract values for the CO binding rate and the rate of hexacoordination. The Y61L mutant protein has a rate of hexacoordination that is nearly 2.5 times greater than that of wild type *synechocystis* hemoglobin. Also, NMR spectra were taken on the wild type protein and it was revealed that a secondary conformation of the protein than that reported is present. Upon further inspection, it was suspected that the secondary conformation was most likely due to a covalent linkage to the heme from the H117 residue.
Chapter 1: Introduction

General Background

The first hemoglobins found were the tetrameric hemoglobin and myoglobin in mammalian blood and muscle. They are commonly known for their functions as oxygen transporters. Hemoglobin will transport oxygen from the lungs throughout the bloodstream, and myoglobin transports oxygen from blood to areas of high activity in tissues. For about the time that myoglobin has been studied, so too has legume hemoglobin (leghemoglobin), found in root nodules in legumes. Leghemoglobin provides the functions of facilitated diffusion of oxygen to the obligate aerobes that are involved in nitrogen fixation as well as oxygen scavenging.

The hemoglobin (Hb) monomer is characterized by a globin fold of eight alpha helices named A through H, which coordinate a heme prosthetic group with an iron in the center which may exist in two forms: the ferrous (2+) form or the ferric (3+) form. The porphyrin ring is planar with four nitrogens that coordinate the central iron. A nitrogen of the histidine on the proximal side of the porphyrin plane also coordinates the iron leading to the pentacoordinate state leaving a pocket on the distal side for exogenous ligands to bind.

The globin fold around the heme group serves three main purposes. First, it increases the heme solubility. Second, it shields the iron from oxidizing to the unreactive ferric state. Third, it regulates the binding of ligands. For example, the distal histidine in myoglobin will form a hydrogen bond with bound oxygen to stabilize the bound ligand. This also
helps discriminate against other ligands such as carbon monoxide. On the other hand, the leghemoglobin stabilizes bound ligand by lacking the hydrogen bonds to the proximal histidine (which is present in myoglobin) that increases the oxygen affinity of the iron (Kundu et al 2003). This change gives leghemoglobin increased oxygen affinity which allows it to both scavenge oxygen and facilitate oxygen diffusion.

So, although the tertiary structure of Hbs remains generally conserved, the residues surrounding the heme pocket exhibit a large degree of variability.

**Hexacoordination**

![Figure 1](image-url)  
**Figure 1.** A, the heme group showing a pentacoordinate state: the ligand binding site is left open in the absence of ligand. B, the heme group showing a hexacoordinate state: the ligand binding site is occupied by a histidine residue from the globin fold in the absence of ligand. *Taken from Trent et al. 2004.*

Recently, a class of Hbs has been found to exhibit the ability to reversibly coordinate the iron on the distal side in the absence of exogenous ligand (Fig. 1). This sixth coordination of the iron occurs by an endogenous histidine residue on the distal side of the heme. The first hexacoordinated hemoglobin was found in barley (Taylor et al 1994). Since then, hexacoordinated hemoglobins have been found to be present in mammals, plants and bacteria. These hexacoordinated hemoglobins exhibit spectral characteristics similar to
those of cytochrome b5 (cyt b5). However, unlike cyt b5, these hexacoordinate hemoglobins (hxHbs) show the ability to reversibly bind the distal histidine. Also, unlike myoglobin and leghemoglobin, hxHbs are not thought to just play a role in oxygen transport for two main reasons. First, their oxygen dissociation rates are significantly lower (sperm whale myoglobin: 15 s\(^{-1}\), soybean leghemoglobin: 5.6 s\(^{-1}\) and rice non-symbiotic hemoglobin (riceHb1): 0.038 s\(^{-1}\)) (Hargrove et al 2000A). Second, their concentrations \textit{in vivo} are in the micromolar to nanomolar range implying that they would be poor candidates for facilitating oxygen diffusion.

HxHbs have been found in various tissues of several organisms, but no specific physiological roles have been assigned to these proteins with any degree of certainty. It has been seen that hxHbs are upregulated in cases of hypoxic stress (Dordas et al 2003) and that they may play a role in nitric oxide detoxification and sensing (Brunori et al 2005). A pertinent step towards understanding the physiological roles of hexacoordinate hemoglobins is to understand the factors involved in hexacoordination.

A comparison of the structure of unliganded riceHb1 to the structure of cyanide bound barley hemoglobin (CN-barHb) reveals some of the factors that allow for the reversible hexacoordination (Hoy et al 2007). First, unfavorable interactions with the Phe\(^{\text{B10}}\) (B10 indicates the 10\(^{\text{th}}\) residue on the B helix from the N-terminus) residue helps to promote a pentacoordinate state (Smagghe et al 2006B). Second, the loops connecting the C and D helix (CD loop), and the loops connecting the E and F helix (EF loop), show a great degree of flexibility. Both loops are stabilized by hydrogen bonds in the pentacoordinate
form by intramolecular or intermolecular hydrogen bonding networks. Cyt b5, on the other hand, has a globin fold with shorter loops which do not allow for the flexibility that is seen in riceHb1 and barHb and so is unable to reversibly coordinate the iron.

A direct comparison of the ligand bound and unliganded structure of a hxHb would be useful for determining what specific factors govern hexacoordination. Currently, the structure of three proteins has been solved in the unliganded and ligand bound state. One is the crystal structure for mouse neuroglobin (Ngb) (de Sanctis et al 2006). The difference between the two structures did not show any large degree of conformational change. However, it was reported that the unliganded crystals were soaked in CO saturated solution and this soaking could shatter the crystals indicating that a conformational change occurred that could not be handled by the crystal. So, this may not be an accurate representation of the change in conformation upon ligand binding.

Another protein whose structures have been solved for the both states is the hemoglobin from the cyanobacterium *synechocystis sp.* PCC 6803 (synHb). The structure of synHb has been solved for the cyanide bound form and the hexacoordinate form (Hoy et al 2004, Trent et al 2004).

**Truncated Hemoglobins**

SynHb is a hexacoordinate protein that is found in the family of hemoglobins called truncated hemoglobins (trHbs). These trHbs are generally 20-40 amino acids shorter. Their defining characteristic is that instead of the “3-on-3” alpha helical fold for normal
Hbs, trHbs have a “2-on-2” fold structure. The trHbs are divided into three groups based on specific residues around the heme binding site and ligand access paths (Wittenberg et al 2002). SynHb falls into group I trHbs of which three other crystal structures have been solved in a ligand bound state: *Chlamydomonas eugametos* (CtrHb), *Paramecium caudatum* (PtrHb), and *Mycobacterium tuberculosis* (trHbN) (Pesce et al. 2000, Milani et al 2004). These structures are seen to have an apolar continuous or nearly continuous “ligand tunnel” connecting solvent to the ligand binding site. Of these three other trHbs only CtrHb is seen to exhibit properties of hexacoordination. However, it is only seen at pH10 with the endogenous coordination coming from the TyrB10 residue (Das et al 1999).

From the structures of synHb and CN-synHb, some distinct traits are seen that set this protein apart from the other three group I trHbs. One is the absence of a continuous ligand tunnel in the hexacoordinate form. However, in the ligand bound state a partial tunnel is seen to exist. Also, the hexacoordinate (hx-synHb) shows a shift in the E-helix inward toward the heme group for the histidine E10 to coordinate the heme iron.

A comparison of the hx-synHb and CN-synHb has been done to reveal the global tertiary structure differences as well as shifts in key residues around the heme pocket upon ligand binding (Trent et al 2004). It is seen that upon ligand binding, the A and upper B helices will shift toward the heme group, while the E helix containing the E10 residue responsible for hexacoordination will swing out and away from the heme (Fig.2, Right). This E helix shift is the largest tertiary structural shift observed so far in hxHbs or trHbs.
The E helix is seen to swivel about the EF loop which is termed the “hinge”. Upon ligand binding, the tyrosine residue (Y61) on the hinge is seen to form a hydrogen bonding network with residues H46 and the heme propionate group. It is thought that the Y61 bond to the distal histidine and the heme propionate group stabilizes the ligand bound conformation (Fig. 2, Left). Upon ligand release, the E helix swings back in so that the distal histidine can coordinate the iron. Here, the D62 residue on the hinge hydrogen bonds with the R64 residue to stabilize the hexacoordinate state (Fig. 2, Left). In order to better understand the effects of the hinge and the movement of the E-helix and factors that regulate it, a mutation of the Tyr61 residue to leucine was made. The leucine should eliminate the hydrogen bonding provided by the tyrosine, while not adversely affecting the protein structure.
Chapter 2: Materials and Methods

Kinetics Studies; general methodology

In order to examine the effects of the Y61 residue, the kinetics rate constants of ligand binding, “histidine on”, and “histidine off” must be measured. The basic scheme of the reaction is that the protein goes from a hexacoordinated state, to a pentacoordinated state, at which point the ligand can associate to the heme.

\[
Hb_{\text{hex}} \xrightarrow{k_{\text{H}}} Hb_{\text{pent}} + L \xrightarrow{k_{\text{L}}} Hb_{\text{L}} \quad \text{(Equation 1).}
\]

Flash Photolysis

In order to simplify this system it is desirable to try to establish a pseudo first order reaction which can be studied to establish some of the rate constants independent of the rest of the reaction. This is accomplished using the flash photolysis method. A laser is pulsed into the sample to photodissociate the bound ligand, in this case CO. The CO exits the protein matrix leaving the protein in a pentacoordinate deoxy state. Once the laser pulse is turned off the CO then rebinds the open ligand site. Using this method we have:

\[
Hb_{\text{hex}} \xrightarrow{k_{\text{H}}} Hb_{\text{pent}} \xrightarrow{k'_{\text{CO}}} Hb_{\text{CO}} \quad \text{(Equation 2).}
\]

This binding is monitored by observing either the increase in absorbance at 417nm (the CO bound peak) or the decrease in absorbance at 423nm (the deoxy peak) (Fig. 3).
Figure 3. Absorption spectra for the hexacoordinate, deoxy and CO bound states of synHb.

The change in absorbance will generally take the form of a single exponential or double exponential.

\[ \text{Abs}(t) = Ae^{-kt} \quad (\text{Equation 3.}) \]

or

\[ \text{Abs}(t) = Ae^{-\gamma t} + Ae^{-\gamma t} \quad (\text{Equation 4.}) \]

From these traces of CO rebinding, a value for \( k_{obs} \) can be extracted for each value of CO concentration. This can be fit to the equation (Hargrove et al 2000B):

\[ k_{obs} = k_H + k_{-H} + k_{CO}^{'} [CO] \quad (\text{Equation 5.}) \]

The slope of this plot yields the value of \( k_{CO}^{'} \).
Rapid Mixing

The principles of rapid mixing are similar to flash photolysis in that the absorbance traces are seen for CO rebinding. However, for this experiment, hexacoordinate protein is mixed in a stopped flow apparatus on extremely short time scales on the order of microseconds. This gives the full reaction. Since the histidine off rate, \(k_{-H}\) is the rate limiting step, the \(k_{obs}\) values will be governed by the values of \(k_H\) and \(k_{-H}\). A fit of \(k_{obs}\) versus [CO] would give a plot that corresponds to the following equation (Smagghe et al 2006A):

\[
k_{obs} = \frac{k_{-H} \times k'_{CO}[CO]}{k_{-H} + k_H' + k'_{CO}[CO]} \quad (Equation \ 6.)
\]

From this, the values of \(k_H\) and \(k_{-H}\) can be extracted. Also, if it is seen that \(k_H \gg k_{-H}\), then the intercept of the \(k_{obs}\) vs. [CO] plot would give a good approximation of \(k_{-H}\).

NMR Studies

To further examine the shift of the E-helix, NMR studies can be performed. The advantage of examining the structural shift using NMR techniques as opposed to crystallography techniques is that NMR allows the study of synHb in a solution state rather than “locked in” in the crystal. It is possible to use NMR techniques to examine the dynamics of the protein backbone. This can give a better understanding of the difference in flexibility of the E-F loop compared to the rest of the protein as well as compare
dynamics rates for the hinge region and E-helix in a ligand bound and hexacoordinated state.

The solution state structure of synHb in the hexacoordinated state has been solved (Falzone et al 2002). From this, the chemical shift values for the backbone amides and associated hydrogen atoms can be taken giving an assignment of chemical shift values for each backbone amide in the protein. Then, $T_1$ and $T_2$ relaxation time values as well as NOE values can be measured. These values fit with the Model-Free formalism can give an order parameter $S^2$ which indicates general rigidity of the amide corresponding to each residue with $S^2 = 0$ meaning complete isotropic movement and $S^2 = 1$ corresponding to complete rigidity. While the $S^2$ order parameter gives a general idea for the “floppiness” of the backbone at each point in the protein, it does so for timescales on the order of nanoseconds. Experiments can be performed that measure the dynamics rates of the protein backbone for timescales on the order of microseconds or even milliseconds.

**Unlabelled Protein Expression and Purification**

The synHb wild type gene cloned into pET28 was obtained. The mutant DNA for Y61L was generated through a standard mutagenesis PCR. The mutant DNA was transformed into the host strain BL21 (Invitrogen). Expression cultures were grown in 20L of 2xYT at room temperature for 24 hours. The culture was centrifuged on a Beckman-Coulter JA-10 rotor at 10,000 RPM for 10 minutes at 4°C and pelleted cells were collected. Cell pellets were resuspended in 20mM Tris buffer (pH 7) and stored at -20°C. Cells were thawed
and lysozyme was added to a concentration of 1mg/mL and left at room temperature for 2 hours. The mixture was cooled to 4°C and sonicated on ice at high power for 2 minutes, then allowed to cool. This mixture was centrifuged on a Beckman-Coulter JA-14 rotor at 14,000 RPM for 10 minutes at 4°C and the red supernatant was kept. The supernatant was brought to an ammonium sulfate concentration of 40% (percent saturation in distilled water) and centrifuged again where the supernatant was kept. The supernatant was brought to a 90% ammonium sulfate concentration and centrifuged again and the pellet was kept. The pellet was resuspended in 10mM sodium phosphate, 300mM sodium chloride (pH 7) and dialyzed to 10mM sodium phosphate, 300mM sodium chloride (pH 7). This solution was loaded onto a cobalt IMAC and washed with the loading buffer (10mM sodium phosphate, 300mM sodium chloride, pH7) until the absorbance at 280nm was below 0.01. The protein was eluted by adding loading buffer with 100mM imidazole added to the washing/loading buffer and fractions were collected. Fractions with ratio (absorbance at 407nm to absorbance at 280nm) lower than 3 and fractions too dilute were discarded. The remaining fractions were combined and concentrated to approximately 1mM hemoglobin concentration as calculated by absorbance at 417nm (CO bound extinction coefficient of 160 mM⁻¹cm⁻¹). 750µL of protein was taken and oxidized by adding 2-3 crystals of potassium ferricyanide. This solution was added onto a sepharose G25 column with 20mM sodium phosphate (pH 7) elution buffer to provide buffer exchange and to remove potassium ferricyanide.
**15N Labeled Protein Expression and Purification**

The wild type DNA was transformed into the host strain BL21 (Invitrogen) and grown on an LB agarose plate overnight. Colonies were then added to 5mL aliquots of LB media and allowed to grow for 10-12 hours. 1mL from these cultures were added into 2L flasks containing 1L of M9 minimal media with the $^{15}$N provided by the added ammonium chloride. The growth was allowed to reach an OD$_{600}$ of 0.9 at which point IPTG was added to a final concentration of 1mM. Protein purification was performed in the same manner as for unlabelled protein with two exceptions. First, cells were lysed using an EmulsiFlex-C5 High Pressure Homogenizer at a pressure of 15,000 PSI until a significant reduction in solution viscosity (approximately 5 times). Second, after lysis and extraction of the supernatant, the protein was stirred for 1-2 hours in a solution of 0.1M NaOH containing hemin chloride. This was done to reconstitute heme into the apoprotein which did not contain heme due to a lack of iron in the minimal media. The pH was reduced to 5.5 to precipitate out excess heme, and centrifuged on a Beckman-Coulter JA-25 rotor at 25,000 RMP for 30 minutes at 4°C.

A second sample of protein was made with the following differences in expression and purification protocols. First, a plasmid lacking the histidine tag was used. Second the M9 minimal media had added ferric citrate. Third, no step was taken to reconstitute heme to the globin as the ferric citrate in the media provided enough iron to have the cell extract contain heme as observed by the red color of the cell extract compared to the white color of the cell extract with no iron added. Fourth, rather than using a cobalt IMAC column,
the protein was loaded onto a DEAE column, washed with 50mM NaCl and eluted with 70mM NaCl. Fractions with ratio greater than 3 were collected and concentrated.

**Flash Photolysis**

Saturated nitrogen and carbon monoxide buffers were prepared by bubbling N\textsubscript{2} and CO gas through 100mM potassium phosphate buffer (pH 7) for approximately 20 minutes at which point dry sodium dithionite was added into solution and the buffers were stored in air tight syringes. Buffers of varying CO concentrations were prepared by diluting saturated CO buffer with N\textsubscript{2} buffer into air tight syringes. Samples were prepared by adding protein until an approximate concentration of 20-40µM with the appropriate CO buffer. Photodissociation of bound ligand was accomplished using a 10 Hz Continuum Surelite I YAG at a 532nm wavelength on a 5ns pulse. The absorption at the CO bound peak of 417nm or deoxy peak of 423nm was monitored with the light source being a 75-W xenon lamp and the detector being an Optometrics DMC1-03 photomultiplier. The detection was triggered electronically to take absorption at the specified wavelength immediately after the pulse. Traces were collected on PicoScope software.

**Rapid Mixing**

Nitrogen and carbon monoxide buffers were prepared as described for flash photolysis. Rapid mixing experiments were done on a BioLogic SFM 400 stopped-flow reactor. Spectra were collected using a MOS-250 spectrophotometer on Biokine software. Protein was added into N\textsubscript{2} buffer to an approximate concentration of 20-40µM and sealed in an
an air tight syringe. 150µL of protein solution and CO buffer were mixed with a dead time of 3µs and the absorbance at 417nm was taken.

**NMR Data Acquisition**

HSQC data were acquired on a Bruker Ultrashield 700 MHz spectrometer. Samples were prepared in 10mM phosphate buffer at pH 7.5 with 10% D$_2$O. Quadrature detection was accomplished with the States-TPPI method. Water suppression was accomplished using a WATERGATE solvent suppression scheme (Piotto *et al* 1992).

**Data Analysis and Figures**

For kinetics experiments data analysis was done and figures were created using Igor Pro. For NMR experiments, spectra were collected on Bruker software, processed using NMRPipe and analyzed on NMRViewJ.
Chapter 3: Results from Kinetics Studies

Flash Photolysis

Data were collected for both synHb wild type and synHbY61L. Data were collected at CO concentrations of 100µM, 200µM, 500µM and 1000µM. Change in absorbance at 417nm for wild type and 423nm for Y61L were recorded as time progressed. Multiple data sets were collected and concatenated together to have data points for the fast and slow phase of CO rebinding.

![Figure 4](image)

Figure 4. Traces of changes in absorbance for synHb. The traces for wild type synHb and synHbY61L are shown with ΔAbs shown normalized, and time shown in seconds on a log₁₀ scale. Left, synHb wild type. It is clearly seen that for the traces of 1000µM CO, 500µM CO and 200µM CO the CO rebinding shows only the single phase while the 100µM CO is biphasic rebinding. Right, SynHbY61L. Traces for all four CO concentrations clearly show biphasic CO rebinding. It should be noted that the wild type data were collected at 417nm, the CO bound Soret peak, and so increase as CO rebinds, while the synHbY61L data were collected at 423nm, the deoxy peak, and so decreases as CO rebinds.

The traces either showed a single exponential fit, indicating a pseudo first order reaction, or a double exponential fit indicating a biphasic reaction (Fig. 4). The fits were dependent
on the protein and CO concentrations. The wild type traces for CO concentrations of 1000µM, 500µM, and 200µM were fit to a single exponential curve (equation 3). All other traces were fit to a double exponential curve (equation 4).

For single exponential traces $k_{obs} = k$, while for double exponential traces $k_{obs} = \gamma_1 + \gamma_2$. $k_{obs}$ was plotted versus [CO].

![Figure 5](image)

**Figure 5.** A plot of $k_{obs}$ vs. [CO]. Linear fits were added in and $k'_{CO}$ and $k_H$ were extracted. $k_H$ is extracted under the assumption that $k_{H} \ll k_H$ which is seen to be the case upon inspection of rapid mixing data.

Linear fits of the data were made and the parameters of $k'_{CO}$ and $k_H$ were extracted using equation 5 (Fig. 5).

**Rapid Mixing**

Data were collected for synHb wild type and synHbY61L. For the wild type, data were collected for 12µM, 25µM, 50µM, 125µM, 250µM and 400µM. For the mutant, data
were collected for 12µM, 25µM, 50µM, 100µM, 150µM, 250µM and 500µM. Absorbance was recorded at 417nm for a time on the order of 1 second. The reason for the differences in CO concentrations for the wild type and mutant are simply because the wild type data was collected at a later date when different CO buffers were made for flash photolysis experiments. However, the data points still give a generally reasonable fit for $k_{obs}$ vs. [CO].

**Figure 6.** Traces for changes in absorbance for synHb. Left, synHb wild type. Right, synHbY61L. Absorbance change versus time in seconds is shown. Fits are shown as dashed lines. Traces for wildtype protein of CO concentration 250µM, and 400µM were fit to a single exponential. All other plots were fit to a double exponential curve.

The traces for the wild type of CO concentrations of 250µM and 400µM were fit to a single exponential (equation 3) $k_{obs}$ was set to $k$. The other traces were fit to a double exponential (equation 4) (Fig. 6). The faster rate was taken as $k_{obs}$. For these traces the faster rate accounted for greater than 75% of the change in absorbance (i.e. if $\gamma_2$ represented the faster rate, $\frac{A_2}{A_2 + A_1} > 0.75$).
Figure 7. A plot of $k_{\text{obs}}$ versus [CO] for synHb wild type and synHbY61L. $k_H$ and $k_{-H}$ are extracted from the fit data and $k'_{CO}$ is taken from flash photolysis data.

$k_{\text{obs}}$ was plotted versus [CO]. The plot was fit to equation 6. $k'_{CO}$ was set to $53 \mu M^{-1}s^{-1}$ and $140 \mu M^{-1}s^{-1}$ for synHbY61L and synHb wild type respectively. $k_h$ and $k_{-H}$ were extracted from the fit data (Fig. 7).

<table>
<thead>
<tr>
<th>Table 1. Ligand association and hexacoordination rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k'_{CO}$ ($\mu M^{-1}s^{-1}$)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>SynHb WT</td>
</tr>
<tr>
<td>SynHbY61L</td>
</tr>
<tr>
<td>RiceHb1*</td>
</tr>
<tr>
<td>Ngb*</td>
</tr>
</tbody>
</table>

*Taken from Smagghe et al 2006A.*
Discussion

As seen from the results of the rapid mixing fit, the values of $k_H$ for the wild type and mutant show a difference by a factor of approximately 2.5. Two other factors verify that this difference in $k_H$ is consistent. First, the value of the intercept from the flash photolysis experiment gives a value for $k_H + k_{-H} \cong k_H$ (assuming $k_H \gg k_{-H}$) that correlates well with the value obtained from the fit of the rapid mixing data. Another interesting feature supporting the fact that the histidine on rate is higher for that of the mutant are the traces from the flash photolysis data exhibiting biphasic ligand binding. This shows that for the mutant, competition for the ligand binding site, most likely from the endogenous histidine, appears much more readily than for the mutant. From this, we can see state with certainty that the Y61 residue plays a significant role in affecting ligand binding rates. The hydrogen bonding network is seen in the ligand bound state, and so removal of the hydrogen bonds by the replacement of Y61 with a leucine allows for an increase in the affinity for the hexacoordinated state.

Also, $k'_{CO}$ for the mutant is significantly smaller than that of the wild type by a factor of roughly 2.5. The difference in rate constants for $k'_{CO}$ support the conclusion that the Y61 hydrogen bonding stabilizes the pentacoordinate state. Once the ligated CO is photodissociated from the protein matrix, the fast phase of ligand binding competes with the distal histidine for rebinding. By stabilizing the E helix in the open form where the distal histidine is moved out and away from the heme pocket, the ligand has less competition for rebinding and thus would show higher values of $k'_{CO}$. 
Both wild type and mutant synHb are seen to have significantly higher rates of hexacoordination. The degree of structural shift between CN-synHb and Hx-synHb and between riceHb1 and CN-barHb is roughly of the same magnitude. The shift for riceHb1 is seen to be movement of the E helix along its helical axis where the \( N^{\varepsilon2} \) atom moves from 2.1Å to 5.1Å from the heme iron. This is accompanied by structural rearrangements in the EF loop and CD loop whose residues are then coordinated by intermolecular and intramolecular hydrogen bonding networks. So the difference in degrees of shift does not account for the large variation in the histidine binding rates. The change associated with removing a hydrogen bond governing structural shift is approximately a factor of 2.5 while the difference between the \( k_H \) of riceHb1 and synHb is a factor of over 50. It has been shown that for riceHb1 the Phe\(^{B10}\) residue hinders the distal histidine in hexacoordination to a point where it allows for the distal histidine to coordinate the iron, but to do so reversibly (Smagghe et al 2006B). SynHb on the other hand contains a tyrosine at the B10 position which is seen to coordinate bound ligand rather than hinder histidine coordination to the heme. So, though the degree of structural change between synHb and riceHb1 are similar, the interaction of the distal histidine with Phe\(^{B10}\) may play be the major factor governing \( k_H \).

**Implications**

The hydrogen bonding system of Y61 could be a regulating factor for ligand binding. SynHb exhibits a tertiary structural shift on a scale not seen in any other hemoglobin. One of the purposes of the globin fold is to shield the heme iron from oxidation to the unreactive ferric species. However, the globin fold of synHb shows a very high degree of
flexibility compared to other Hbs as it has the highest tertiary conformational shift yet observed upon reversible ligand binding. In order to serve the role of shielding the iron, the globin flexibility has to be regulated. The variation in the histidine on rate for the synHb wild type versus synHbY61L shows that the tyrosine hydrogen bonding is one of the regulating factors of this flexibility. The hydrogen bonds formed by the Y61 residue could provide stability to the helix in the ligand bound state thereby helping to restrict the overall flexibility of the E helix.

It may be important to note the difference in effect that the changing the hinge residue has compared to changing one of the heme pocket residues. Mainly the hinge residue accounts for a factor of about 2.5 for the difference in $k_H$ while the Phe$^{B10}$ residue accounts for a factor of 250 (Smagghe et al 2006B). However, hinge residues may play a greater role in directing the movement and position of the E-helix rather than regulating the histidine affinity. Either way, it is clear that the hinge residue Y61 plays a role in regulating the conformational change.

While other Hbs are known to serve as O$_2$ transporters, they can manage to do so without evolving the capability to reversibly shift the globin fold. This shift has to be regulated so that the globin fold can shield the heme iron allowing for specificity in ligand binding and to prevent autooxidation of the iron. It is unlikely that synHb evolved the ability to exist in two different ligated states in order to serve as a gas transportation protein. Furthermore, synHb concentrations in vivo have not been seen that can support the possibility of an O$_2$ transportation role.
Also, unlike other trHbs, synHb has a discontinuous ligand tunnel in the hexacoordinate state. However, the partial ligand tunnel exists with the E helix shifted outwards. It is possible that the hinge residues serve in a ligand sensing mechanism allowing for the conformational change to occur in the presence of ligand. It is also possible that the hinge residues help stabilize the ligand bound confirmation with the E helix shifted outwards to allow for faster ligand exit by providing the open ligand tunnel.

**Future Studies**

Another amino acid in the hinge region is D62. While Y61 stabilizes the ligand bound state, the D62 hydrogen bonding is seen to occur in the hexacoordinate state. So, if a mutation to the Y61 residue were to increase the value of $k_H$, a mutation to the D62 residue could most likely serve to decrease the value of $k_H$ or increase the value of $k_{-H}$. 
Chapter 4: Results form NMR Studies

Heteronuclear single quantum correlation (HSQC) spectra were taken for uniformly $^{15}$N labeled protein samples. It was seen from these spectra, that there were nearly double the number of expected peaks.

![Figure 8. HSQC Spectra of synHb wild type. Shown is a contour plot of the peak intensities in two dimensions. D1 represents the HN dimension and D2 represents the $^{15}$N dimension. Overlayed squares represent previously reported chemical shift data. Upon close inspection, it is seen that there is generally an unassigned peak for each assigned peak.](image)

Approximately half the peaks in the HSQC correlated to the reported chemical shift data. However, the remaining half seemed to have chemical shift values that were significantly different. This leads to the conclusion that there must be heterogeneity in the protein in giving a pair of peaks for each amide corresponding to the two different conformations. The original intent of this study was to establish $T_1$, $T_2$ and NOE values for each peak.
Then using the model-free formalism, the order parameter of $S^2$ could be obtained to give general dynamics properties of the protein backbone. However, residue assignments are needed for each peak in the HSQC spectrum to be able to assign dynamics data. Although, the unknown peaks on the edges of the spectrum can be obviously attributed to the known peaks, the central region shows too much spectral density leaving many ambiguous peaks. It would be possible to calculate dynamics parameters for ambiguous peaks, but without the residue assignment, there would be no way to know what part of the protein correlated with the calculated parameters.

Several factors were thought to play a role in this secondary set of unknown peaks, and being unable to identify the cause of the heterogeneity, the line of studies involving backbone dynamics was not pursued.

**Discussion**

One possibility for the second conformation is that the protein could be existing in the pentacoordinate and hexacoordinated state, and the shift of the E-helix could result in a significant variation in the chemical shift values. From kinetics data, however, it is seen that the ratio of hexacoordinate to pentacoordinate proteins $\left( K_H = \frac{k_H}{k_{-H}} \right)$ is greater than 0.99, so the protein should exist as a majority in the hexacoordinate state in the absence of ligand. However, peak volumes for the reported conformation and the unknown conformation are on the order of 2:3 indicating that there must be some other factor.
It was also thought that there could be a variation in the chemical shift due to the histidine tag region of the protein on the N-terminus. Also the addition of excess heme to reconstitute apoprotein, as was done for the original solution structure of synHb, could have led to excess heme in solution or the heme associating with the globin fold in unpredicted ways such as associating with the histidine tag. These theories were tested by expressing and purifying a protein which lacked the histidine tag and that did not require addition of heme (see Materials and Methods).

Compared to the crystal structure of synHb, there is one major difference in the solution structure: the absence of a covalent heme linkage to the globin fold from H117 to the heme vinyl group. Before the existence of the heme covalent bond was discovered (Vu et al 2002), two forms of the protein were noted. Studies were performed on only the one form which lacked the covalent bond and so the solution structure of synHb lacks the covalent bond. Partial NMR structures of the two wild type species and the H117A mutant which is incapable of forming the covalent linkage have been solved and show little variation in the global conformation (Falzone et al 2002, Vu et al 2004A, Vu et al 2004B). However, an alignment of the Ca atoms of the crystal structure to the solution structure yielded and average r.m.s.d. value of 2.6Å. On the other hand, an alignment based on heme position yielded an average r.m.s.d. of 3.76Å suggesting that the presence of the covalent bond serves to reorient the heme (Hoy et al 2004). Heme rearrangement could be a cause of the appearance of the secondary set of peaks as even a small rearrangement in structure can lead to variations in electron deshielding which could give rise to variation in chemical shift.
So, a strong possibility for the secondary set of peaks is the formation of the covalent bond between the heme and globin fold at position H117. The solution structure of synHb was reported in the absence of a covalent bond, while the crystal structure was reported with the bond present. However, since it is unknown currently what factors affect the formation of the covalent bond, it is unknown whether this is the case.

**Future Studies**

One of the main factors to examine in the unknown conformation is actual structure. If it would be at all possible to solve the solution structure of the second conformation, then a clearer understanding of the cause of the secondary conformation (whether it is the covalent bond or some other factor) could be ascertained. Furthermore, assigning the chemical shift values to residues, would allow for the study of the dynamics of the protein backbone to give a clearer understanding of the E-helix shift. Also, a similar expression and purification method that was used for obtaining wild type synHb could be used for obtaining the H117A mutant labeled with $^{15}$N and would serve to answer this question as the H117A mutant is unable to form the covalent linkage to the heme. However, the wild type protein with the covalent linkage formed has been shown to be more stable under variations in temperature and pH compared to the H117A mutant (Lecomte *et al* 2001, Vu *et al* 2000A,B). This could indicate that of the two states that have been seen to exist in solution NMR, the state with the covalent linkage may be the best representation of what exists *in vivo*. A study of dynamics of the H117A mutant
could provide an inaccurate conclusion about the dynamics of the EF loop and E helix as it exists *in vivo*.

Finally, if the covalent bond is indeed the cause of the secondary structure, it could be possible by subjecting the protein to different conditions such as pH changes, temperature changes, or heme iron reduction or oxidation, and then measuring the HSQC and examining the ratio in peak volumes for the reported and unknown peaks, to establish factors that influence the formation of the covalent bond.
Chapter 5: Conclusions

Examination of CN-synHb and hx-synHb revealed a large structural shift in the E helix that was hypothesized to be partially regulated by the formation of a hydrogen bonding network between the heme propionate group, Y61 on the EF loop, and the distal histidine H46 in the cyanide bound state. Upon examining the ligand binding and hexacoordination kinetics rates, it was seen that changing the Y61 residue to a leucine, eliminating the hydrogen bonding network, increased the hexacoordination rate by a factor of approximately 2.5. This revealed that the Y61 hydrogen bonding network plays a role in regulating the flexibility of the globin fold that allows for the large conformational shift of the E-helix. However, the role of the hydrogen bonding network in the hexacoordinate state with the D62 residue still remains uncertain and further studies may be needed to verify it as stabilizing the hexacoordinate state.

The NMR spectra taken of the wild type synHb showed a heterogeneity in the protein that was neither attributed to the presence of the histidine tag, addition of heme to reconstitute apo-protein or a presence of a significant concentration of the deoxy form of the protein, leaving the main probability as a heme reorientation due to the formation of a covalent bond with the H117 residue. Solution of the unknown structure could confirm this, as well as allow for the study of the dynamics properties of the globin fold for the transition from the hexacoordinate to ligand bound state.
References


Das, T.K. et al. (1999) Identification of the ligands to the ferric heme of Chlamydomonas chloroplast hemoglobin: evidence for ligation of tyrosine-63 (B10) to the heme. *Biochemistry* 38 (46), 15360-15368


Lecomte, J.T. et al. (2001) Binding of ferric heme by the recombinant globin from the cyanobacterium Synechocystis sp. PCC 6803. *Biochemistry* 40 (21), 6541-6552


Smagghe, B.J. et al. (2006) Slow ligand binding kinetics dominate ferrous hexacoordinate hemoglobin reactivities and reveal differences between plants and other species. *Biochemistry* 45 (2), 561-570


