Published in final edited form as:
Biochemistry. 2018 October 09; 57(40): 5785–5796. doi:10.1021/acs.biochem.8b00752.

Replacement of the Distal Histidine Reveals a Non-Canonical Heme Binding Site in a 2-on-2 Hemoglobin

Dillon B. Nye and **Juliette T. J. Lecomte**[*]
T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD, 21218, United States

Abstract

Heme ligation in hemoglobin is typically assumed by the "proximal" histidine. Hydrophobic contacts, ionic interactions, and the ligation bond secure the heme between two α- helices denoted E and F. Across the hemoglobin superfamily, several proteins also use a "distal" histidine, making the native state a *bis*-histidine complex. The group 1 truncated hemoglobin from *Synechocystis* sp. PCC 6803, GlbN, is one such *bis*-histidine protein. Ferric GlbN, in which the distal histidine (His46 or E10) has been replaced with a leucine, though expected to bind a water molecule and yield a high-spin iron complex at neutral pH, has low-spin spectral properties. Here, we applied NMR and electronic absorption spectroscopic methods to GlbN modified with heme and amino acid replacements to identify the distal ligand in H46L GlbN. We found that His117, a residue located in the C-terminal portion of the protein and on the proximal side of the heme, is responsible for the formation of an alternative *bis*-histidine complex. Simultaneous coordination by His70 and His117 situates the heme in a binding site different from the canonical site. This new holoprotein form is achieved with only local conformational changes. Heme affinity in the alternative site is weaker than in the normal site, likely because of strained coordination and a reduced number of specific heme–protein interactions. The observation of an unconventional heme binding site has important implications for the interpretation of mutagenesis results and globin homology modeling.

Graphical Abstract

---

[*] To whom correspondence should be addressed: lecomte_jtj@jhu.edu, Tel: (410) 516-7019.

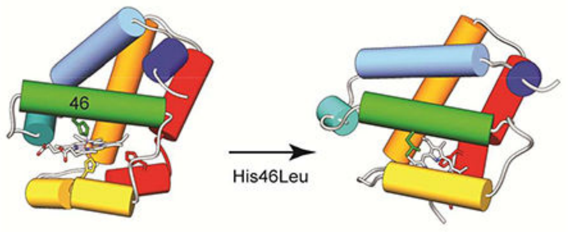
ASSOCIATED CONTENT

Supporting Information

Tables of chemical shifts, longitudinal relaxation times, and nuclear Overhauser effects; NMR spectra for longitudinal relaxation time determination and intensity recovery plots; assigned ^1H-^{15}N HSQC spectra; portions of NOESY spectra, chemical shift perturbation plots, NOE diagrams and ^1H-^{15}N HSQC spectra supporting the reversible relocation of the heme and minor secondary structure alterations; apoprotein ^1H-^{15}N HSQC spectra and pressure response of NMR data; electronic absorption spectra and their response to pH.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest with the contents of this article.



Keywords

truncated hemoglobin; heme coordination; hexacoordinate hemoglobin; mesoheme; NMR; *Synechocystis*

Steady improvement in genome sequencing has provided structural biologists with an enormous number of new sequences across all superfamilies of proteins. Hemoglobins (Hbs) are no exception to this sudden increase in information. Thousands of hypothetical globin genes are now available with multiple representatives in nearly all forms of life. Traditionally viewed as oxygen transporters, Hbs assume other functions related to the management of small molecules, principally reactive oxygen and nitrogen species.[1–3] These other functions are emphasized in unicellular organisms under aerobic and anaerobic conditions. Unlike for proteins devoid of cofactors, however, functional prediction in globins and rational design based on the globin fold are complicated by the exquisite sensitivity of the reactive heme group to small variations in structure. To this day, it is a challenge to anticipate such a fundamental property as protein ligands to the iron given a primary structure unequivocally corresponding to a globin. Without clear determinants of heme ligation and consequently reactivity, the wealth of information contained in expanding sequence databases cannot be fully exploited.

In efforts to explore and rationalize the natural range of chemistry supported by the Hb scaffold, we and others have used GlbN, the monomeric Hb from the cyanobacterium *Synechocystis* sp. PCC 6803, as a model protein.[4–6] GlbN is a representative of the “truncated” lineage of the superfamily; as such, it is several residues shorter than mammalian myoglobins, but it preserves essential structural features of the globin fold: the B, E, G and H helices assemble in an orthogonal bundle and form a hydrophobic core while the *b* heme group is captive between the E and F helices. The F helix provides the strictly conserved proximal His70 (F8 in myoglobin notation) as axial ligand to the heme iron. In GlbN, a second axial ligand, His46 (E10), is found on the opposite (distal) side of the heme.[7] *bis*-Histidine ligation classifies GlbN as a “hexacoordinate” globin, so-called because the heme iron is coordinated by two protein side chains as well as the four pyrrole nitrogens of the porphyrin.

Endogenous hexacoordination is observed in many globins throughout the superfamily, as a native property,[8] a stable misfolded state,[9] or a species appearing transiently in folding processes.[10,11] As a native state property, *bis*-histidine ligation tunes the iron reduction potential, establishes activation energy barriers for substrate binding, and determines the propensity for various types of reactions. For these reasons, the attributes that allow or prevent a distal histidine to serve as a ligand have been a focus of investigation.[12] The difference between GlbN and myoglobin, which uses only the proximal histidine as axial ligand in the native state, highlights the importance of scaffold rigidity and the need for the description of the conformational space accessible to various globins.

GlbN is an excellent subject to study the role and extent of plasticity in the globin fold. Figure 1 illustrates the *bis*-histidine structure[13] and the changes occurring when His46 is displaced by an exogenous ligand, cyanide in this case.[14] A trio of residues, Tyr22 (B10), Gln43 (E7), and Gln47 (E11) is introduced to the distal pocket to form a hydrogen bond network with bound cyanide. When using the coordinate transformation matrix minimizing the overall Cα rmsd between the two structures, displacements are systematically greater than 2 Å in the A and E helices and the beginning of the B helix. Gly63, which initiates the F helix in the *bis*-histidine structure, moves by more than 11 Å. Overall, helices rotate and translate as rigid objects while remaining nearly intact. Loops and turns, endowed with some conformational freedom, are key to the rearrangement.

Whether structural strains are imposed in the *bis*-histidine state or in the exogenously coordinated state has been investigated with studies of the wild-type (WT) GlbN,[5] Zn-substituted GlbN,[15] and His46 variants.[12,14] For the distal variant proteins, a reasonable expectation is that, by eliminating an axial ligand, the structure will relax and the iron will assume a high-spin character, as a pentacoordinate complex (ferrous state, $S = 2$, or ferric state $S = 5/2$) or a water-bound hexacoordinate complex (ferric state, aquomet, $S = 5/2$). Electronic absorption, EPR, and NMR studies of ferric H46A and H46L GlbN, however, do not fulfill this expectation. Ferric H46L is almost entirely low-spin ($S = 1/2$)[16] whereas H46A is a mixture of a low-spin and an aquomet high-spin complex.[7,16] Ferrous H46A and H46L GlbN also show a degree of low-spin ($S = 0$) character by electronic absorption spectroscopy or EPR.[5,7,16] Because these observations are made in the absence of exogenous ligand, the coordination status of the His46 variants betrays an unusual behavior and merits elucidation.

The structural rearrangement shown in Figure 1 suggests that any of the residues forming the hydrogen bond network in the cyanomet complex may be able to reach the iron in the H46L variant. These candidates were tested with puzzling results: Y22F/H46L GlbN displays NMR signatures of an unchanged ferric complex compared to H46L GlbN,[16] and the Q4-L/H46L replacement does not eliminate fully the low-spin properties of the ferrous state.[5] Inspection of the *bis*-histidine and cyanide-bound structures leaves as possibilities Gln47 and other less obvious candidates. In this work, our goal was to identify the alternative distal ligand(s) in H46L GlbN to document the plasticity of the GlbN fold. We resorted to mutagenesis and the use of a modified heme in NMR spectroscopy studies to analyze the distorted GlbN structure. We show that a single amino acid replacement can lead to a profound perturbation of the heme–protein topology without extensive alteration of the polypeptide fold. The results contribute to an improved understanding of the role of protein flexibility in limiting the properties of globins and emphasize the need for systematic experimental characterization of newly discovered globin sequences.

MATERIALS AND METHODS

Protein Purification.

Variant GlbN proteins were overexpressed in *E. coli* BL21 cells using a pET3c plasmid vector. H46L and H46L/Q47L apoGlbNs partitioned primarily into the cell lysate, as opposed to inclusion bodies, and were purified according to published procedures.[16,17] Porcine hemin chloride (Sigma, 20 mg/mL in 0.1 M NaOH) was added to the crude lysate until full holoprotein reconstitution was observed by electronic absorption spectroscopy.[16] This procedure produced a variable amount of ferrous O_2-bound complex, and 1 mM $K_3[Fe(CN)_6]$ was added to ensure a homogenous ferric product. The holoprotein was then purified by size exclusion and anion exchange chromatography. For the H46L/H117A GlbN variant, the apoprotein was purified directly from the cell lysate by size exclusion and anion exchange chromatography. Hemin chloride was added to reconstitute the holoprotein, which was separated from excess hemin by an additional round of size exclusion chromatography. Purity was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and mass spectrometry. Typical yield was 100 mg/L. Uniformly ^{15}N or ^{15}N, ^{13}C labeled protein was purified in the same manner from cells grown in M9 minimal medium containing $^{15}NH_4Cl$ and ^{13}C-glucose as the sole nitrogen and carbon sources. Following the final chromatography step, purified holoproteins were exchanged into storage buffer (0.5 mM Na/K phosphate pH 7), lyophilized, and stored at −20 °C.

Reconstitution with Mesoheme.

Purified H46L and H46L/Q47L GlbN holoproteins were converted to their apoprotein forms using the procedure of Teale.[18] Fe(III) mesoporphyrin IX chloride (mesoheme, Frontier Scientific, 5–10 mg/mL in 0.1 M NaOH) was added in ~2-fold molar excess to the stirred apoprotein solution (200 μM GlbN, 20 mM Na/K phosphate, pH 7.2, 5–10 mL) and the reconstitution reaction was allowed to proceed for several hours at 4 °C. The modified holoprotein was separated from excess mesoheme by anion exchange chromatography, exchanged into storage buffer, and lyophilized if not immediately used.

Electronic Absorption Spectrophotometry.

Electronic absorption spectra were collected using a Cary50 UV-Vis spectrophotometer. Spectra of ferric complexes were collected from 800 to 260 nm in 1-nm steps using a 0.1-s averaging time. Ferrous deoxy samples were generated by addition of 2 mM sodium dithionite (DT, Alfa Aesar) and spectra were collected from 650 to 350 nm every 1 min. pH titrations were performed in 5 mM Na/K phosphate as reported previously.[19] Concentrations were determined on a per-heme basis using extinction coefficients obtained by the pyridine (SigmaAldrich) hemochromogen assay.[20,21]

Circular Dichroism.

Circular dichroism (CD) spectra of GlbN variants (15 μM, 25 mM Na phosphate pH 7.4–7.5) were collected using an Aviv Model 420 spectropolarimeter. Far-UV spectra were collected with a 1-mm path cuvette length over the range of 300 to 190 nm in 1 nm steps using a 3-s averaging time. Soret CD spectra were collected on the same sample over the range 500–300 nm using a 1-cm path length cuvette. The sample temperature was maintained at 25 °C.

Heme Transfer Kinetics.

Horse myoglobin (Mb, Sigma) was converted to the apoprotein state by the procedure of Teale, as above, and further purified using a 100 × 2.5 cm G-50 (Sigma) size exclusion column. Concentrations of apoMb were determined using the calculated extinction coefficient of 13,980 M^{-1} cm^{-1} at 280 nm.[22] ApoMb was added to a ferric H46L GlbN sample (5 or 10 μM, 100 mM Na phosphate pH 7.2) and absorbance spectra were acquired in intervals of 70 s for 2 h. The reaction kinetics were found to be independent of H46L GlbN and apoMb concentrations (above a 2.5-fold molar excess apoMb). Singular value decomposition was applied to kinetic data sets and two abstract vectors were globally fit to a single exponential using Mathematica (Wolfram Research, Champaign, IL). The apparent first-order rate constant is taken to represent the heme k_{off}.[2-] In a similar experiment, the addition of apoMb was immediately followed by addition of 2 mM DT.

Heme Crosslinking in H46L/Q47L GlbN.

GlbN is capable of spontaneously forming a covalent linkage between His117 Nε2 and the heme 2-vinyl Cα in the *bis*-histidine ferrous state.[24] The linkage is made rapidly ($k \sim 0.4\ s^{-1}$ at pH 7[25]) by addition of DT to recombinant ferric protein. H46L GlbN is less reactive, but aided by added imidazole or cyanide.[16] Based on these results, the covalent linkage was generated in samples of H46L/Q47L GlbN (1–5 mM, 100 mM Na/K phosphate pH 7–7.5) by overnight incubation with 2–5 mM DT, followed by oxidation using 5 mM $K_3[Fe(CN)_6]$ and passage over a 1 × 50 cm G-25 (Sigma) desalting column equilibrated in 10 mM Na/K phosphate pH 7. Formation of the cross-link was verified by mass spectrometry and comparison to previously published NMR data.[16]

NMR Spectroscopy.

NMR spectra were acquired at a field strength of 14.1 T using a Bruker Avance or Avance II spectrometer, each equipped with a TXI cryoprobe, or at a field strength of 18.8 T using a Varian INOVA spectrometer. ^{1}H chemical shifts were referenced indirectly through the residual water signal. ^{15}N and ^{13}C chemical shifts were referenced indirectly using the Ξ ratios.[26] Topspin 3.1 was used to process and analyze 1D data. Multidimensional data sets were processed with NMRPipe 3.0[27] and analyzed with Sparky3.[28] Select 3D correlation experiments were acquired using a Poisson-Gap sampling schedule in both indirect dimensions.[29] Non-uniformly sampled spectra were reconstructed using the NESTA- NMR software package implemented on the NMRbox virtual machine.[30,31]

Mesoheme ^{1}H assignments for ferric H46L mesoGlbN (0.5 mM, 20 mM Na/K phosphate pH* 7.3, 99% $^{2}H_2O$) and ferric H46L/Q47L mesoGlbN (2 mM, 30 mM Na/K phosphate pH* 7.2, 99% $^{2}H_2O$) were achieved using standard homonuclear experiments (NOESY, WEFT-NOESY, DQF-COSY, and TOCSY). ^{1}H Non-selective spin-lattice relaxation times (T_1) were determined with a simple inversion recovery sequence. ^{1}H, ^{15}N and ^{13}C assignment of H46L/Q47L GlbN (2 mM, 25 mM Na phosphate pH 7.2, 5% $^{2}H_2O$) and H46L/Q47L mesoGlbN (2.5 mM, 20 mM Na phosphate pH 7.1, 5% $^{2}H_2O$) were obtained using ^1H-^{15}N HSQC, ^1H-^{13}C HSQC, ^1H-^{15}N-^1H TOCSY-HSQC[32] (τ_{mix} = 45 ms), HNCO,[33] CBCA(CO)NH,[34] HNCACB,[35] HBHA(CO)NH,[34] and C(CO)NH[36] and H(CCO)NH−[6] (τ_{mix} = 15 ms) experiments. Amide NOEs were measured with ^1H-^{15}N-^1H NOESY-HSQC[32] spectra (τ_{mix} = 80 ms), and ^{13}C H selective ^1H-^{13}C HSQC and ^1H-^{13}C-^1H NOESY-HSQC[37] spectra (τ_{mix} = 70 ms) were acquired on H46L/Q47L mesoGlbN exchanged into 99% ^2H_2O.

RESULTS

Ruling Out Gln47 and Hydroxide Ligation.

The residues composing the hydrogen bond network of cyanomet GlbN (Figure 1B) are not far from the iron and could coordinate in the H46L variant with a modest distortion of the three-dimensional structure. To complement prior work on Y22F/H46L GlbN[16] and Q43L/H46L GlbN,[5] H46L/Q47L GlbN was prepared and compared to H46L GlbN. The electronic absorption spectra of the single and double variants at neutral pH (Figure 2) are nearly identical to one another, each displaying a Soret peak at 414 nm and lacking the long-wavelength charge transfer band of a high-spin complex. In the NMR spectra, hyperfine shifted ^1H signals are observed in the 30 to −20 ppm window only (Figure 3A and B). Non-selective T_1 values were determined for peaks tentatively assigned to the heme methyl groups (Figure S1) were found to be ~65 ms at the shortest. Both ^1H shifts and T_1 values are consistent with a principally low-spin iron.[38] Thus, barring additional ligand switching, the set of Tyr22, Gln43, and Gln47 variants rejects these residues from contention.

A structurally conservative interpretation of the ferric low-spin H46L GlbN data implicates hydroxide ions. His–Fe–OH^- complexes exist in low-spin/high-spin equilibria with a low-spin ground state.[38] The pK_a of the transition between hydroxymet and aquomet species varies, with values generally above 7.[39,40] Stabilization of the hydroxymet state over the high-spin aquomet state at neutral pH is occasionally observed, M80A iso-1-cytochrome *c* providing one example.[41,42] The pH titration of H46L GlbN monitored by electronic absorption spectroscopy[16] shows that the onset of the transition to an aquomet species occurs below pH 5, suggesting that, if hydroxide is the ligand, it is highly stabilized by the protein structure. Such enhanced stabilization of the hydroxymet species could be accomplished by hydrogen bond donation by Tyr22, as in the cyanomet complex, assisted by Gln43 and Gln47. However, the Y22F[16] and Q47L (Figure S2A) replacements in the H46L background do not change the pH response compared to H46L GlbN. If hydroxide coordination is indeed the source of the low-spin complex, some other stabilization mechanism must be at work.

The apparent pK_a of the coordinated water molecule depends in part on the heme electronic structure.[39,43] Replacement of the *b* heme in Mb with mesoheme (Figure 3E) raises the apparent pK_a by ~0.5 units by increasing the electron density at the iron.[40,44] To test the hypothesis of a hydroxide ligand further, the H46L and H46L/Q47L variants of GlbN were prepared with mesoheme. The electronic absorption spectra of H46L and H46L/Q47L mesoGlbN (Figure 2) match those of the parent proteins, after accounting for a 10-nm hypsochromic shift caused by a decrease in heme conjugation. The low-pH dependence of the electronic absorption spectra of H46L/Q47L mesoGlbN is similar to that of the heme *b* complex, showing protein denaturation and heme release with an apparent pK_a of ~4 (Figure

S2B). Decoupling of ligand protonation and acid denaturation is not observed, so that the pK_a of a bound hydroxide would have to be remarkably low to explain the result. The NMR relaxation and shift data also suggest a thermally inaccessible high-spin state. An isotope effect on chemical shift, expected in the event of strong hydrogen bonding,[45] is absent when comparing spectra collected in H_2O and ^2H_2O (not shown), which weakens this hydroxide hypothesis further. The alternative to hydroxide binding is that one strong-field ligand is provided by the protein. At the simplest, the same protein residue would coordinate the iron in both the ferric and ferrous states.

Identification of the Heme Ligands in H46L GlbN.

The ^{1}H NMR spectra of H46L mesoGlbN and H46L/Q47L mesoGlbN are shown in Figure 3. Reduced dispersion relative to the *b* heme counterparts is attributed to an increase in electronic distribution symmetry[46,47] and a possible alteration of the axial ligand orientation.[48] From a spectral analysis perspective, the loss of dispersion in mesoGlbN is compensated by sharpened resonances, and, of H46L mesoGlbN and H46L/Q47L mesoGlbN, the double variant presented higher quality data. This protein was therefore chosen for further study. Mesoheme assignments (Table S1) were achieved using 2D homonuclear data. Resolved mesoheme methyl resonances have T_1 values between 80 and 120 ms (Figure S1 and Table S1), shorter than those of *b* heme in WT GlbN (130–260 ms),[17] but nevertheless consistent with a low-spin iron.[49] Fast-relaxing nuclei belonging to the protein were also detected in 1D data as shown in Figure 4. For a ferric $S = 1/2$ porphyrin complex, neglecting spin delocalization, the T_1 of protons near the paramagnetic center is dominated by the dipolar mechanism with R^6 dependence on distance from the iron.[50] Crude distance estimates situate these protons within 6 Å of the iron, a distance appropriate for heme ligands.

To identify the fast-relaxing protons, H46L/Q47L apoGlbN was uniformly labeled with ^{15}N and ^{13}C and reconstituted with mesoheme. The ^{1}H-^{15}N HSQC spectrum of the resulting complex is shown in Figure S3 to illustrate data quality. Unambiguous sequential assignments from Leu4 to Gln124 (except prolines) was achieved with 3D triple-resonance experiments. Assignments for selected residues are given in Table S2. Of note is the ^{68}EAHKE string of shifted signals, which identifies His70. This residue has one efficiently relaxed CβH proton, and the $^{13}C^\alpha$ and $^{13}C^\beta$ signals are significantly shifted from the mean diamagnetic value for this type of residue (Table 1). The spectral characteristics of His70 condition expectations for the resonances of the distal ligand. Indeed, a second set of signals with shifts and relaxation properties similar to His70's is detected, embedded in the ^{116}AHKR sequence (Figure 5, Table 1): we conclude that His117 is the distal ligand in H46L/Q47L mesoGlbN. ^{1}H assignments for His70 and His117 were readily transferred to the H46L variant confirming that the Q47L replacement had no influence on the identity of the axial ligands.

If His117 is a ligand, ferric H46L/H117A GlbN is expected to bind the heme with the proximal histidine and form an aquomet (or hydroxymet) complex, as long as no further distortion of the fold and no additional ligand switching occurs. The ^{1}H NMR spectrum of the double variant acquired at neutral pH (Figure S4A) is consistent with a predominantly aquomet species (hyperfine shifts greater than 60 ppm). The electronic absorption spectra agree and, as the pH is raised, reveal an aquomet to hydroxymet transition with an apparent pK_a of ~8.6 (Figure S4B). In the ferrous state, the H117A replacement eliminates the spectral features of endogenous hexacoordination observed in H46L GlbN (Figure 6). These results are consistent with His117 being a heme ligand in both the ferric and ferrous oxidation states of H46L (meso)GlbN. In what follows, we refer to the conformation of the holoprotein with His70-Fe-His117 ligation as GlbN*.

Effects of Heme Substitution in H46L/Q47L GlbN.

Hyperfine ^{1}H chemical shifts are highly sensitive to the nature of the heme ligation (ligand identity, geometry, H-bonding, etc.). Comparison of data collected on GlbN* and mesoGlbN* reveals that the heme substitution in the H46L or H46L/Q47L variants does not alter the His70/His117 scheme (Table 1). However, the ^{1}H-^{15}N HSQC spectrum of H46L/Q47L GlbN* (Figure S5) shows two forms in a ~8:2 ratio, whereas the H46L/Q47L mesoGlbN* counterpart displays only one form. The population of two holoprotein forms conserving the same ligand set but accommodating the heme in two orientations related by a ~180° rotation about the α–γ meso axis is commonly observed in heme proteins.[38] This "heme rotational isomerism" has modest structural consequences and is modulated by both direct and remote heme–protein interactions.[51–53] The small differences in ^{1}H-^{15}N chemical shifts between the major and minor forms of H46L/Q47L GlbN* (Figure S6A) support that this heme rotational heterogeneity is the origin of NMR signal duplication.[54] Backbone amide assignments for the major heme *b* isomer closely match those of the mesoheme complex (Figure S6B). To a first approximation, structural information available from the high quality NMR spectra of H46L/Q47L mesoGlbN* holds for both the major isomer of the heme *b* complex and the H46L single variant. As expected because of changes in structure and in the paramagnetic susceptibility tensor, large chemical shift differences are observed between WT GlbN and H46L/Q47L GlbN* (Figure S6C).

Structural Properties of GlbN*.

The NMR model of WT GlbN (PDB ID:1MWB)[55] shows that His117 is positioned on the same face of the heme as the proximal histidine and points towards the heme in some conformers or away from it in others. Evidence for axial iron ligation by His70 and His117 in GlbN* implies not only a displacement of the heme but inevitably some structural reorganization of the protein. Circular dichroism was used to assess secondary structure perturbation. The far-UV CD spectra (Figure 7A) of GlbN*, with heme *b* or mesoheme, display signatures of high α-helical content consistent with expectations informed by the WT structure.[4] Larger differences are observed in the visible region of the spectrum (Figure 7B). In contrast to the negative Soret CD band of WT GlbN,[4] the double variants display positive Soret peaks, which may be related to different interactions with aromatic side chains and conformations of the heme propionates.[56]

CD data show similar helical content of GlbN and GlbN*, but offer limited insight into the structural changes that accommodate the unusual ligation scheme. To identify residue-specific distortions in secondary structure associated with His117 coordination, TALOS+ analysis[57] was applied to backbone ^{1}H, ^{13}C and ^{15}N chemical shifts of H46L/Q47L mesoGlbN* and WT GlbN (Figure 8). TALOS+ secondary structure predictions are supported by the detection of helical NOEs in the variant (diagrammed in Figure S7) and the WT protein.[58] There is substantial agreement between GlbN and GlbN* and between GlbN and its solution structure. Exceptions include residues 32–35, which TALOS+ classifies as a loop in GlbN but actually form the 3_{10} C helix in solution both in WT GlbN[58] and GlbN*, and residues in the vicinity of His117. Distortion of this region is an expected consequence of heme ligation by His117. Predicted order parameters, although to be taken with caution,[59] suggest that in H46L/Q47L mesoGlbN* the final turn of the H helix becomes a mobile random coil. Unfolding the C-terminus of GlbN creates a necessary opening between His70 and His117. Overall, however, the helices of GlbN are largely maintained in GlbN*, which explains the relatively unchanged far-UV CD spectrum.

The interfaces formed between helices in H46L/Q47L mesoGlbN* were probed with a $^{13}CH_3$-edited NOESY experiment, portions of which are shown in Figure S8. Many of the observed long-range NOEs (Table S3) could be predicted from the crystal structure of WT GlbN bound to cyanide (PDB ID: 1S69).[14] NOEs among Leu4, Leu8, Phe55, Leu104 and Val108 define the interfaces of the A, E and G helices. Hydrophobic residues on the B, E and G helices form the core of GlbN, and an extensive network involving Phe21, Tyr22, Val25, Leu47, Val87, and Leu91 indicates a conserved arrangement of these helices in the distal variant. Dipolar contacts between Leu92 and Ala109 (not shown) are consistent with the interface between the G helix and the N-terminal portion of the H helix.

In a flagrant departure from the WT structure, NOEs are observed between the E and F helices, arising from Gln43, Leu47, Phe50, Ala69, His70, and Leu73 (Figure 9). These helices are on opposite faces of the heme in all globins but are brought into close proximity in H46L/Q47L mesoGlbN*. These new contacts provide strong evidence for the displacement of the heme group. Also of note is the environment of Tyr22 (B10). This residue is in close proximity to Phe21, faces inwards toward the canonical distal heme pocket, and has an OηH proton in slow exchange with solvent (chemical shift time scale). NOEs to Val18, Val25, Leu47 and Leu73 describe a hydrophobic site likely responsible for the protection of the hydroxyl group (Figure S9). At the beginning of the G helix the upfield shift of Asn80 NH in both ^{1}H and ^{15}N and low protection factor determined by hydrogen-deuterium exchange (not shown) demonstrate perturbation of the conserved G helix N-cap.[60] This perturbation provides a convenient signature for GlbN*.

Repositioning the heme within GlbN is accompanied with the creation of an E–F interface and, as mentioned above, involves a separation of the F and H helices. Residues forming the binding site in H46L/Q47L mesoGlbN* were identified with NOEs between the polypeptide and the mesoheme. In WT GlbN, the B and C pyrroles are buried among hydrophobic residues in the core of the protein. Contacts between this portion of mesoheme and Phe50, Leu79, Phe84 and Val87 are observed in GlbN* and highlight common features of the two conformations. In contrast, the A and D pyrroles normally fit between the E and F helices to face solvent but, in GlbN*, Leu122 is found in proximity to the 5-CH_3. Also informative are contacts between Phe50 and His70, and His117 and Phe84. A few key protein–mesoheme NOEs are shown in Figure S10 and listed in Table S2.

In summary, replacement of His46 in GlbN causes the heme to pivot about the pyrrole B/C half of the heme molecule and to slide between the F and H helices. The F helix packs against the E helix, filling the void created by heme displacement, while the C-terminus of the protein unfolds, opening the needed gap between His70 and His117. The heme propionates remain exposed to solvent. The overall globin fold and the core formed by hydrophobic residues in the B, E and G helices are maintained in the new structure, schematized in Figure 10.

Heme Release from Ferric H46L GlbN*.

A change in heme binding site is expected to affect heme affinity. Monitoring the transfer of ferric *b* heme from GlbN to apoMb is a convenient means of determining the dissociation rate of the heme from the holoprotein.[62] Addition of excess apoMb to a sample of H46L GlbN* resulted in conversion of the optical spectrum to that of aquomet Mb (Figure 11). Singular value decomposition of the spectra and global analysis display monophasic kinetics with an apparent rate constant of $3.1 \pm 0.1 \times 10^{-3}$ s^{-1}. The mechanistic details (e.g., relation to rotational isomer population) are unclear, but the observed rate is ~1000-fold faster than that observed for the major kinetic phase of WT GlbN under similar conditions.[4] Assuming that the association rate constant, k_{on}, is unchanged,[63] the increased off-rate represents a lower heme affinity in the variant. Consistent with a loss of binding energy and stability, a decrease in T_m of ~10 °C was previously reported for H46L GlbN,[16] and raising hydrostatic pressure causes a loss of heme resonances and spectral resolution relatively early[64] in the pressure titration (Figure S11).

Heme Crosslinking in Ferrous GlbN Variants.

GlbN is remarkable for forming a covalent linkage between the heme and His117[24] (Figure S12A). Recombinant GlbN is prepared in the ferric state, and addition of a reducing agent, typically DT, causes complete formation of the bond on the time-scale of milliseconds at neutral pH.[65] H46L GlbN* undergoes the same post- translational modification (PTM) but is less reactive.[16] Prolonged incubation of H46L/Q47L GlbN* with DT followed by reoxidation and removal of DT byproducts revealed partial formation of the heme–protein adduct, which displays high-spin features consistent with the loss of His117 iron coordination (Figure S12B). Crosslinking requires the heme to adopt the correct orientation with respect to His117, the sliding of the heme back to its canonical site perhaps contributing to the decreased reaction rate in the H46 variants.

The ability of His117 to both coordinate the ferrous iron and undergo irreversible covalent modification was explored with heme transfer experiments and by making use of mesoGlbN. Because vinyl groups are absent from mesoheme, mesoGlbN complexes provide access to the ferrous state without PTM. The optical spectrum of ferrous H46L/Q47L mesoGlbN displays two resolved Soret peaks, one presumably corresponding to a high-spin form with a maximum at ~426 nm and the other to a low-spin form with maximum at ~415 nm (Figure 12A). The mixture represents partial His117 coordination, and the appearance of the NMR spectrum (not shown) demonstrates that interconversion between low- and high-spin species occurs rapidly on the chemical shift time scale. The coexistence of GlbN* and GlbN illustrates that *bis*-histidine coordination in GlbN* is weaker in the ferrous state than in the ferric state.[66,67]

To inspect the kinetics of PTM formation, H46L GlbN containing a *b* heme was placed in the presence of apoMb and immediately reduced by DT (Figure 12B). In the absence of apoMb, freshly reduced H46L GlbN shows spectral changes over the course of hours (not shown). When excess apoMb is available, nearly complete loss of the ferrous heme occurred in the dead-time of the manual mixing experiment. For comparison, reduced H117A GlbN releases heme with a half-life of ~1 h,[67] and WT GlbN forms the PTM immediately.[16] Thus, ferrous H46L GlbN is a dynamic species that exhibits transient His117 ligation (GlbN*), gradual PTM formation, and rapid dissociation of the heme group.

DISCUSSION

Factors Allowing His117 Coordination in H46L GlbN.

Positioning the heme between the F and H helices requires conformational adjustments to the WT structure (Figure 9). The disordered EF loop undergoes a large structural change when WT GlbN binds an exogenous ligand (Figure 1);[14] the same flexible loop allows the F helix to form a new interface in mesoGlbN*. In the H-helix beyond Ala112, rapid backbone NH/ND exchange[55] indicates a propensity for structure opening. This region also participates in the rearrangement. Although the intrinsic plasticity of the EF loop and low conformational stability at the end of the H helix were known features of GlbN expected to facilitate a low-energy deformation of the structure, the perturbation of a conserved helix capping H-bond at the beginning of the G helix[60] was not. Formation of the Fe–His117 coordination bond apparently offsets the cumulative losses in conformational stability associated with the distortion of the canonical holoprotein structure.

Further insight can be derived with cyanide binding to GlbN*. As observed in prior work, the NMR spectrum of cyanomet H46L GlbN resembles closely that of WT GlbN,[16] including evidence for Tyr22–cyanide interaction. This implies that cyanide displaces His117 from the iron and restores the native (cyanomet) GlbN structure. In support of this interpretation, binding of cyanide, imidazole or azide to H46L/Q47L GlbN restores the helix capping H-bond between Asn80 and His8- (Figure S13A). Subsequent removal of imidazole or azide by extensive buffer exchange yields the starting GlbN* complex (Figure S13B), demonstrating reversible movement of the heme within the protein matrix. A plausible interpretation is that simultaneous coordination of His117 and His70 propagates strains from each Fe–Nε2 linkage (or Nδ1 linkage, a possibility that is not eliminated by the available data) to the backbone, strains that are compensated by ligation. Once an exogenous ligand binds, the H helix is released, strains on the proximal side are relieved, and the heme migrates to the WT, energetically more favorable site lined with hydrophobic residues and consolidated by the hydrogen bond network composed of Tyr22, Gln43, and Gln47.

A related GlbN from the cyanobacterium *Synechococcus* PCC sp. 7002 provides additional information on the determinants of His117 ligation. *Synechococcus* GlbN also coordinates the heme iron with His70 and His46 and is capable of the same heme–His117 PTM.[68] The H46L variant of *Synechococcus* GlbN, however, forms an aquomet complex in the ferric state with the heme presumably in its normal site.[65] Although it is not possible to infer binding energies from structure, striking differences in the H helices of the two proteins are likely relevant to the relative affinities of the alternative binding site. In *Synechocystis* GlbN the sequence directly upstream of His117 is rich in alanines (^{109}AAVA<u>G</u>APA, Gly113 underlined) whereas in *Synechococcus* GlbN, it is proline-less and has several β-branched residues, (^{109}VTIV<u>G</u>SVQ). On the basis of the sequence alone, helical propensity in this region is higher in *Synechocystis* GlbN than *Synechococcus* GlbN, which may contribute to a favorable orientation of His117 for coordination.

A survey of sequences related to GlbN mapped onto known three-dimensional structures reveal that most proteins have a two-residue deletion at positions 114–115. The sequences containing a histidine at GlbN's positions 117 do not have this deletion and fall roughly in two categories, either the 109–116 segment is Ala-rich or not. Furthermore, His117 is systematically found in combination with His46, except in *Microcystis aeruginosa* HbN (WP_004161553), which harbors the Tyr46/His117 pair and the ^{109}VQIV<u>G</u>SVT sequence. In GlbN*, the mesoheme contacts Ala112. Val112, in *Synechococcus* GlbN and presumably other proteins containing this residue, forms a hydrophobic core with the conserved Phe50 (E14) and Phe84 (G5) along with Val109 and Ile111.[69] Compared to those afforded by Ala112, these interactions may be sufficient to counteract the energetic benefit of His117 ligation. Hypotheses such as these can be tested with a combination of mutagenesis and NMR characterization to refine the understanding of heme site preference.

Relation to Apoprotein Properties.

The thermodynamic view presented above draws attention to the apoprotein as an agent guiding the preference for GlbN or GlbN*. It is reasonable to assume that an initial state endowed with pre-existing, stable and distinctive holoGlbN features would be less likely to populate the holoGlbN* conformation. Conversely, an unstable apoprotein may be more permissive. In *Synechococcus* apoGlbN, portions of the B, G and H helices form a core that unfolds cooperatively on heating.[70] The ^{1}H-^{15}N HSQC spectrum is resolved, and chemical shifts suggest that the N-terminal portion of the H-helix is folded. In contrast, *Synechocystis* apoGlbN, which has much reduced helical content compared to the holoprotein[4] and no native baseline in urea denaturation experiments,[71] has a poorly resolved ^{15}N-^{1}H HSQC spectrum, a property shared by H46L/Q47L apoGlbN (Figure S14).

Apoprotein stability and folding mechanism have not yet been extensively explored in the 2-on-2 globin lineage. To condition expectations, it is useful to draw a parallel to 3-on-3 globins. In sperm whale apomyoglobin the A, B, E helices and parts of the G and H helices pack to form a hydrophobic core, whereas the E/F loop, F helix, F/G turn, and C-terminal end of the H helix are dynamic elements.[72–75] In general, 3-on-3 apoglobins display a variety of folding pathways and intermediate population to reach their native state.[76–79] The stability of this state is known to depend on the amino acid make-up of the core and heme binding site.[80–82] However, even highly destabilized 3-on-3 apoproteins can refold into the expected holoprotein structure on heme binding.[80] The truncated A helix and shorter H helix of the group 1 2-on-2 fold eliminate several of the interactions that are observed in the 3-on-3 proteins. Among 2-on-2 globins the extent of the hydrophobic core, e.g., the participation of Val109 and Val112 in *Synechococcus* GlbN, varies and likely contributes to both the structural stability of the apoprotein and the accessibility of alternative holoprotein conformations. Globins that have low level of apoprotein structure, such as *Synechocystis* GlbN, may have a higher propensity for non-canonical heme binding, although this initial state property is clearly not a sufficient condition. Additional data factoring in final state characteristics are needed to probe such a connection.

Distal Ligation Competition.

Dioxygen binding is a defining feature of globins, whether for transport, storage, or chemistry. Competition between endogenous and exogenous coordination is a well-known property of most “hexacoordinate” hemoglobins.[8] The persistence of a non-native *bis*-histidine structure upon the replacement of a single amino acid in GlbN highlights the facility with which a heme group can associate with nitrogenous ligands and lead to internal competition for the iron. His/Lys ligand switching, for example, can be driven by altering solution conditions, most notably pH as shown in cytochrome *c*[8–,84] and *Synechococcus* GlbN.[85] Internal ligand exchange is observed in other guises as well, illustrated in the signaling mechanism of CooA[86] and the heme transfer mechanism of PhuS.[87,88] Non-specific or transient binding is also a hallmark of heme–protein interactions.[63,89,90] Most often, the architecture and flexibility of the protein create a unique heme binding site with a constant set of ligands. In a sense, *Synechococcus* and *Synechocystis* GlbNs illustrate a gradation in the strength of “non- specific” binding, with *Synechocystis* H46L GlbN an extreme example of competition and stabilization.

Reconciliation of H46 Variant Observations.

In light of alternative ligation, the role of the interfering His117 should be scrutinized anew. Covalent attachment of the heme to His117 by addition to a vinyl is observed in *Synechococcus* cells[91] and presumably other cyanobacterial HbNs. The PTM offers the advantage of heme retention, which may be important under some cellular conditions. We have proposed that His46 drives the heme to the correct site for PTM.[16] We now show that free His117 has a strong tendency to coordinate the heme iron and reinforce our structural explanation for the requirement of His46. PTM and rogue ligation may also explain the rare occurrence of a histidine at the 117 topological position. One possible role for His117 besides PTM is to assist apoGlbN in the acquisition of free heme. To return to an extensively studied system, folding of apoMb occurs through a three-state mechanism and the unfolded, intermediate and native states display progressively increasing heme affinities.[10] Partially folded apoMb incorporates heme to give a *bis*-histidine hemichrome that readily converts to the structured holoprotein, illustrating how metal ions and coordination chemistry can shape protein folding pathways. On-pathway intermediates must undergo ligand exchange to produce the folded protein efficiently. That transient His117 coordination is observed in ferrous H46L GlbN may speak to this point. Gradual formation of the covalent heme modification in this protein demonstrates that the *bis*-histidine GlbN* complex is capable of converting to GlbN. His117, attached to a terminal and flexible part of the protein, could offer a means to recruit free ferrous heme to apoGlbN by coordinating the iron opposite His70, rapidly exchanging with His46 to generate the native state, and securing the cofactor with a covalent linkage.

The unusual structure of GlbN* clarifies other observations made on H46 variants. A greater degree of high-spin character in ferric H46A GlbN,[7,16] as compared to ferric H46L GlbN, is ascribable to a competition between native GlbN and GlbN* conformations in the variants. Leu46 contributes to the hydrophobic core of GlbN* while Ala46 could promote the GlbN structure by facilitating the coordination of a solvent molecule on the distal iron site. CO binding kinetics have been measured for H46A GlbN.[5] Discrepancies between results from rapid mixing and flash photolysis experiments can be explained by movement of the cofactor within the polypeptide. A slow phase (~0.01 s^{-1}, [CO] independent) observed by rapid mixing may reflect migration of the cofactor back to the standard heme pocket.

Perspective on Globin Mutagenesis and Homology Modeling.

Beyond the novelty of the GlbN* structure, our results offer some generic perspective on heme binding. They emphasize that a single amino acid replacement in the heme cavity can have hard-to-detect but profound consequences. It is customary to replace a suspected heme axial ligand to confirm a coordination scheme or pursue thermodynamic and kinetic studies of hexacoordination, but as shown here, data collected with such variants may reflect much larger perturbations than breaking a bond and allowing solvent in the heme pocket. The complexity of heme protein analysis is well illustrated with GlbN, which displays heme orientational disorder, PTM, and uniquely-behaved distal variants. GlbN* also highlight the potential for distortion of a holoprotein without extraordinary global reshaping. The results call into question the validity of simple homology modeling or computation resting solely on wild-type structural assumptions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Dr. Katherine Tripp for help with CD data collection, Dr. Ananya Majumdar of the JHU Biomolecular NMR Center for help with NMR data collection, Dr. Matthew Preimesberger for helpful discussions, Dr. Christopher Falzone for careful reading of the manuscript, and Christos Kougentakis for preliminary experiments and assistance with non-uniformly sampled NMR data.

FUNDING SOURCES

This work was supported by the National Science Foundation grant MCB-1330488 to JTJL and National Institutes of Health grant T32 GM080189 (DBN).

Abbreviations:

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
apoMb	horse skeletal muscle apomyoglobin
CD	circular dichroism

DT	sodium dithionite
GlbN	*Synechocystis* sp. PCC 6803 hemoglobin
GlbN-A	GlbN with covalently attached heme
GlbN*	GlbN with His70 and His117 as heme axial ligands
heme *b*	Fe-protoporphyrin IX
HSQC	, heteronuclear single quantum coherence
Mb	myoglobin
mesoGlbN	GlbN reconstituted with mesoheme
mesoheme	Fe-mesoporphyrin IX
met	ferric
NOE	nuclear Overhauser effect
PDB	Protein Data Bank
pH*	pH uncorrected for isotope effect
PTM	post-translational modification
rmsd	root-mean-square deviation
T_1	spin-lattice (longitudinal) relaxation time
WT	wild-type

References

(1). Wajcman H, Kiger L, and Marden MC (2009) Structure and function evolution in the superfamily of globins. C. R. Biologies 332, 273–282. [PubMed: 19281958]

(2). Vinogradov SN, Tinajero-Trejo M, Poole RK, and Hoogewijs D (2013) Bacterial and archaeal globins — A revised perspective. Biochim. Biophys. Acta 1834, 1789–1800. [PubMed: 23541529]

(3). Gardner PR (2012) Hemoglobin: A nitric-oxide dioxygenase. Scientifica 2012, 34.

(4). Lecomte JTJ, Scott NL, Vu BC, and Falzone CJ (2001) Binding of ferric heme by the recombinant globin from the cyanobacterium *Synechocystis* sp. PCC 6803. Biochemistry 40, 6541–6552. [PubMed: 11371218]

(5). Hvitved AN, Trent JT, 3rd, Premer SA, and Hargrove MS (2001) Ligand binding and hexacoordination in *Synechocystis* hemoglobin. J. Biol. Chem. 276, 34714–34721. [PubMed: 11438545]

(6). Das TK, Couture M, Ouellet Y, Guertin M, and Rousseau DL (2001) Simultaneous observation of the O–O and Fe–O_2 stretching modes in oxyhemoglobins. Proc. Natl. Acad. Sci. U. S. A. 98, 479–484. [PubMed: 11209051]

(7). Couture M, Das TK, Savard PY, Ouellet Y, Wittenberg JB, Wittenberg BA, Rousseau DL, and Guertin M (2000) Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC 680- reveal a unique distal heme pocket. Eur. J. Biochem. 267, 4770–4780. [PubMed: 10903511]

(8). Kakar S, Hoffman FG, Storz JF, Fabian M, and Hargrove MS (2010) Structure and reactivity of hexacoordinate hemoglobins. Biophys. Chem. 152, 1–14. [PubMed: 20933319]

(9). Rifkind JM, Abugo O, Levy A, and Heim J (1994) Detection, formation, and relevance of hemichromes and hemochromes. Methods Enzymol. 231, 449–480. [PubMed: 8041268]

(10). Culbertson DS, and Olson JS (2010) Role of heme in the unfolding and assembly of myoglobin. Biochemistry 49, 6052–6063. [PubMed: 20540498]

(11). Mollan TL, Khandros E, Weiss MJ, and Olson JS (2012) Kinetics of α-globin binding to a-hemoglobin stabilizing protein (AHSP) indicate preferential stabilization of hemichrome folding intermediate. J. Biol. Chem. 287, 11338–11350. [PubMed: 22298770]

(12). Halder P, Trent JT, 3rd, and Hargrove MS (2007) Influence of the protein matrix on intramolecular histidine ligation in ferric and ferrous hexacoordinate hemoglobins. Proteins 66, 172–182. [PubMed: 17044063]

(13). Hoy JA, Kundu S, Trent JT, 3rd, Ramaswamy S, and Hargrove MS (2004) The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. J. Biol. Chem. 279, 165–5–16542.

(14). Trent JT, 3rd, Kundu S, Hoy JA, and Hargrove MS (2004) Crystallographic analysis of *Synechocystis* cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. J. Mol. Biol. 341, 1097–1108. [PubMed: 15289104]

(15). Lecomte JTJ, Vu BC, and Falzone CJ (2005) Structural and dynamic properties of *Synechocystis* sp. PCC 6803 Hb revealed by reconstitution with Zn-protoporphyrin IX. J. Inorg. Biochem. 99, 1585–1592. [PubMed: 15961161]

(16). Nothnagel HJ, Love N, and Lecomte JTJ (2009) The role of the heme distal ligand in the post-translational modification of *Synechocystis* hemoglobin. J. Inorg. Biochem. 103, 107–116. [PubMed: 18992944]

(17). Scott NL, and Lecomte JTJ (2000) Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803. Protein Sci. 9, 587–597. [PubMed: 10752621]

(18). Teale FWJ (1959) Cleavage of heme-protein link by acid methylethylketone. Biochim. Biophys. Acta 35, 543. [PubMed: 13837237]

(19). Johnson EA, Rice SL, Preimesberger MR, Nye DB, Gilevicius L, Wenke BB, Brown JM, Witman GB, and Lecomte JTJ (2014) Characterization of THB1, a *Chlamydomonas reinhardtii* truncated hemoglobin: linkage to nitrogen metabolism and identification of lysine as the distal heme ligand. Biochemistry 53, 4573–4589. [PubMed: 24964018]

(20). de Duve C (1948) A spectrophotometric method for the simultaneous determination of myoglobin and hemoglobin in extracts of human muscle. Acta Chem. Scan. 2, 264–289.

(21). Antonini E, Brunori M, Caputo A, Chiancone E, Fanelli AR, and Wyman J (1964) Studies on the structure of hemoglobin: III. Physiochemical properties of reconstituted hemoglobins. Biochim. Biophys. Acta 79, 284–292. [PubMed: 14165019]

(22). Gill SC, and von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182, 319–326. [PubMed: 2610349]

(23). Hargrove MS, Singleton EW, Quillin ML, Ortiz LA, Phillips GN, Jr., Olson JS, and Mathews AJ (1994) His64(E7)-->Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation. J. Biol. Chem. 269, 4207–4214. [PubMed: 8307983]

(24). Vu BC, Jones AD, and Lecomte JTJ (2002) Novel histidine-heme covalent linkage in a hemoglobin. J. Am. Chem. Soc. 124, 8544–8545. [PubMed: 12121092]

(25). Preimesberger MR, Pond MP, Majumdar A, and Lecomte JTJ (2012) Electron self-exchange and self-amplified posttranslational modification in the hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. J. Biol. Inorg. Chem. 17, 599–609. [PubMed: 22349976]

(26). Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, and Sykes BD (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J. Biomol. NMR 6, 1–5–140. [PubMed: 22911575]

(27). Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, and Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293. [PubMed: 8520220]

(28). Goddard TD, and Kneller DG (2006) SPARKY 3. University of California, San Francisco.

(29). Hyberts SG, Takeuchi K, and Wagner G (2010) Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. J. Am. Chem. Soc. 132, 2145–2147. [PubMed: 20121194]

(30). Sun S, Gill M, Li Y, Huang M, and Byrd RA (2015) Efficient and generalized processing of multidimensional NUS NMR data: the NESTA algorithm and comparison of regularization terms. J. Biomol. NMR 62, 105–117. [PubMed: 25808220]

(31). Maciejewski MW, Schuyler AD, Gryk MR, Moraru II, Romero PR, Ulrich EL, Eghbalnia HR, Livny M, Delaglio F, and Hoch JC (2017) NMRbox: A resource for biomolecular NMR computation. Biophys. J. 112, 1529–1534. [PubMed: 28445744]

(32). Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, and Clore GM (1989) Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann- Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β. Biochemistry 28, 6150–6156. [PubMed: 2675964]

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

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

(35). Wittekind M, and Mueller L (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. J. Magn. Reson. B 101, 201–205.

(36). Montelione GT, Lyons BA, Emerson SD, and Tashiro M (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence- specific resonance assignments of isotopically-enriched proteins. J. Am. Chem. Soc. 114, 10974–10975.

(37). Bax A, and Grzesiek S (1993) Methodological Advances in Protein NMR, In NMR of Proteins (Clore GM, and Gronenborn AM, Eds.), pp 33–52, Macmillan Education UK, London.

(38). La Mar GN, Satterlee JD, and de Ropp JS (2000) Nuclear magnetic resonance of hemoproteins, In The Porphyrin Handbook (Smith KM, Kadish K, and Guilard R, Eds.), pp 185–298, Academic Press, Burlington, MA.

(39). Antonini E, and Brunori M (1971) Hemoglobin and myoglobin in their reactions with ligands, Vol. 12, North-Holland, Amsterdam.

(40). Nagao S, Hirai Y, Suzuki A, and Yamamoto Y (2005) ^{19}F NMR characterization of the thermodynamics and dynamics of the acid−alkaline transition in a reconstituted sperm whale metmyoglobin. J. Am. Chem. Soc. 127, 4146–4147. [PubMed: 15783177]

(41). Bren KL, and Gray HB (1993) Structurally engineered cytochromes with novel ligand-binding sites: oxy and carbon monoxy derivatives of semisynthetic horse heart Ala80 cytochrome *c*. J. Am. Chem. Soc. 115, 10382–10383.

(42). Banci L, Bertini I, Bren KL, Gray HB, and Turano P (1995) pH-dependent equilibria of yeast Met80Ala-iso-1-cytochrome *c* probed by NMR spectroscopy: a comparison with the wild-type protein. Chem. Biol. 2, 377–383. [PubMed: 9383439]

(43). McGrath TM, and La Mar GN (1978) Proton NMR study of the thermodynamics and kinetics of the acid in equilibrium base transitions in reconstituted metmyoglobins. Biochim. Biophys. Acta 534, 99–111. [PubMed: 26418]

(44). Shibata T, Nagao S, Fukaya M, Tai H, Nagatomo S, Morihashi K, Matsuo T, Hirota S, Suzuki A, Imai K, and Yamamoto Y (2010) Effect of heme modification on oxygen affinity of myoglobin and equilibrium of the acid-alkaline transition in metmyoglobin. J Am Chem Soc 132, 6091–6098. [PubMed: 20392104]

(45). Qin J, La Mar GN, Dou Y, Admiraal SJ, and Ikeda-Saito M (1994) 1H NMR study of the solution molecular and electronic structure of engineered distal myoglobin His64(E7) Val/Val68(E11) His double mutant. Coordination of His64(E11) at the sixth position in both low-spin and high-spin states. J. Biol. Chem. 269, 1083–1090. [PubMed: 8288565]

(46). La Mar GN, Viscio DB, Smith KM, Caughey WS, and Smith ML (1978) NMR studies of low-spin ferric complexes of natural porphyrin derivatives. 1. Effect of peripheral substituents on the π electronic asymmetry in biscyano complexes. J. Am. Chem. Soc. 100, 8085–8092.

(47). Hirai Y, Yamamoto Y, and Suzuki A (2000) ^{19}F NMR study of the heme orientation and electronic structure in a myoglobin reconstituted with a ring-fluorinated heme. Bull. Chem. Soc. Jpn. 73, 2309–2316.

(48). Bertini I, Luchinat C, Parigi G, and Walker FA (1999) Heme methyl ^{1}H chemical shifts as structural parameters in some low-spin ferriheme proteins. J. Biol. Inorg. Chem. 4, 515–519. [PubMed: 10555585]

(49). La Mar GN, and Walker FA, (Eds.) (1979) Nuclear magnetic resonance of paramagnetic metalloporphyrins, Vol. 4, Academic Press, New York.

(50). Unger SW, Jue T, and La Mar GN (1985) Proton NMR dipolar relaxation by delocalized spin density in low-spin ferric porphyrin complexes. J. Magn. Reson. 61, 448–456.

(51). Mortuza GB, and Whitford D (1997) Mutagenesis of residues 27 and 78 modulates heme orientation in cytochrome b_5. FEBS Lett. 412, 610–614. [PubMed: 9276476]

(52). Cocco MJ, Barrick D, Taylor SV, and Lecomte JTJ (1992) Histidine 82 influences heme orientational isomerism in sperm whale myoglobin. Long-range effect due to mutation of a conserved residue. J. Am. Chem. Soc. 114, 11000–11001.

(53). Yang F, Zhang H, and Knipp M (2009) A one-residue switch reverses the orientation of a heme b cofactor. Investigations of the ferriheme NO transporters nitrophorin 2 and 7 from the blood-feeding insect *Rhodnius prolixus*. Biochemistry 48, 235–241. [PubMed: 19140692]

(54). Sarma S, DiGate RJ, Banville DL, and Guiles RD (1996) 1H, 13C and 15N NMR assignments and secondary structure of the paramagnetic form of rat cytochrome b5. J. Biomol. NMR 8, 171–183. [PubMed: 8914273]

(55). Falzone CJ, Vu BC, Scott NL, and Lecomte JTJ (2002) The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state. J. Mol. Biol. 324, 1015–1029. [PubMed: 12470956]

(56). Nagai M, Kobayashi C, Nagai Y, Imai K, Mizusawa N, Sakurai H, Neya S, Kayanuma M, Shoji M, and Nagatomo S (2015) Involvement of propionate side chains of the heme in circular dichroism of myoglobin: experimental and theoretical analyses. J. Phys. Chem. B 119, 1275–1287. [PubMed: 25525834]

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

(58). Falzone CJ, and Lecomte JTJ (2002) Assignment of the 1H, 13C, and 15N signals of Synechocystis sp. PCC 680- methemoglobin. J. Biomol. NMR 2-, 71–72.

(59). Berjanskii MV, and Wishart DS (2005) A simple method to predict protein flexibility using secondary chemical shifts. J. Am. Chem. Soc. 127, 14970–14971. [PubMed: 16248604]

(60). Preimesberger MR, Majumdar A, Rice SL, Que L, and Lecomte JTJ (2015) Helix-capping histidines: Diversity of N-H...N hydrogen bond strength revealed by $^{2h}J_{NN}$ scalar couplings. Biochemistry 54, 6896–6908. [PubMed: 26523621]

(61). Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004) UCSF Chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612. [PubMed: 15264254]

(62). Hargrove MS, and Olson JS (1996) The stability of holomyoglobin is determined by heme affinity. Biochemistry 35, 11310–11318. [PubMed: 8784185]

(63). Hargrove MS, Barrick D, and Olson JS (1996) The association rate constant for heme binding to globin is independent of protein structure. Biochemistry 35, 11293–11299. [PubMed: 8784183]

(64). Dellarole M, Roumestand C, Royer C, and Lecomte JTJ (2013) Volumetric properties underlying ligand binding in a monomeric hemoglobin: A high-pressure NMR study. Biochim. Biophys. Acta 1834, 1910–1922. [PubMed: 23619242]

(65). Nothnagel HJ, Preimesberger MR, Pond MP, Winer BY, Adney EM, and Lecomte JTJ (2011) Chemical reactivity of *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 hemoglobins: covalent heme attachment and bishistidine coordination. J. Biol. Inorg. Chem. 16, 539–552. [PubMed: 21240532]

(66). Cowley AB, Kennedy ML, Silchenko S, Lukat-Rodgers GS, Rodgers KR, and Benson DR (2006) Insight into heme protein redox potential control and functional aspects of six-coordinate ligand-sensing heme proteins from studies of synthetic heme peptides. Inorg. Chem. 45, 9985–10001. [PubMed: 17140194]

(67). Hoy JA, Smagghe BJ, Halder P, and Hargrove MS (2007) Covalent heme attachment in *Synechocystis* hemoglobin is required to prevent ferrous heme dissociation. Protein Sci. 16, 250–260. [PubMed: 17242429]

(68). Scott NL, Falzone CJ, Vuletich DA, Zhao J, Bryant DA, and Lecomte JTJ (2002) The hemoglobin of the cyanobacterium *Synechococcus* sp. PCC 7002: Evidence for hexacoordination and covalent adduct formation in the ferric recombinant protein. Biochemistry 41, 6902–6910. [PubMed: 12033922]

(69). Wenke BB, Lecomte JTJ, Héroux A, and Schlessman JL (2014) The 2/2 hemoglobin from the cyanobacterium *Synechococcus* sp. PCC 7002 with covalently attached heme: comparison of X-ray and NMR structures. Proteins 82, 528–534. [PubMed: 23999883]

(70). Vuletich DA, Falzone CJ, and Lecomte JTJ (2006) Structural and dynamic repercussions of heme binding and heme-protein cross-linking in *Synechococcus* sp. PCC 7002 hemoglobin. Biochemistry 45, 14075–14084. [PubMed: 17115702]

(71). Knappenberger JA, Kuriakose SA, Vu BC, Nothnagel HJ, Vuletich DA, and Lecomte JTJ (2006) Proximal influences in two-on-two globins: Effect of the Ala69Ser replacement on *Synechocystis* sp. PCC 6803 hemoglobin. Biochemistry 45, 11401–11413. [PubMed: 16981700]

(72). Cocco MJ, and Lecomte JTJ (1990) Characterization of hydrophobic cores in apomyoglobin: a proton NMR spectroscopy study. Biochemistry 29, 11067–11072. [PubMed: 2176892]

(73). Hughson FM, Wright PE, and Baldwin RL (1990) Structural characterization of a partly folded apomyoglobin intermediate. Science 249, 1544–1548. [PubMed: 2218495]

(74). Lecomte JTJ, Sukits SF, Bhattacharya S, and Falzone CJ (1999) Conformational properties of native sperm whale apomyoglobin in solution. Protein Sci. 8, 1484–1491. [PubMed: 10422837]

(75). Landfried DA, Vuletich DA, Pond MP, and Lecomte JTJ (2007) Structural and thermodynamic consequences of *b* heme binding for monomeric apoglobins and other apoproteins. Gene 398, 12–28. [PubMed: 17550789]

(76). Nishimura C, Lietzow MA, Dyson HJ, and Wright PE (2005) Sequence determinants of a protein folding pathway. J. Mol. Biol. 351, 383–392. [PubMed: 16005892]

(77). Eun Y-J, Kurt N, Sekhar A, and Cavagnero S (2008) Thermodynamic and kinetic characterization of apoHmpH, a fast-folding bacterial globin. J. Mol. Biol. 376, 879–897. [PubMed: 18187151]

(78). Zhu L, Kurt N, Choi J, Lapidus LJ, and Cavagnero S (2013) Sub-millisecond chain collapse of the *Escherichia coli* globin apoHmpH. J. Phys. Chem. B 117, 7868–7877. [PubMed: 23750553]

(79). Codutti L, Picotti P, Marin O, Dewilde S, Fogolari F, Corazza A, Viglino P, Moens L, Esposito G, and Fontana A (2009) Conformational stability of neuroglobin helix F--possible effects on the folding pathway within the globin family. FEBS J. 276, 5177–5190. [PubMed: 19674102]

(80). Hargrove MS, Krzywda S, Wilkinson AJ, Dou Y, Ikeda-Saito M, and Olson JS (1994) Stability of myoglobin: a model for the folding of heme proteins. Biochemistry 33, 11767–11775. [PubMed: 7918393]

(81). Smith LP (2003) The Effects of Amino Acid Substitution on Apomyoglobin Stability, Folding Intermediates, and Holoprotein Expression. Ph.D. Thesis, Rice University, Houston, TX.

(82). Samuel PP, Smith LP, Phillips GN, Jr., and Olson JS (2015) Apoglobin stability is the major factor governing both cell-free and in vivo expression of holomyoglobin. J. Biol. Chem. 290, 23479–23495. [PubMed: 26205820]

(83). Ferrer JC, Guillemette JG, Bogumil R, Inglis SC, Smith M, and Mauk AG (1993) Identification of Lys79 as an iron ligand in one form of alkaline yeast iso-1- ferricytochrome *c*. J. Am. Chem. Soc. 115, 7507–7508.

(84). Rosell FI, Ferrer JC, and Mauk AG (1998) Proton-linked protein conformational switching: Definition of the alkaline conformational transition of yeast iso-1- ferricytochrome *c*. J. Am. Chem. Soc. 120, 11234–11245.

(85). Nye DB, Preimesberger MR, Majumdar A, and Lecomte JTJ (2018) Histidine–lysine axial ligand switching in a hemoglobin: A role for heme propionates. Biochemistry 57, 631–644. [PubMed: 29271191]

(86). Aono S, Ohkubo K, Matsuo T, and Nakajima H (1998) Redox-controlled ligand exchange of the heme in the CO-sensing transcriptional activator CooA. J. Biol. Chem. 273, 25757–25764. [PubMed: 9748246]

(87). Block DR, Lukat-Rodgers GS, Rodgers KR, Wilks A, Bhakta MN, and Lansky IB (2007) Identification of two heme-binding sites in the cytoplasmic heme- trafficking protein PhuS from *Pseudomonas aeruginosa* and their relevance to function. Biochemistry 46, 14391–14402. [PubMed: 18020455]

(88). Tripathi S, O'Neill MJ, Wilks A, and Poulos TL (2013) Crystal structure of the Pseudomonas aeruginosa cytoplasmic heme binding protein, Apo-PhuS. J. Inorg. Biochem. 128, 131–136. [PubMed: 23973453]

(89). Mukhopadhyay K, and Lecomte JTJ (2004) A relationship between heme binding and protein stability in cytochrome b_5. Biochemistry 4-, 12227–12236.

(90). Robinson CR, Liu Y, O'Brien R, Sligar SG, and Sturtevant JM (1998) A differential scanning calorimetric study of the thermal unfolding of apo- and holo- cytochrome b_{562}. Protein Sci. 7, 961–965. [PubMed: 9568902]

(91). Scott NL, Xu Y, Shen G, Vuletich DA, Falzone CJ, Li Z, Ludwig M, Pond MP, Preimesberger MR, Bryant DA, and Lecomte JTJ (2010) Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. Biochemistry 49, 7000–7011. [PubMed: 20669934]

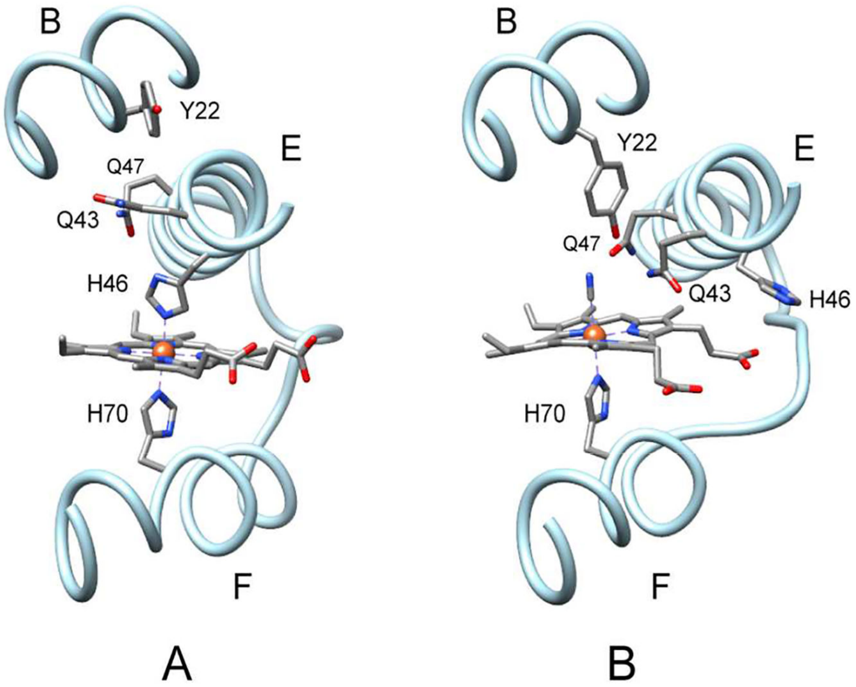


Figure 1.

Residues of interest in the heme binding site of GlbN (A) in the *bis*-histidine state (PDB ID: 1RTX)[13] and (B) in the cyanomet state (PDB ID: 1S69).[14] The E and F helices are shown, along with a portion of the B helix.

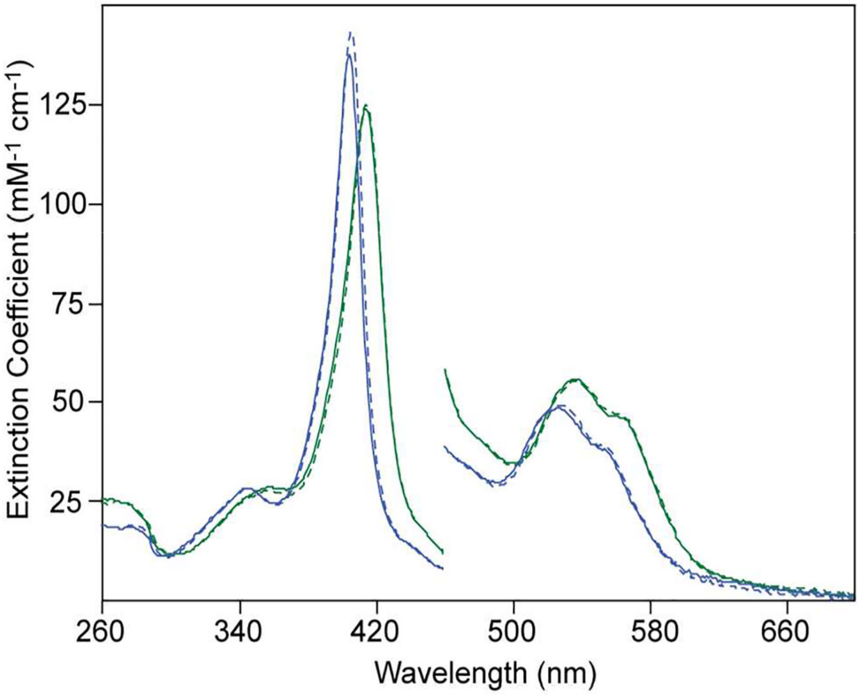


Figure 2.

The electronic absorption spectra of ferric GlbN variants containing mesoheme (blue) or *b* heme (green) at neutral pH. H46L GlbN (solid lines) and H46L/Q47L GlbN (dashed lines) are shown.

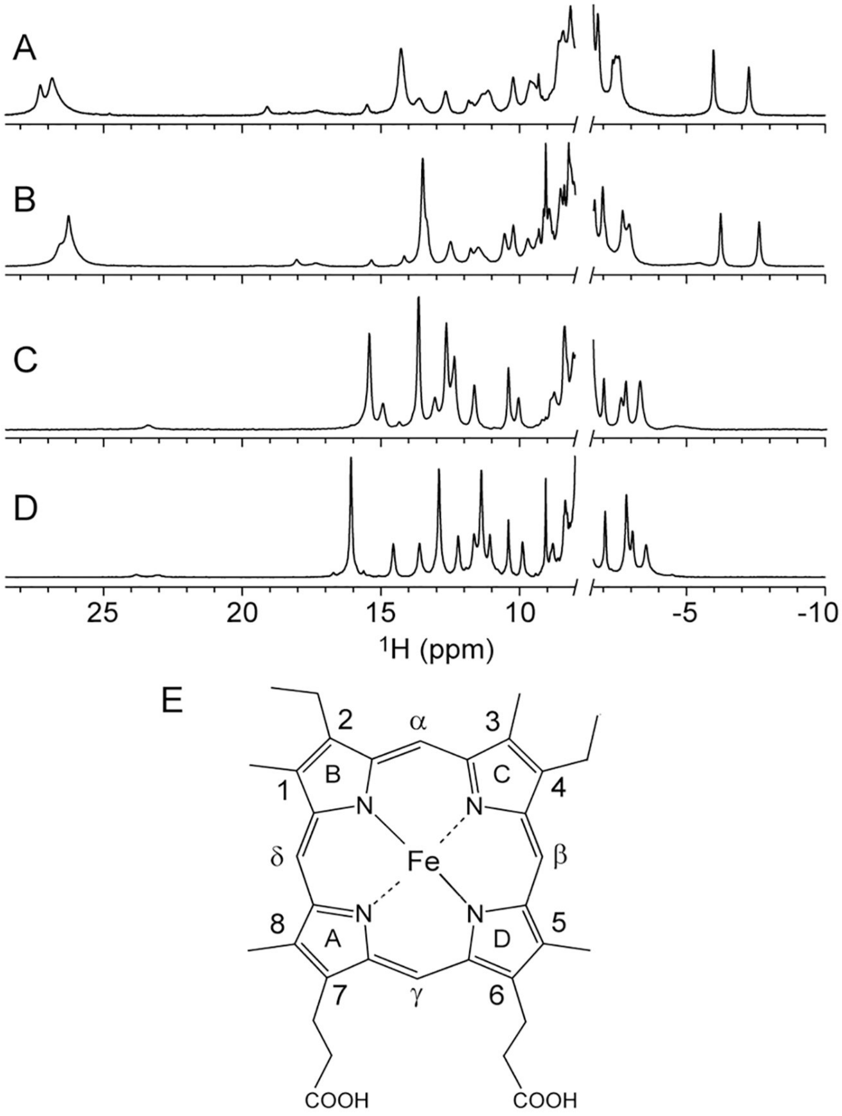


Figure 3.

The ^{1}H NMR spectra of ferric (A) H46L GlbN,[16]. (B)H46L/Q47L GlbN, (C) H46L mesoGlbN and (D) H46L/Q47L mesoGlbN. All samples contain 99% ^2H_2O, pH* 7.0–7.5. Vertical scaling is arbitrary. The *b* heme spectra (A and B) display a pair of broad peaks near −15 ppm that are not shown. (E) The structure of mesoheme and the nomenclature used in the text. Heme *b* has vinyl groups at positions 2 and 4.

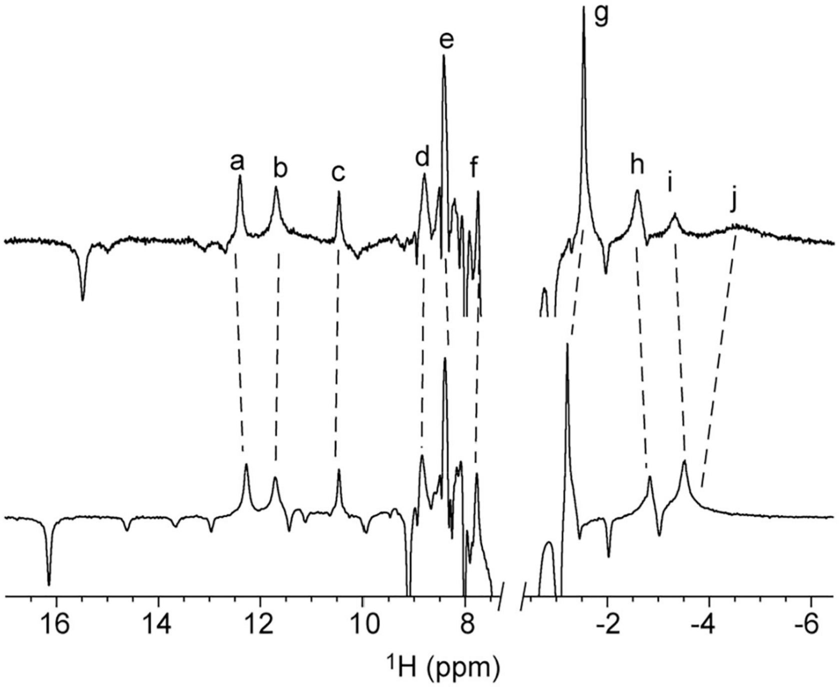


Figure 4.

The ^{1}H NMR inversion recovery spectra of H46L mesoGlbN (top, 0.5 mM, 99% ^2H_2O, pH* 7.3) and H46L/Q47L mesoGlbN (bottom, 2 mM, 99% ^2H_2O, pH* 7.2) using a 50 ms recovery delay. Positive peaks identify protons within ~6 Å of the ferric iron. a, His70 $H^{\beta3}$; b, His117 $H^{\beta3}$; c, Phe84 H^{ζ}; d, mesoheme δ-meso; e, Phe84 $H^{\varepsilon1/\varepsilon2}$; f, His117 H^{α}; g, Val74 $H^{\gamma1}$; h, mesoheme a-meso H; i, mesoheme γ-meso H; j, unassigned.

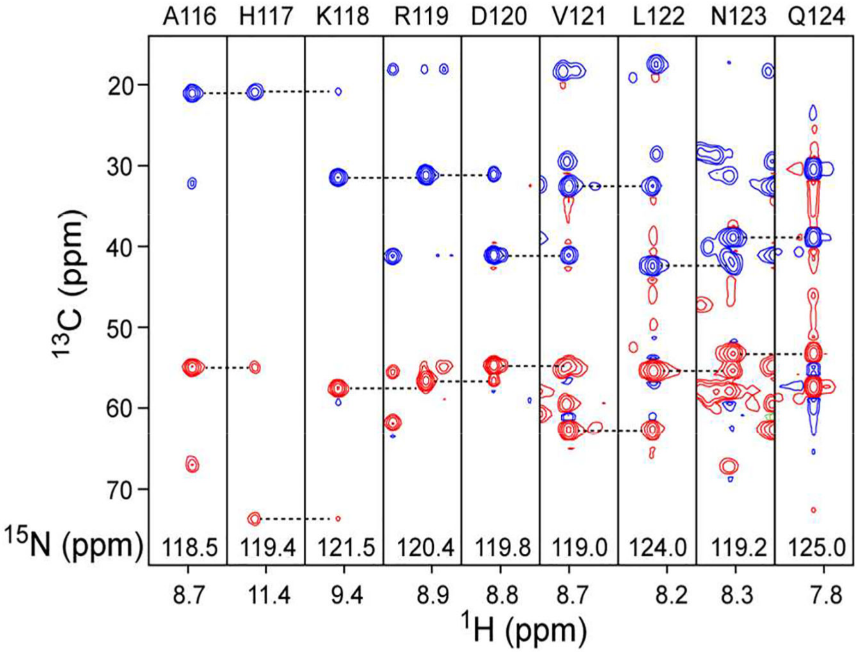


Figure 5.

A portion of the HNCACB spectrum of H46L/Q47L mesoGlbN (2.5 mM GlbN, 20 mM Na phosphate pH 7.1, 5% ^2H_2O). An isolated string of nine residues was observed and assigned to A116–Q124, the C-terminus stretch preceded by P115. The plane labeled H117 is amplified two-fold.

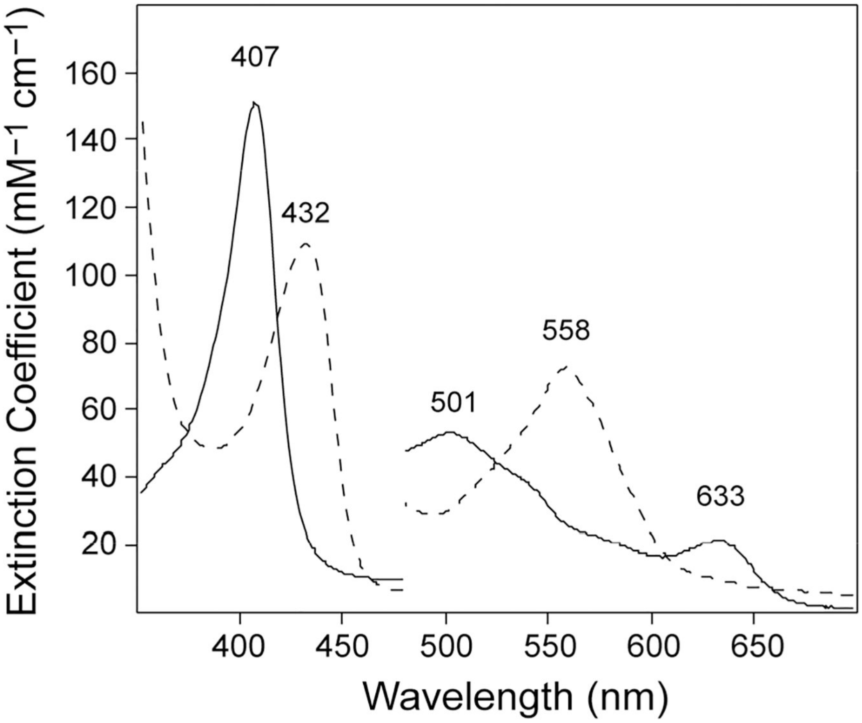


Figure 6.

Electronic absorption spectra of H46L/H117A GlbN at neutral pH. The ferric aquomet state (solid line) was converted to the ferrous deoxy state (dashed line) by addition of DT. The visible regions of the spectra have been amplified 5-fold.

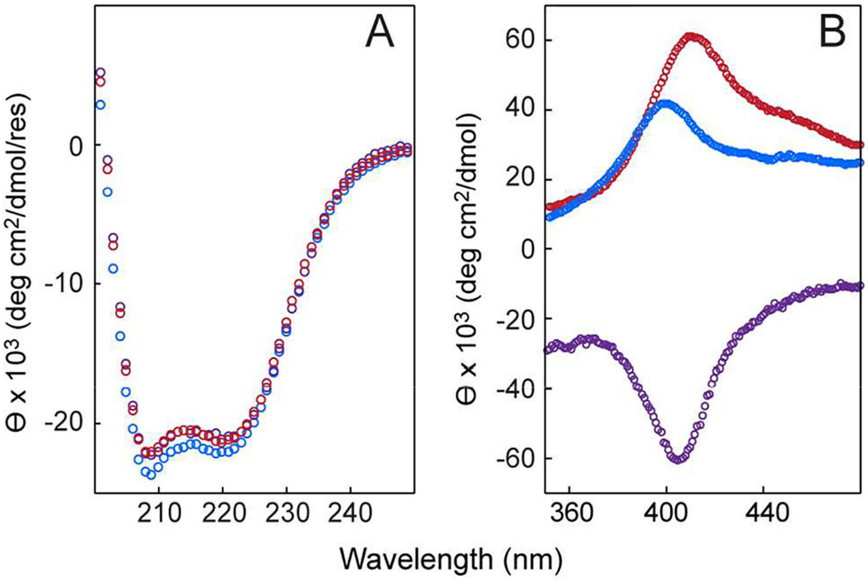


Figure 7.

Circular dichroism spectra of WT GlbN (purple, 15 μM, pH 7.4), H46L/Q47L GlbN* (red, 15 μM, pH 7.4), and H46L/Q47L mesoGlbN* (blue, 15 μM, pH 7.5). (A) Far-UV region and (B) Soret region.

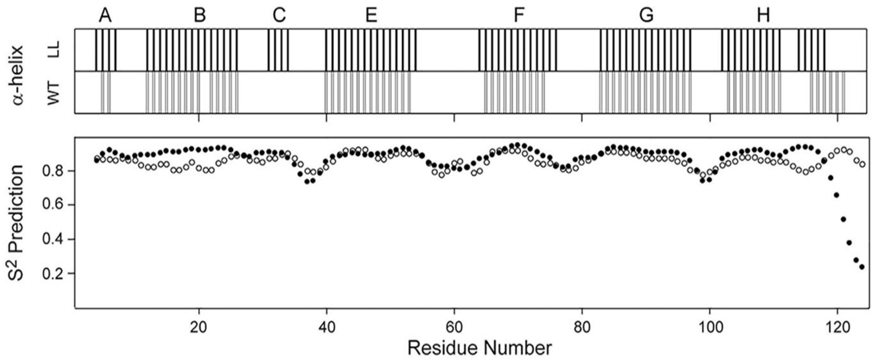


Figure 8.

TALOS+ prediction of α-helical secondary structure (top, confidence level above 0.3) and ^{15}N order parameters (bottom) using the H^N, H^α, C^α, C^β, N and C’ chemical shifts of H46L/Q47L mesoGlbN (“LL”, solid symbols) and WT GlbN (open symbols, BMRB: 5269). The letters above the figure denote the helices in WT GlbN.

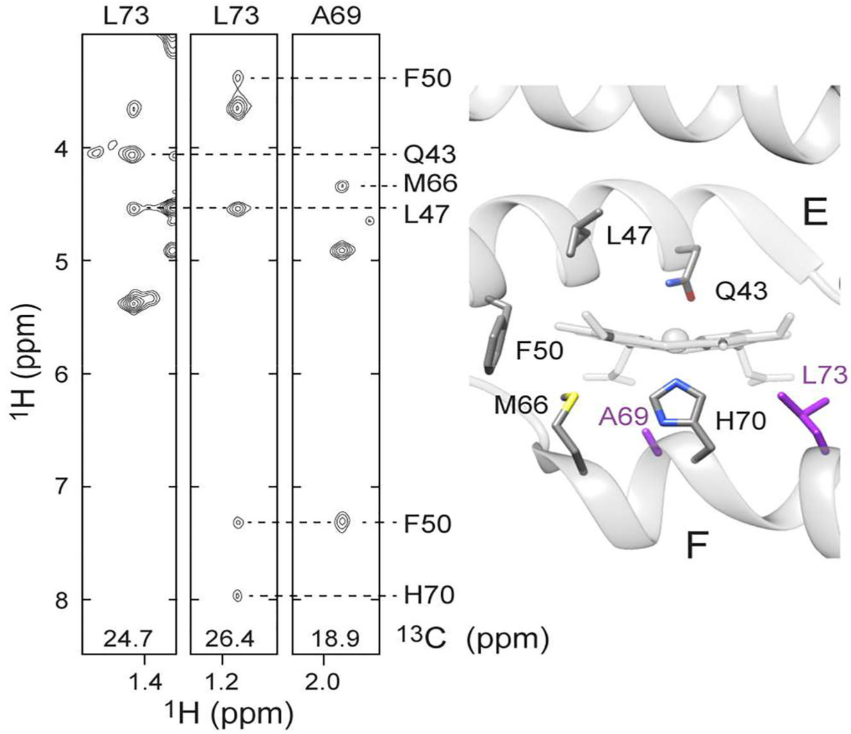


Figure 9.

NOESY cross peaks observed for Leu73 and Ala69 in H46L/Q47L mesoGlbN. Portions of the ^{13}CH$_{-}$ selective NOESY data displaying the CH$_{-}$ groups of Leu73 and Ala69 are shown on the left. Cross peaks to Phe50 H$^{\beta 3}$, Gln43 H$^{\alpha}$, Met66 H$^{\alpha}$, Leu47 H$^{\alpha}$, Phe50 H$^{\delta 1/\delta 2}$ and His70 H$^{\alpha}$ are labeled. Unlabeled peaks correspond to intraresidue NOEs. The relevant residues are shown on the right using the crystal structure of cyanomet GlbN (PDB ID: 1S69).[14] Leu73 and Ala69 are colored purple. The Gln47Leu replacement was generated with Chimera.[61] In this structure the distance between Leu7- and Leu47 is greater than 15 Å. The observed NOEs require the relative repositioning of the E and F helices and the removal of the intervening heme.

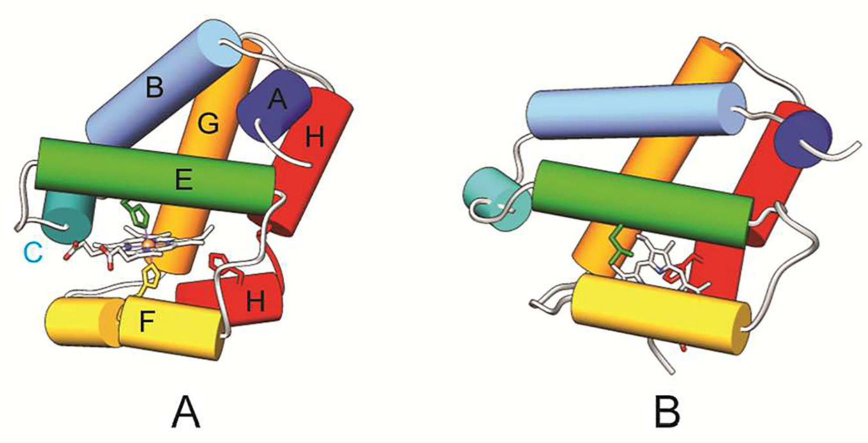


Figure 10.

Structural representations of GlbN. Left: bis-histidine GlbN (PDB ID: 1RTX); right: model of GlbN* accounting for TALOS+ secondary structure and NOEs detected in H46L/Q47L mesoGlbN*. Note the changed position of the heme group.

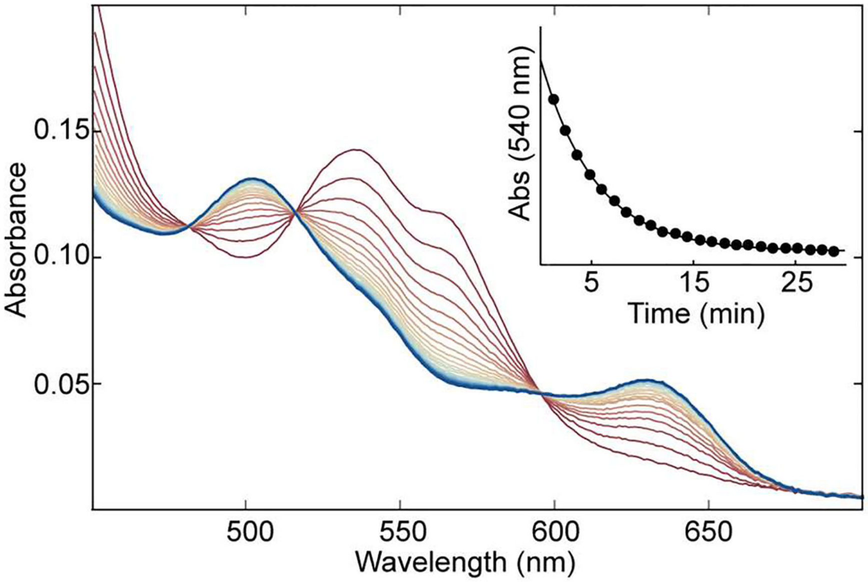


Figure 11.

The transfer of ferric *b* heme from H46L GlbN (10 μM) to apoMb (100 μM) at pH 7.2. The spectrum of H46L GlbN (red) is converted to the spectrum of aquomet Mb (blue). The inset shows the decrease in absorbance at 540 nm in time and the best fit to a single exponential decay.

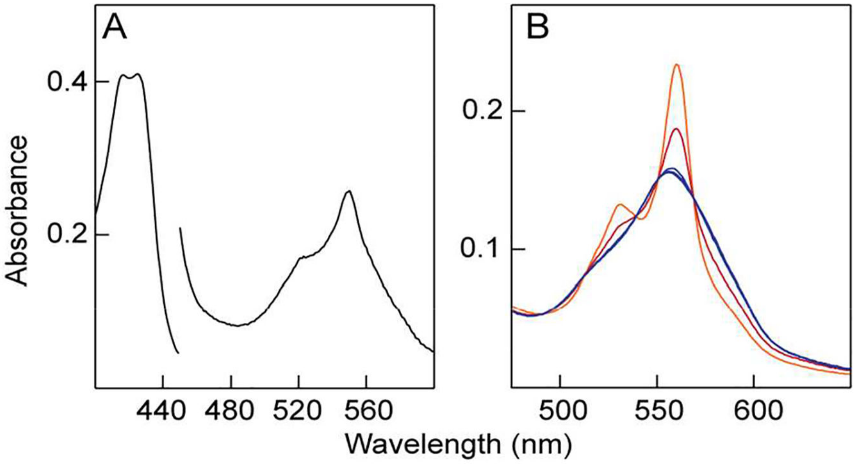


Figure 12.

(A) The optical spectrum of ferrous H46L/Q47L mesoGlbN at neutral pH. The visible region of the spectrum has been amplified 5-fold. (B) The transfer of ferrous *b* heme from H46L GlbN to apoMb. The spectrum of ferrous H46L GlbN immediately following reduction is shown in orange. Upon addition of reducing agent and apoMb, a fraction of ferrous H46L GlbN is observed (red trace, ~10 s after mixing) before complete conversion to deoxy Mb (blue traces). The orange spectrum was scaled according to the isosbestic point at 559 nm for ease of comparison.

Table 1.

Selected ^{1}H chemical shifts and T_1 values in H46L/Q47L GlbN and H46L/Q47L GlbN mesoGlbN[a]

		MesoH46L/Q47L		**H46L/Q47L**	
Residue	**Nucleus**	**δ (ppm)**	**T_1 (ms)**	**δ (ppm)**	**T_1 (ms)**
His70	NH	10.39		10.35	
	H^{α}	7.97		7.80	
	$H^{\beta3}$	12.32	45	12.55	35
	$H^{\beta2}$	8.02		8.10	
	N	120.1		119.8	
	C^{α}	82.9		84.8	
	C^{β}	22.5		22.2	
Hisll7	NH	11.49		11.23	
	H^{α}	7.78	60	7.49	
	$H^{\beta3}$	11.70	50	11.56	35
	$H^{\beta2}$	8.18		8.39	
	N	119.6		119.2	
	C^{α}	73.7		74.2	
	C^{β}	20.9		N.D.	

[a] 25 °C, 5% ^2H_2O, pH 7.1. T_1 values have a standard error of the fit lower than 5% (Figure S1).