HISTIDINE-HEME POSTTRANSLATIONAL MODIFICATION IN 2/2 HEMOGLOBINS: MECHANISM, CONSEQUENCES, AND ENGINEERING

by

Matthew Robert Preimesberger

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October, 2014

Abstract

Proteins harness the intrinsic reactivity of heme to carry out a remarkable array of chemical tasks within the cell. Hemoglobins (Hbs) are heme proteins classically characterized by their reversible oxygen binding and storage function within the red blood cells of vertebrates. In the last twenty years, genomics initiatives lead to the discovery of genes encoding Hb in a wide variety of organisms including: plants, bacteria, cyanobacteria, fungi, and archaea. Phylogenetic studies indicate an ancient origin for the Hb superfamily and that Hbs preceded the cyanobacterial invention of photosynthesis and oxygenation of Earth's atmosphere. Indeed, many microbial Hbs behave as enzymes and protect the cell by detoxifying reactive oxygen and reactive nitrogen species. As such, Hbs are a paradigm for protein functional diversification over evolutionary time.

Hbs fold around their cofactor and via specific interactions tune heme reactivity to achieve functional chemistry. The genome of the cyanobacteria *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. 6803 each encodes a single Hb (GlbN, 59% identity) with unusual structural features. In addition to the strictly conserved "proximal" histidine, these two GlbNs coordinate the heme iron using a 2nd histidine ("distal"). *Bis*-histidine hexacoordination is expected to lower reduction potential, accelerate electron transfer and condition ligand binding kinetics relative to classical "pentacoordinate" Hbs. GlbNs are further distinguished from other Hbs by their ability to bind heme covalently. The spontaneous irreversible modification links a non-coordinating histidine Nε2 to the heme 2-Cα and is analogous to the ubiquitous cysteine thioether (Cys-Sγ-

$C\alpha H$-$C\beta H_3$-heme) linkages observed in *c* cytochromes. The unusual His-heme posttranslational modification (PTM) in GlbN is thought to be triggered by reduction to the ferrous state. In vivo, GlbNs have been implicated in protecting their host organisms from reactive nitrogen species stress. Interestingly, both non-crosslinked and covalently modified (GlbN-A) proteins are detected in cyanobacterial cells.

The following dissertation attempts to answer several questions regarding the unusual structural, chemical, and functional features of GlbNs. Chapter 1 introduces the reader to the diverse world of heme proteins and attempts to provide some background necessary to appreciate the GlbN results. Chapters 2-3 are NMR and UV-visible spectroscopic and protein engineering studies aimed at understanding the mechanism by which ferrous GlbN spontaneously converts to the covalently modified state (GlbN-A). In this work, we show that a cysteine can substitute for histidine in the GlbN covalent modification and also test several key mechanistic features of the electrophilic addition reaction. My most significant contribution includes demonstrating that reducing *any* amount of GlbN, in a pool of oxidized GlbN, will result in a facile electron transfer chain reaction ultimately converting the entire sample to GlbN-A. This observation constitutes a novel signal amplification mechanism for a heme protein. Additionally, we made (to my knowledge) the first quantitative measurement of electron self-exchange in a *native* form of Hb.

Chapter 4 addresses differences between GlbN and GlbN-A with regard to their reactivity toward nitric oxide. In this work, I show that the covalent modification is required to prevent rapid heme dissociation upon binding nitric

oxide in the ferrous state. Formation of the PTM also appears to extinguish the unusual nitric oxide reductase activity of GlbN.

Chapters 5-6 are NMR structural studies in which we demonstrate that a histidine can be placed at different positions within the GlbN heme pocket and that nonnative PTM can be induced by reduction. We also show that the PTM can be successfuly transplanted in another Hb, in support that the covalent modification may be generally useful for heme protein engineering purposes.

Chapter 7 uses sophisticated NMR methodologies to explore the structural consequences of native and non-native PTM within GlbNs. I demonstrate that hydrogen bond scalar couplings can be used to probe directly minute strain and relaxation in hydrogen bonding. I believe this work will be generally useful for understanding the biophysical chemistry of proteins, and also provides motivation for future studies of hydrogen bond perturbation within enzyme active sites.

Overall, the results provide insight into the expanding chemical repertoire of the hemoglobin superfamily and pave a way for future structure-reactivity and protein engineering studies.

Advisor:

Professor Juliette Lecomte (primary reader)

Committee:

Professor Doug Barrick (2nd reader), Professor Jim Stivers, Professor Greg Bowman, Professor Dominique Frueh

Acknowledgments

My first acknowledgement goes to my advisor, Juliette Lecomte. Her intelligence, tireless work ethic, adherence to reason, healthy skepticism, communication skills, sense of humor, and empathy have provided me a model for the kind of scientist (and person) I want to be. Thank you so much Juliette, the last 7 years have been a great journey and I can't imagine having spent them any other way.

I am also lucky to have had so many great people to work with over the years. Hank Nothnagel and Matt Pond took me under their collective wing and taught me much of what I know about heme chemistry and NMR spectroscopy. In addition to being incredibly caring and challenging mentors, I consider them good friends. I must also thank other former and current Lecomte lab members: Dan Landfried, Ben Winer, Emily Adney, Selena Rice, Belinda Wenke, Lukas Gilevicius, Eric Johnson, Dillon Nye, and Chris Falzone. Thank you for providing a pleasant workplace environment and especially for the productive collaboration over the years.

Ananya Majumdar was an incredible mentor and good friend. We shared many exciting moments in the NMR facility (and laughs at One World). Thank you to my thesis advisory committee: Doug Barrick, Jim Stivers, Greg Bowman, and Dominique Frueh. Your teaching skills, insights, critique, and collaboration have improved the way I think about science. Bertrand Garcia-Moreno was an excellent teacher. Thank you for being so gracious with lab space and equipment. Thank you to Aaron Robinson for the many great

science and non-science conversations over the years. Additional thanks are due to Ryan Peterson, Richard Himes, and Kenneth Karlin for their assistance in acquiring stopped-flow data.

I also owe thanks to the Jenkins administration staff. Special thanks go to Jess Appel and Program in Molecular Biophysics coordinator Ranice Crosby. Thank you to my 2007 classmates for your support over the years; especially important to me my good friends Jackson Buss, Thuy Dao, and Helen Jun. In addition, I owe special thanks to role models from my formative years: college wrestling coach Pete Rogers, college advisor and teacher Brian Frink, high-school wrestling coach Dan Scharenbrock, and high-school math teacher Mike Allen. I would not have ever made it to where I am without you.

Thanks to my best friend, Mike Lee-Thompson, for moving to Baltimore and helping me keep sane over the years. Thanks to my closest friends back home: Adam, Heidi, Anna, and Nick. Thank you to my wonderful girlfriend Carla Coltharp, whose love has empowered me to be the best man I can be. Finally, I owe inestimable gratitude to my family for their tireless love and encouragement. Thank you Mom, Dad, Scott, and Mark.

This thesis is dedicated with love to the memory of Marian and Robert Preimesberger.

Contents

List of Tables

List of Figures

Chapter 1.

Introduction

Abbreviations: Fe, iron; PPIX, protoporphyrin IX; heme *b*, iron-protoporphyrin IX; Fe^{II}, ferrous iron; Fe^{III}, ferric iron; Mb, myoglobin; Hb, hemoglobin; CcP, cytochrome *c* peroxidase; SHE, standard hydrogen electrode; CO, carbon monoxide; O_2, dioxygen; CN^-, cyanide anion; NO·, nitric oxide; $NADP^+$, oxidized nicotinamide adenine dinucleotide phosphate; FAD, oxidized flavin adenine dinucleotide; GlbN, hemoglobins from the cyanobacteria *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803; GlbN-A, GlbN with histidine-heme posttranslational modification; PTM, posttranslational modification; DT, dithionite; ROS, reactive oxygen species; RNS, reactive nitrogen species; CtrHb, globin domain of *Chlamydomonas eugametos* chloroplastic hemoglobin; THB1, *Chlamydomonas reinhardtii* hemoglobin 1

Introduction

To survive and reproduce in the evolutionary battlefield that is life on Earth, organisms must accelerate the rate of thousands of chemical reactions, each by many orders of magnitude. Cells achieve this stunning feat primarily through the action of proteins. Proteins are composed principally of twenty genetically-encoded amino acids, strung together to form a polypeptide chain. The amino acids vary substantially in their hydrophobicity, polarity, charge state, aromaticity, size, and oxidation-reduction activity; these properties confer proteins the ability to specifically interact with a wide variety of substances and perform highly diverse chemistry. Christian Anfinsen demonstrated that the amino acid sequence of a given polypeptide chain determines its three-dimensional structure.[1] Over evolutionary time, proteins have adopted structures that facilitate their biochemical function and therefore confer a selective advantage to their hosts. The protein sequence → structure → function paradigm has contributed enormously to our understanding of how these complex molecules work at the molecular level.

1.1 Metals in Biology

In many instances the amino acid residues of a protein lack sufficient reactivity to catalyze the challenging oxidation-reduction (redox) reactions (e.g., oxidation of H_2O) necessary for life. In these instances, proteins have evolved to bind auxiliary molecules (cofactors), which often contain the ions of transition metals (Mn, Fe, Co, Ni, Cu). Iron (Fe) is ubiquitously used in biochemical reactions. Indeed, most organisms have evolved dedicated Fe-acquisition pathways and build molecules specifically to stably bind iron.[2]

The cyclic tetrapyrrole protoporphyrin IX (PPIX) is the primary molecule biosynthesized to chelate free iron. The four pyrrole coordination bonds enable tight binding of the metal to the center of the porphyrin, forming Fe-protoporphyrin IX (Fe-PPIX, Figure 1.1).

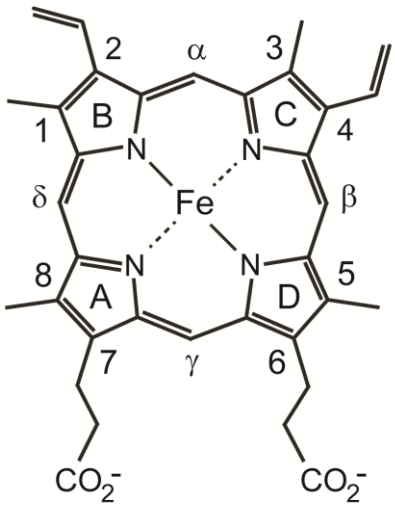


Figure 1.1. Structure and nomenclature for iron protoporphyrin IX (Fe-PPIX, heme *b*). Note that heme has C_{2v} pseudosymmetry about its α-γ meso axis.

The heme molecule is aromatic and contains methyl, vinyl, and propionate substituents. In the ferrous state (Fe^{II}), heme iron is formally neutral owing to the two negative charges on pyrrole ring nitrogens (as drawn in Figure 1.1, pyrroles B and D). At neutral pH in aqueous solution, the propionate groups are usually deprotonated leading to a net -2 charge. Heme oxidation to the ferric state (Fe^{III}) results in a net -1 charge. Along with the stable Fe^{III} and Fe^{II} oxidation states, higher valence states are also possible (Figure 1.2). Heme can also bind a multitude of small molecules that behave as Lewis bases. These qualities endow heme the capability to catalyze many redox and group transfer reactions. Indeed, proteins use heme to accomplish a wide array of chemistry including: small molecule transport and storage, gas sensing/signaling, peroxidation, monoxygenation, dioxygenation, and electron transfer. How do heme binding proteins achieve such functional diversity?

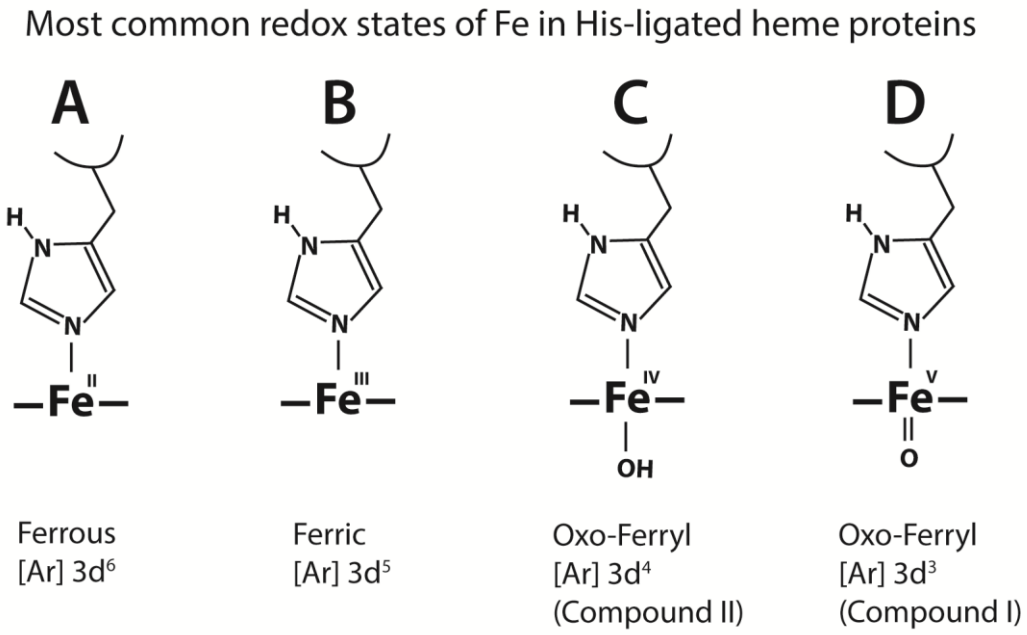


Figure 1.2. Primary oxidation states for iron in heme proteins. The hemes are depicted from an "edge-on" perspective with proximal and distal iron ligands above and below the plane of the ring, respectively. Elemental iron has the electronic configuration $[Ar]3d^64s^2$. A) Fe^{II} is a dication with configuration $[Ar]3d^6$. B) Fe^{III} is a trication with configuration $[Ar]3d^5$. C) Compound II is a reactive intermediate formed during peroxidase and monoxygenase oxidation reactions. D) Compound I contains a formal Fe^{V} and is the more strongly oxidizing precursor of Compound II.

1.2 Factors governing heme protein reactivity

Polypeptides fold around the cofactor and alter heme chemical environment by engaging in specific interactions. As exemplified in Figure 1.2, heme proteins often use amino acid side chains to coordinate the axial position(s) of the central iron. Five-coordinate complexes adopt square pyramidal geometries whereas six-coordinate complexes are octahedral. The identity of the amino acids and geometry of the ligand field significantly alters the intrinsic magnetic properties of iron (Figure 1.3).

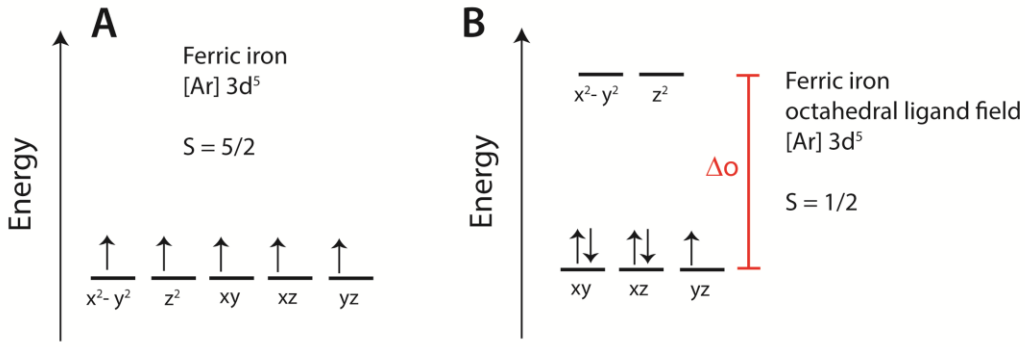


Figure 1.3. Ligand field splitting diagram for the d-orbitals of ferric iron. A) In the absence of a ligand field, the five d-orbitals of free ferric iron are degenerate. According to Hund's rule, the five d-electrons remain unpaired (maximizing total angular momentum) leading to the high spin (S=5/2) state. B) In naked heme, the porphyrin ring chelates the central iron using four pyrrole nitrogen ligands resulting in a square planar geometry. In heme proteins, amino acid side chains (or water) adopt the axial coordination positions resulting in octahedral iron geometry. For strong ligand-metal bonds, a large increase in energy for both Fe^{III} d_{x2-y2} and d_{z2} orbitals is expected owing to their electrostatic repulsion with the octahedral ligand field. If this is the case, the d-orbital octahedral crystal field splitting energy (Δo) becomes larger than the spin-pairing energy and the ferric iron adopts the low spin (S=1/2) state.[2] Weak field ligands or distortions from ideal geometry will lead to an altered splitting pattern and potentially form the high spin (S=5/2) state.[3]

The most common amino acids used for heme iron coordination bonds are: histidine (Nε2), methionine (Sδ), and cysteine (Sγ), although examples of tyrosinate (Oη, catalase), N-terminal amine (NH_2, cytochrome *f*), and lysine ($N\zeta H_2$, cytochrome *c* nitrite reductase) ligands do exist.[4] Among these, histidine, methionine, cysteine, and amine ligands are typically strong enough to generate a low-spin Fe center. Beyond spin state, additional heme properties can be tuned by the nature of the ligand field and immediate environment. For example, the cysteine thiolate ligand of cytochromes P_{450} imparts negative charge density to the iron, lowering reduction potential and stabilizing its high-valent redox states. On the other hand, stable His-Fe-Met coordination in both ferric and ferrous states precludes the binding of exogenous ligands and allows cytochrome *c* to behave as an electron shuttle. Other electron transfer proteins such as cytochrome b_5 utilize *bis*-histidine coordination to block the metal center and condition redox qualities. Pentacoordinate heme proteins such as ferrous myoglobin (Mb) contain an empty distal iron site and bind exogenous ligands such as O_2 reversibly. A summary of different coordination modes used by heme proteins is shown in Figure 1.4.

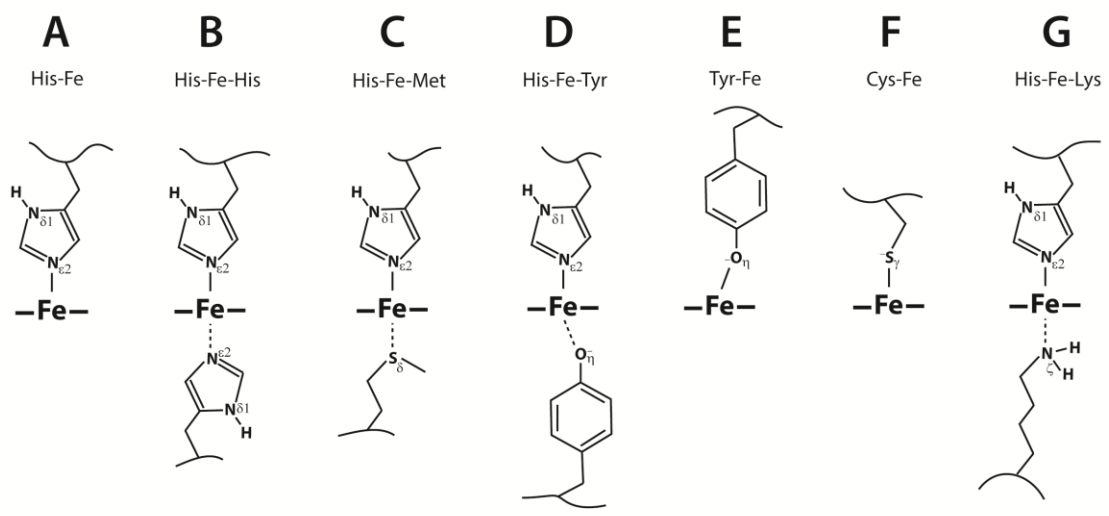


Figure 1.4. Heme physicochemical properties such as redox potential, electron transfer kinetics, and accessibility to exogenous ligands are strongly modulated by the nature of the amino acid coordination bonds. For clarity, only protein side chains are shown; in general, heme is capable of binding a wide array of small molecules such as: CO, NO·, O_2, CN^-, HS^-, H_2S, F^-, N_3^-, and H_2O. A) His-Fe coordination in myoglobin.[5] B) *Bis*-histidine (His-Fe-His) coordination in cytochrome b_5.[6,7] C) His-Fe-Met coordination in cytochrome *c*.[8] D) Proposed His-Fe-Tyr coordination in *C. eugametos* truncated hemoglobin, CtrHb.[9,10] E) Tyrosinate-Fe-H_2O coordination in ferric heme catalase.[11] F) Cysteine thiolate-Fe coordination in ferrous cytochrome P_{450}.[12] G) His-Fe-Lys iron coordination observed for *C. reinhardtii* truncated hemoglobin, THB1.[13]

In general, heme proteins alter the chemical properties of the cofactor through several effects. For example, the solvent accessibility of heme can greatly alter its reduction potential (typically stabilizing the charged Fe^{III} state).[14] The presence of negatively charged groups near the iron can also result in decreased reduction potential.[15] Conversely, a highly hydrophobic heme binding pocket tends to favor the neutral (ferrous) state. Combination of distinct heme ligation, solvent accessibility, and electrostatic environment (along with other features, vide infra) allows a single protein family such as the

cytochromes *c* to modulate the reduction potential of heme over a range of 0.8 V (from -400 to 400 mV vs. SHE).[16,17]

Heme proteins can also achieve similar chemistry using distinct heme environments. Figure 1.5 presents key residues within the heme pocket of the H_2O_2-detoxification enzyme yeast cytochrome *c* peroxidase (CcP). The enzyme adopts His-Fe^{III}--H_2O coordination in its ferric resting state and catalytically reduces H_2O_2 to $2H_2O$ by oxidation of two cytochrome *c* molecules ($H_2O_2 + 2H^+ + 2e^- \rightarrow 2H_2O$). The activity of heme peroxidases such as CcP is thought to involve a "push-pull" mechanism, in which electron donation ("push") from H75 stabilizes the two electron oxidation of iron by H_2O_2 to form Compound I (Fe^{V}=O or Fe^{IV}=$O^{\cdot+}$) and loss of H_2O.[18] The distal residues H52 and R48 are important for protonation of the leaving water ("pull"). The porphyrin radical generated by the two electron oxidation is rapidly transferred to Trp191.[19] Sequential electron transfers from cytochrome *c* (with corresponding protonation events) quench the radical and return the enzyme to its resting state (His-Fe^{III}--H_2O).

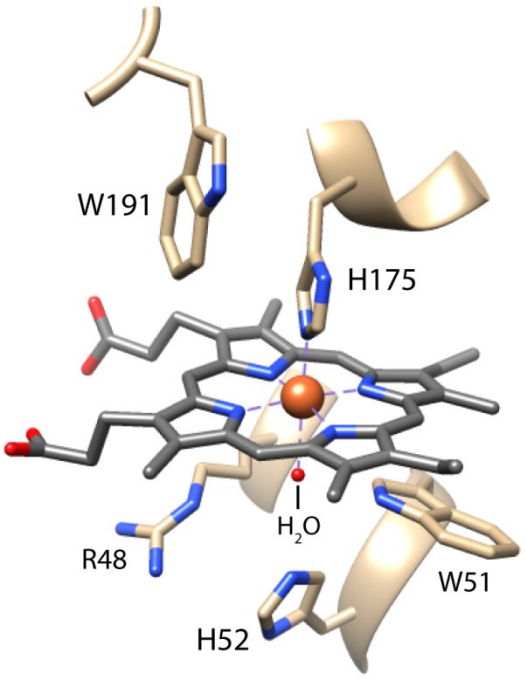


Figure 1.5. Heme environment in the ferric "resting state" of yeast cytochrome *c* peroxidase (PDB: 2CYP). H_2O_2 displaces water to initiate chemistry.

To contrast, Figure 1.6 presents the heme pocket environment of a different heme enzyme, cytochrome P_{450}. P_{450}s are a diverse class of monoxygenases that have the potent ability to oxidize specific C-H bonds to form hydroxyl groups. Similar to the heme peroxidases, cytochromes P_{450} oxidize substrates through formation of a Compound I type intermediate.[20,21] However, in the P_{450} instance, a thiolate side chain (C357) donates negative charge to the heme iron for stabilization of positive charge buildup upon metal oxidation.[22,23] Unlike the heme peroxidases, which react with H_2O_2 in the ferric state, cytochromes P_{450} (usually) undergo NADPH dependent-reduction and initiate chemistry in the ferrous state by binding O_2 ($O_2 + H^+ + C\text{-}H + NADPH \rightarrow C\text{-}OH + H_2O + NADP^+$).

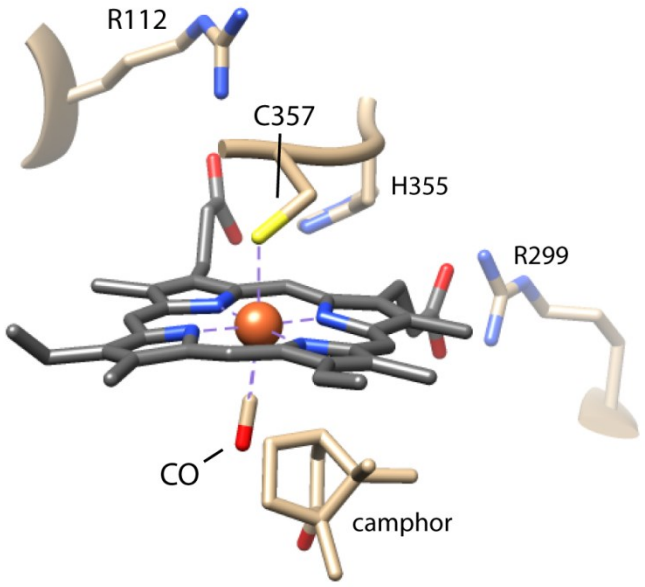


Figure 1.6. Heme environment in the ternary complex of *Pseudomonas putida* cytochrome $P450_{cam}$ (PDB: 3CPP) with CO (O_2 mimic) and substrate (camphor). The camphor substrate binds within 3 Å of the CO oxygen atom.

It is clear from comparison of the CcP and $P450_{cam}$ heme pockets that distinct coordination modes can be used to generate the same reactive intermediate (Compound I). However, Compound I generated by peroxidases and P450s is utilized for different purposes (H_2O_2 detoxification and substrate oxidation).

As discussed above, proteins bind their cofactor and modulate heme chemistry using specific interactions (coordination bonds, H-bonds, hydrophobic packing). In this manner, stable heme binding and functional chemistry is selected for and undesired side reactions limited. In many instances, however, non-covalent interactions are insufficient, and the heme group is chemically modified.

1.3 Heme covalent modification

An additional mechanism by which a polypeptide chain can tune heme chemistry is through covalent modification of its cofactor. Several positions along the heme ring can undergo reaction and in many cases result in its covalent attachment to the protein. Mammalian peroxidases and some P_{450} enzymes are capable of oxidative modifications of heme methyl, vinyl and meso groups.[24,25,25–27] In some instances, these heme-protein crosslinks appear to protect the enzyme from heme damage and other catalytically-driven deactivation processes.[24] An example of a modified heme not attached to polypeptide can be found in the respiratory proton pump cytochrome *c* oxidase.[28] This important enzyme uses heme *a*, a cofactor which contains a formyl substituent at position 8, and an isoprenoid extension at position 2.

By far, the most prevalent chemical modification found in heme proteins is that of the *c* cytochromes. These proteins often act as shuttles within photosynthetic and respiratory electron transfer chains. Cytochromes *c* contain conserved CXXCH motifs, whereby each cysteine sulfur forms a thioether bond resulting in saturation of the heme 2- and 4-vinyl groups (Cys-Sγ-CαH-CβH_3-heme linkages). The histidine within the CXXCH motif acts as proximal ligand to the iron. Figure 1.7 illustrates these Markovnikov adducts (vinyl Cα reacted), along with the His-Fe-Met coordination scheme for cytochrome c_{553}.[8]

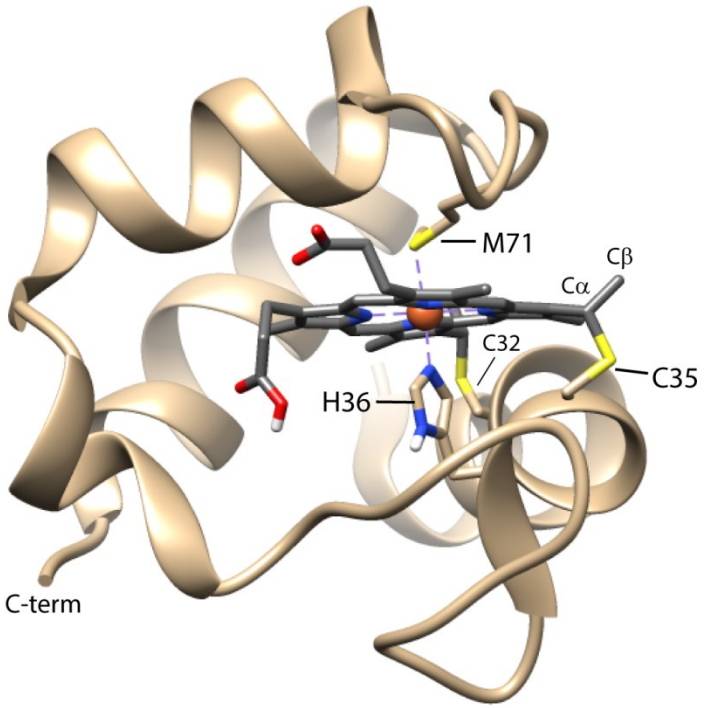


Figure 1.7. X-ray ribbon model of cytochrome c_{553} from the gram-positive bacteria *Bacillus pasteurii* (PDB ID: 1C75). Many *c* cytochromes utilize stable His-Fe-Met coordination, and a solvent-exposed heme edge to facilitate rapid electron transfer. Covalent heme binding is achieved through formation of cysteine-heme thioether linkages. In cytochrome c_{553}, C32 and C35 react with the 2- and 4-vinyl groups to form Cys-Sγ-CαH-CβH_3-heme moieties.

The heme-cysteine thioether linkages observed in cytochrome *c* can be formed in vitro, although rather inefficiently.[29–31] Additionally, protein engineering work has demonstrated the ability to transplant a *c*-type linkage into *b*-type cytochrome, but a mixture of products is often obtained.[32] Several factors can influence whether correct Markovnikov addition (Sγ-CαH-CβH_3-heme) occurs. For example, apocytochromes *c* (without cofactor) typically lack well-defined tertiary structure and do not bind *b* heme tightly. Poor heme seating will inhibit the posttranslational modification through geometric effects.

Additionally, the cysteines must remain in their reduced state in order to prevent the formation of a CXXCH disulfide bond. However, while in the reduced state (-SH), a cysteine near heme is prone to oxidation by ferric (or ferryl) iron resulting in thiyl radical formation and incorrect reaction (anti-Markovnikov, Sγ-CβH_2-CαH_2-heme linkage).[32] In vitro studies on the mechanism of *c*-heme formation indicate that reduction of the heme iron is a key requirement to prevent radical formation and generate proper product.[29] Collectively, these observations indicate that the intrinsic redox reactivity of cysteine results in competing side reactions and an inherent challenge to *c* heme production.

In vivo, however, several dedicated systems exist to ensure proper cytochrome localization, folding, heme binding, and thioether bond formation.[33,34] In *E. coli*, the cytochrome *c* maturation system I (CcM I) involves the protein products of at least 12 genes. Clearly, the cell invests considerable energetic resources into ensuring that the thioether bonds form properly. Yet, the functional preference for *c* heme over *b* heme in these proteins is still unknown.[35]

Formation of *c* heme has several possible effects on heme protein properties. 1. Heme affinity: proteins with covalently and irreversibly bound heme do not lose their cofactor. This property may be especially useful for systems with low intrinsic affinity for *b* heme.[36] 2. Redox chemistry: saturation of the heme vinyl groups decreases the degree of conjugation along the porphyrin ring, altering the heme electronic structure. However, it is difficult to predict the direction and magnitude of this effect.[37] 3. Protein stability: it may be expected that

covalently bound heme will enhance the stability of the attached protein.[38] Increased stability could alter a protein's lifetime and degradation pathway within the cell. However, forming covalent bonds does not guarantee an increase in protein stability.[39] 4. Packing density: in multi-heme *c*-cytochromes, formation of the thioether bonds is thought to allow a low amino acid:heme ratio. This minimizes the amount of polypeptide required for stable heme binding, and also enables tight packing and specific orientation of adjacent heme groups.[35] These features allow multi-heme cytochromes to serve efficiently in a variety redox reactions involving long-range directional electron transfer.[40] Figure 1.8 illustrates the heme packing effect for the tetra-heme cytochrome *c* from *Shewanella oneidensis*.[41]

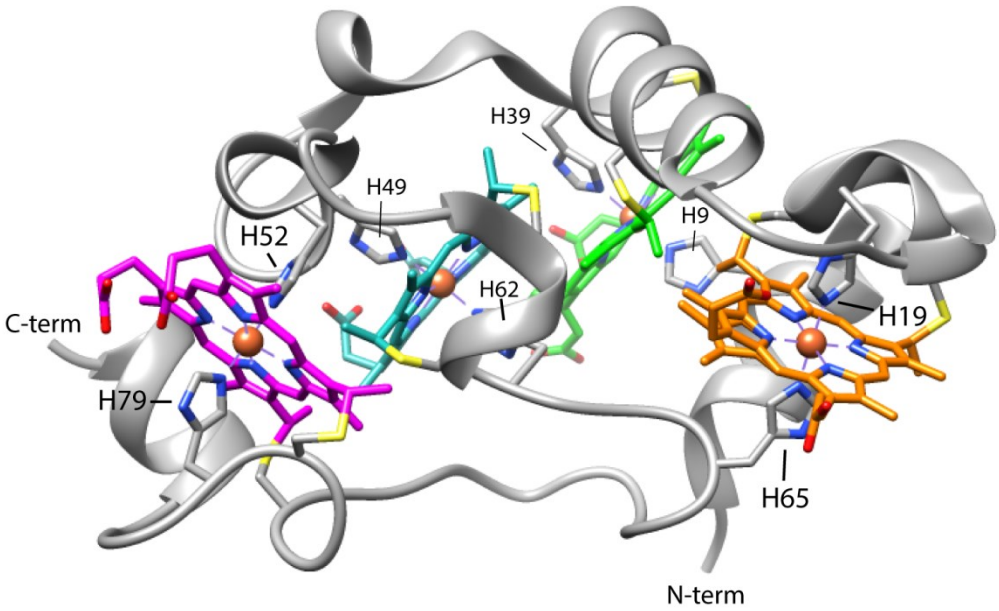


Figure 1.8. X-ray ribbon model of the ~90 residue tetra-heme *c* cytochrome from *Shewanella oneidensis* (PDB ID: 1M1Q). The protein contains four tightly packed covalently bound *c*-heme groups. Each heme shows *bis*-histidine coordination (histidines ligands labeled). Multi-heme cytochromes *c* often act in nitrogen and sulfur metabolism as mediators for oxidoreductase reactions requiring multiple electron transfers. Remarkably, some multi-heme cytochromes *c* can contain > 25 hemes per polypeptide chain.[42] Relative to *b* heme proteins, *c* cytochromes can achieve stable heme binding using strikingly low amino acid:heme ratios (<25 residues per heme).

1.4 Classic view of hemoglobin function

The oxygen binding molecules hemoglobin (Hb) and myoglobin (Mb) are arguably the best studied proteins. Pioneering work by Christian Bohr demonstrated that under a given pO_2, the extent of Hb-O_2 binding is linked to the concentration of other ligands such as H^+. Specifically, under constant pO_2, increasing concentration of H^+ causes a decrease in O_2 saturation. This phenomenon has come to be known as the Bohr effect and is the classic example of (negative) heterotropic allostery in a protein. The oxygen binding curves of Mb and Hb provide additional insight into differences in their functional chemistry. Hb is composed of four subunits, whereas Mb is monomeric. Mb binds oxygen tightly and exhibits a hyperbolic binding curve. A Hb $\alpha_2\beta_2$ tetramer binds four oxygen molecules but with lower affinity than Mb. Additionally, the sigmoidal shape of O_2 binding curve for Hb is interpreted to indicate that the four oxygen binding sites are not independent; instead, the binding of an oxygen molecule at one site promotes binding at the others in a cooperative fashion.[43] The cooperative O_2 binding behavior by Hb has come to be the classic example of (positive) homotropic allostery. These ideas were further developed by Monod, Wyman, and Changeux to account quantitatively for cooperativity in Hb.[44] In vertebrates, Hb and Mb are present at high concentrations (Hb estimated as ~5 mM in red blood cell). This observation, along with their distinct O_2 binding behaviors, lead to the textbook view that Hb and Mb serve as the efficient O_2 transport and storage team required for cellular respiration.[45]

Myoglobin and hemoglobin also provided the first atomic-resolution pictures of proteins.[5,46] The X-ray crystal structures determined for Mb and Hb allowed unprecedented and intricate rationalization of their chemical properties. For example, Perutz proposed a detailed mechanism for cooperative oxygen binding behavior in Hb. In the absence of oxygen, the tetramer adopts a T-state (tense) quaternary arrangement with low oxygen affinity (and square pyramidal iron geometry); upon a single O_2 binding event, the iron forms an octahedral geometry and is pulled within the plane of the heme ring leading to strain of the proximal histidine-iron bond (negative *trans* effect). The proximal strain is transmitted through inter-subunit interactions and promotes a conformational change to the R-state (relaxed) quaternary arrangement, which has high affinity for O_2. Thus, the binding of a single oxygen molecule can be sensed by all members of the tetramer through communicated structural deformation.[47] In support of the Perutz mechanism, uncoupling of the proximal histidine from the polypeptide backbone and *trans* rescue with imidazole results in a Hb with significantly decreased cooperativity and heightened O_2 affinity.[48]

The three-dimensional structures of Hb β chain and Mb are virtually superimposable. Subsequent structure determination of hemoglobins from seal and human (among others) unequivocally demonstrated the evolutionarily-conserved globin fold. Figure 1.9 presents the X-ray crystal structure of a representative member of the globin superfamily: sperm whale oxy-Mb.[49]

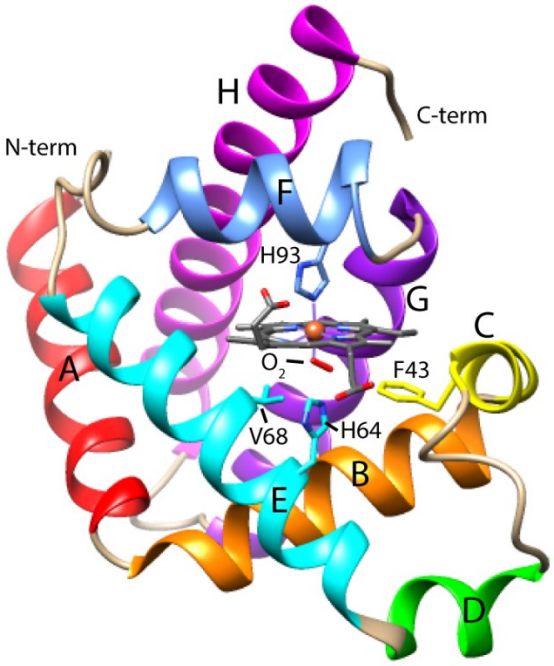


Figure 1.9. X-ray ribbon model of the classic O_2-storage protein, sperm whale oxy-myoglobin (Mb-O_2, PDB ID: 1MBO). α-Helices are labeled A-H according to Perutz's notation and arrange into a "3-on-3" sandwich. The strictly conserved "proximal" histidine (H93) occurs at position F8 and coordinates the heme iron using its Nε2 atom.

Dioxygen (O_2, red) is bound to the "distal" ferrous iron site and stabilized in a bent configuration via an H-bond interaction with H64 Nε2-H (E7). Additional important residues in the distal pocket are highlighted in Figure 1.10: F43 (position CD1) is highly conserved and engages in π-stacking with the porphyrin, whereas V68 modulates ligand access to the iron (Figure 1.10).[50,51]

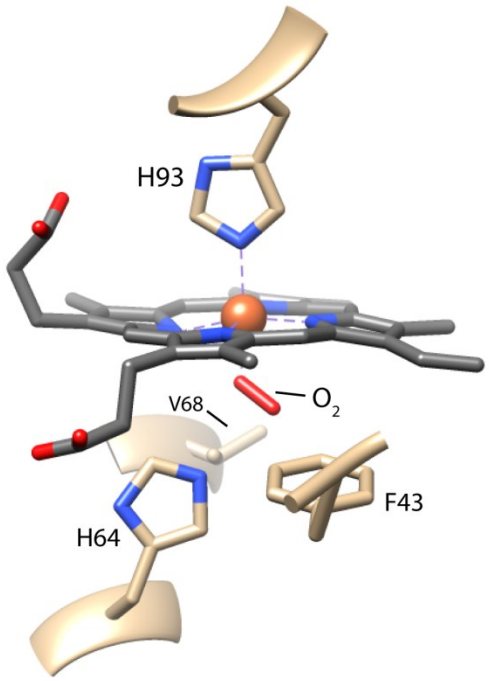


Figure 1.10. Close-up view of the heme pocket in sperm whale oxy-myoglobin (PDB ID: 1MBO). The Fe^{II}-O_2 complex is better represented as Fe^{III}-O_2·$^{-}$ and can form a stable bent geometry by interaction with H64. Carbon monoxide prefers to bind ferrous iron in a linear geometry, which takes full energetic advantage of its capacity as a π back-bond acceptor. In the distal heme pocket of Mb, the presence of H64 and V68 decrease CO affinity by sterically straining its preferred coordination geometry.[52]

1.5 Polyfunctionality in the Hb superfamily

Although hemoglobins have long been known to exist in organisms such as nematodes, lamprey, and plants, the genomic revolution has proven enormously useful in understanding hemoglobin phylogeny and functional diversity. In particular, the identification of genes encoding Hbs in organisms spanning all three domains of life (eukaryotes, bacteria, and archaea) indicates an ancient origin for the superfamily (estimated ~3.5 bya).[53] The presence of

Hbs in such diverse organisms may be somewhat surprising, since dioxygen was not present on Earth at high concentrations until the invention of photosynthesis (estimated ~2.6 bya). Thus, the well-known dioxygen transport and storage role for Hbs appears to be a relatively recent adaptation in higher organisms. Instead, many hemoglobins are thought to function in reactive oxygen species/reactive nitrogen species (ROS/RNS) stress response.[54]

Recent phylogenetic analyses indicate that the Hb superfamily comprises three primary lineages.[55,56] The M (myoglobin-like) family of Hbs contains the vertebrate Hbs, and is characterized by the classic "3-on-3" architecture. Many M-family Hbs are chimeric proteins containing a reductase domain capable of binding flavin adenine dinucleotide ($FADH_2$). In vivo and in vitro studies have demonstrated that these flavohemoglobins act as potent nitric oxide ($NO\cdot$) detoxifiers via the so-called "NO dioxygenase" reaction (NOD, $O_2 + NO\cdot + e^- \rightarrow NO_3^-$).[57,58] To undergo efficient NOD chemistry, the protein must be capable of rapid Fe^{III}/Fe^{II} redox cycling. The structure presented in Figure 1.11 provides clear insight into how this activity is achieved in *E. coli* flavohemoglobin.[59]

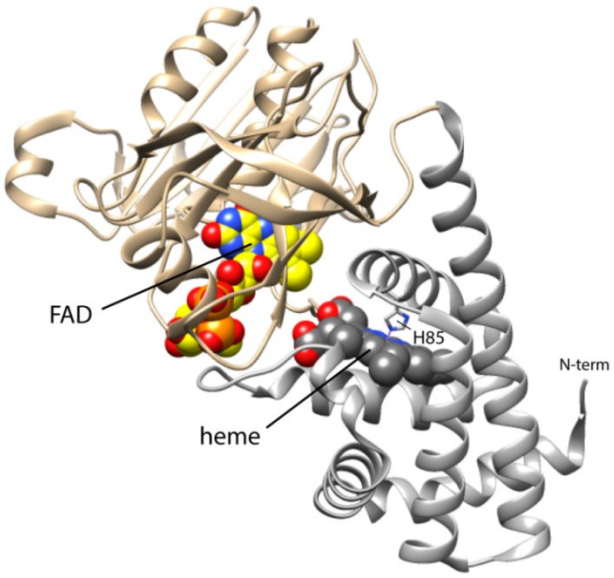


Figure 1.11. X-ray model of *E. coli* ferric flavohemoglobin (PDB: 1GVH). The proximal Fe-coordinating histidine (H85) is labeled. Flavohemoglobins are NODs, which contain heme (gray) and FAD-binding (beige) domains within a single polypeptide chain. The redox cofactors are positioned near one another in order to achieve fast first order electron transfer from $FADH_2$ to the Fe^{III} center.

NOD chemistry is a ubiquitous activity for hemoglobins.[60] For example, in humans, red blood cell Hb NOD activity appears to be the predominant elimination mechanism for endogenously produced NO·.[61] Along with their ability to *destroy* NO· via the NOD reaction, most hemoglobins are also capable of *producing* NO· via nitrite reductase chemistry ($NO_2^- + 2H^+ + 1e^- \rightarrow NO\cdot + H_2O$).[62–65] These observations indicate that even vertebrate hemoglobins are more than just O_2 containers and function in regulating in vivo NO· concentration.

Hemoglobins within the S (Sensor) lineage are phylogenetically distinct from the M-family, but adopt a similar "3-on-3" fold. Sensor hemoglobins occur in bacteria and archaea and are composed of three sub-families: globin-coupled sensors, sensor single domain globins, and the single domain protoglobins (Pgb).[56] Although little functional information is available for these unusual proteins, many contain gene regulatory domains suggesting a function in ligand-dependent gene regulation. A high-resolution crystal structure of Pgb[66] is presented in Figure 1.12.

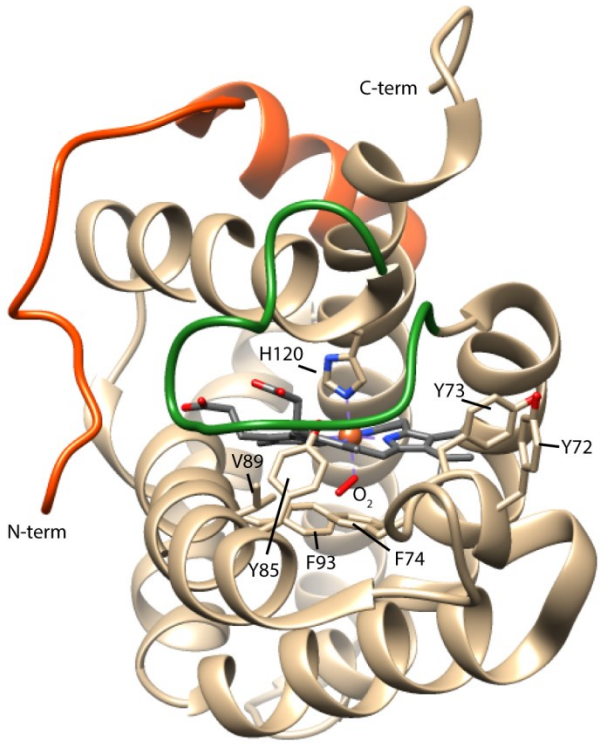


Figure 1.12. X-ray model of Pgb, an Hb from the anaerobic archaea *Methanosarcina acetivorans* C2A (PDB: 2VEE). An N-terminal extension (orange) and extended FG loop (green) fold back over the heme propionates leading to a completely buried cofactor. Dioxygen and the proximal histidine (H120) are highlighted to orient the reader. Note the highly hydrophobic nature of the distal pocket (F93, V89, Y85, and F74).

The T (truncated) family constitutes the third phylogenetic lineage of globins and can be further sub-classified into Groups I, II, and III.[67] These proteins are structurally distinct from M and S family Hbs. Specifically, T globin domains are ~30 residues shorter than their "3-on-3" counterparts (~150 residues), and have short or missing A and D-helices.[68] These proteins are found in organisms such as unicellular eukaryotes, bacteria, fungi, and plants, and are suggested to play a variety of functional roles.[69]

The unicellular alga *Chlamydomonas eugametos*, contains genes encoding at least two truncated hemoglobins. Analysis of gene expression data indicates that these globins are light-induced and associated with photosynthesis.[70] In support, cellular localization by electron-microscopy indicates their presence in the chloroplast, which suggests a novel function. The globin domain of the product of one of these genes (light-induced gene 637) is CtrHb. CtrHb was subsequently expressed recombinantly and crystallized to provide an initial view of the truncated globin fold (Figure 1.13).[68]

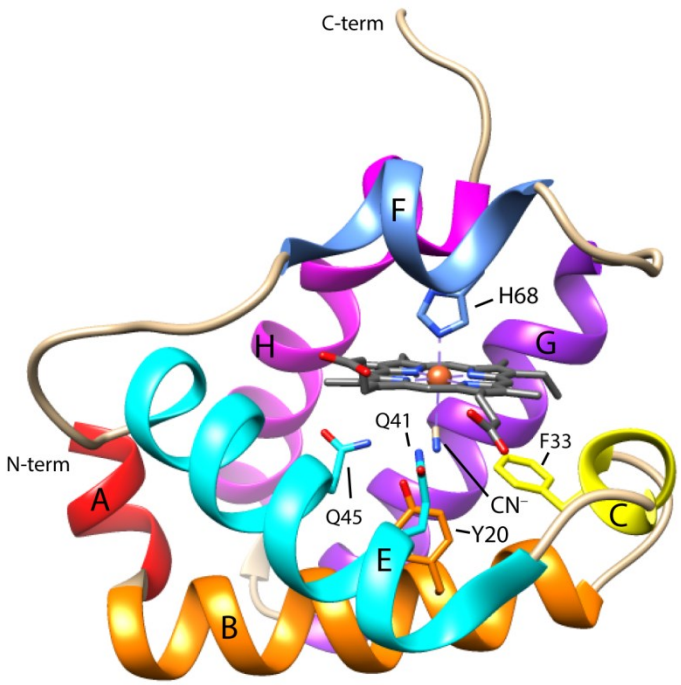


Figure 1.13. X-ray model of the globin domain of *Chlamydomonas eugametos* chloroplastic hemoglobin, CtrHb (PDB ID: 1DLY), a representative member of the Group I T-family of globins. Relative to Mb, the protein adopts an abbreviated "2-on-2" α-helical bundle fold. Helices are labeled according to Perutz notation, which is used to describe specific positions. Cyanide (CN^-) is bound to the ferric iron on the distal side and serves as an O_2 mimic. The distal ligand is stabilized by a conserved H-bond network involving Y20 Oη-H (position B10), and the NεH2 of Q41 (position E7) and Q45 (position E11). F33 (position CD1) and H68 (position F8, "proximal" histidine) are also highlighted for reference. In the absence of exogenous ligand, the $O\eta^-$ of deprotonated Y20 has been proposed to coordinate the iron.[9,10]

Many additional truncated hemoglobins have been identified and studied since determination of the "2-on-2" CtrHb structure.[13,71–73] Two particularly interesting examples (GlbNs) occur in the cyanobacteria *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 (GlbNs). These GlbNs have high sequence conservation with one another (~59 % identity) and also CtrHb (~45 % identity). However, homology modeling using cyanide-bound CtrHb as a target

supports many key differences in the GlbN heme binding pocket, therefore offering an excellent opportunity for comparative studies.

1.6 Unusual features of GlbNs

Initial in vitro investigations of the GlbNs from *Synechocystis* and *Synechococcus* revealed several interesting features. For example, in the absence of exogenous ligand, GlbNs display *bis*-histidine iron coordination in both Fe^{III} and Fe^{II} states.[74–76] Addition of strong ligands such as CN^- (to Fe^{III} GlbN) or CO (to Fe^{II} GlbN) results in displacement of the distal histidine, conformational rearrangement, and stable binding to generate a structure analogous to CN^- bound CtrHb.[77] A growing number of other globins have been shown to contain *bis*-histidine heme coordination, including the human proteins cytoglobin[78] and neuroglobin[79]. As mentioned previously, *bis*-histidine iron hexacoordination is expected to impart a decreased heme reduction potential, and possibly slow ligand binding rates. *Bis*-histidine ligation in both ferric and ferrous states can also facilitate rapid electron transfer[80] relative to a pentacoordinate Hb (e.g., Mb).

GlbNs also contain an unprecedented feature within the Hb superfamily. Specifically, dithionite reduction of the ferric protein to the ferrous state results in a spontaneous, irreversible modification. Detailed NMR analysis of the product proved that a covalent crosslink forms between a non-coordinating histidine (H117) and the heme 2-vinyl group (forming GlbN-A), reminiscent of

the thioether linkage in *c*-cytochromes.[81,82] Figures 1.14 and 1.15 present two views of the heme binding pocket of ferric *bis*-histidine *Synechococcus* GlbN-A.

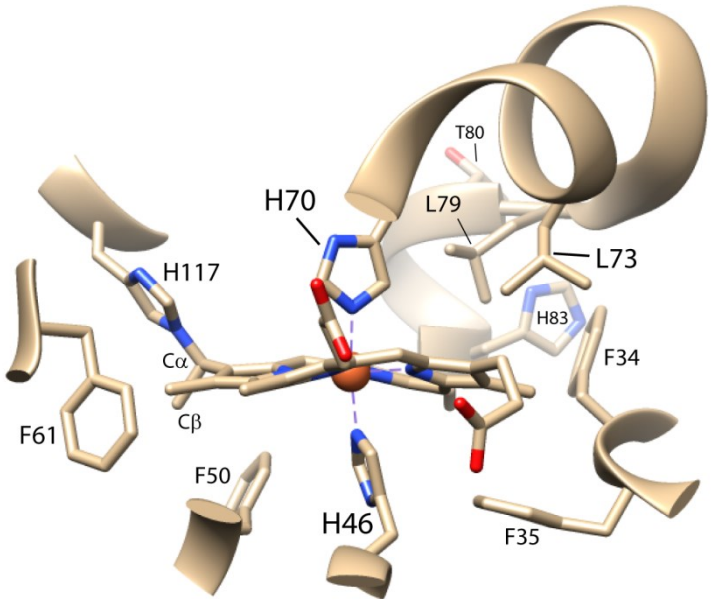


Figure 1.14. Heme pocket determined from the 1.4 Å X-ray model of *Synechococcus* GlbN-A (PDB ID: 4MAX).[83] GlbN-A coordinates the heme iron using two histidines: H70 (position F8) is the tightly bound proximal residue, whereas H46 (position E10) reversibly binds on the distal side and can be displaced by exogenous ligands. GlbN-A contains an unusual covalent bond between H117 Nε2 and the heme 2-Cα, analogous to the thioether linkage characteristic of *c* cytochromes. Several aromatic and nonpolar residues pack tightly with the heme and are also highlighted. Note that the porphyrin ring of GlbN-A deviates significantly from planarity.

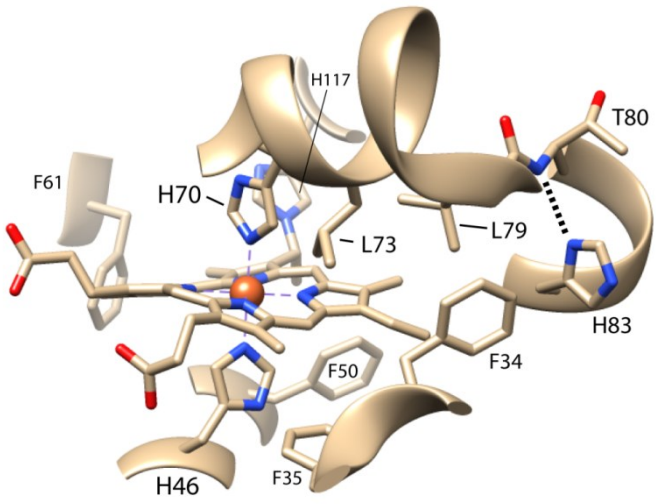


Figure 1.15. Alternate view of the GlbN-A heme pocket (PDB ID: 4MAX). The proximal F-helix (H70) connects to the G-helix via a short FG loop. L73 and L79 are separated by a helical turn and pack against the heme 5-methyl and 4-vinyl groups, respectively. The G-helix is initiated by a conserved N-capping H-bond (marked with black dashed line) between H83 Nδ1 and the N-H of T80.

Although the histidine-heme crosslink in GlbNs causes significant stabilization of the holoprotein, it does not grossly alter its structure.[38,75,84,85] Experiments investigating ligand binding kinetics show modest differences between GlbN (H117A, non-crosslinked) and GlbN-A (crosslinked).[86] In principle, formation of the His-heme posttranslational modification could have multiple effects on GlbN properties (Figure 1.16). Additionally, GlbNs extracted from synechococcal cells grown under microoxic conditions contain covalently attached heme.[87] At this point, however, the functional role for the His-heme PTM is not known.

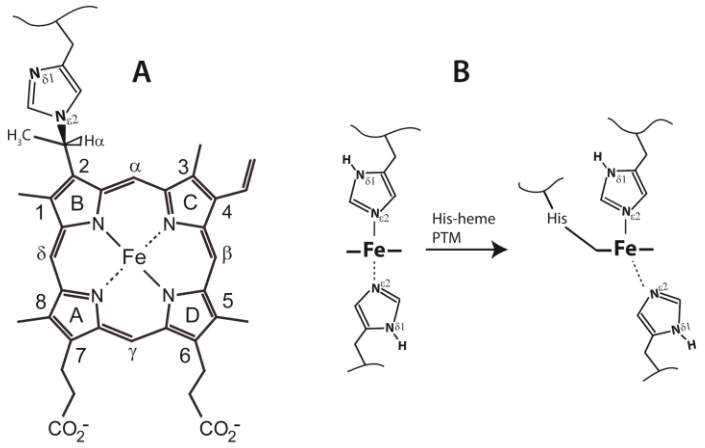


Figure 1.16. A) Structure and nomenclature for the modified *b*/*c* heme observed in GlbN-As from *Synechococcus* and *Synechocystis*. The spontaneous posttranslational modification (PTM) results in saturation of a vinyl group upon reduction to the ferrous state. NMR and crystallographic studies demonstrate the formation of an adduct linking H117 Nε2 to the heme 2-Cα. B) In addition to changing the electronic structure of the heme and enhancing holoprotein stability, the PTM may alter the coordination strength of axial ligands, or modulate reduction potential through His ionization.

1.7 Overview of the dissertation

This dissertation sought to provide answers to three main questions with regard to the posttranslational modification of GlbN:

1. By what mechanism does the ferrous heme-histidine PTM form? As a corollary, why do GlbNs appear to occasionally (and unpredictably) undergo conversion in the ferric state?
2. What are the consequences of His-heme PTM on GlbN chemistry? What is the functional role for heme crosslinking?

3. Can the His-heme crosslink be transplanted into different systems for heme protein engineering purposes? If so, what are the consequences of nonnative crosslinking?

Chapter 2 presents a collaborative effort to understand the His-heme PTM mechanism and its relationship to cytochrome *c* chemistry. Chapter 3 continues this mechanistic work and proposes a comprehensive solution as to how ferric GlbN samples occasionally convert to the crosslinked state. Chapter 4 investigates GlbN functional hypotheses and the consequences of His-heme PTM on GlbN chemistry towards nitric oxide, a common intermediate in cyanobacterial nitrogen metabolism. Chapter 5 initiates the crosslink engineering studies by showing that an additional position within the GlbN heme pocket supports His-heme covalent modification. Chapter 6 is a collaborative chapter that expands this engineering work and demonstrates that the PTM can be successfully implanted into another heme protein. Chapter 7 utilizes sophisticated NMR methodologies to observe directly subtle structural perturbations (H-bond strain) that occur upon forming the nonnative His-heme PTM. Overall, the work provides new insight into the mechanism and consequences of covalent heme binding in GlbNs, and shows that the His-heme linkage may be a promising tool for heme protein engineering.

References

1. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181,** 223–230 (1973).

2. Lippard, S. J. & Berg, J. M. *Principles of Bioinorganic Chemistry*. (University Science Books, 1994).

3. Walker, F. A. Models of the Bis-Histidine-Ligated Electron-Transferring Cytochromes. Comparative Geometric and Electronic Structure of Low-Spin Ferro- and Ferrihemes. *Chem. Rev.* **104,** 589–616 (2004).

4. Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **96,** 2239–2314 (1996).

5. Kendrew, J. C. *et al.* Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185,** 422–427 (1960).

6. Banci, L., Bertini, I., Ferroni, F. & Rosato, A. Solution structure of reduced microsomal rat cytochrome b_5. *Eur. J. Biochem. FEBS* **249,** 270–279 (1997).

7. Arnesano, F., Banci, L., Bertini, I. & Felli, I. C. The solution structure of oxidized rat microsomal cytochrome b_5. *Biochemistry* **37,** 173–184 (1998).

8. Benini, S. *et al.* Crystal structure of oxidized *Bacillus pasteurii* cytochrome c_{553} at 0.97-Å resolution. *Biochemistry* **39,** 13115–13126 (2000).

9. Couture, M. *et al. Chlamydomonas* chloroplast ferrous hemoglobin. Heme pocket structure and reactions with ligands. *J. Biol. Chem.* **274,** 6898–6910 (1999).

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

11. Reid, T. J. *et al.* Structure and heme environment of beef liver catalase at 2.5 Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4767–4771 (1981).
12. Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P_{450}cam. *J. Mol. Biol.* **195,** 687–700 (1987).
13. Johnson, E. A. *et al.* 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 (2014).
14. Fantuzzi, A., Sadeghi, S., Valetti, F., Rossi, G. L. & Gilardi, G. Tuning the reduction potential of engineered cytochrome c_{553}. *Biochemistry* **41,** 8718–8724 (2002).
15. Shifman, J. M., Gibney, B. R., Sharp, R. E. & Dutton, P. L. Heme redox potential control in de novo designed four-alpha-helix bundle proteins. *Biochemistry* **39,** 14813–14821 (2000).
16. Reedy, C. J., Elvekrog, M. M. & Gibney, B. R. Development of a heme protein structure-electrochemical function database. *Nucleic Acids Res.* **36,** D307–313 (2008).
17. Zheng, Z. & Gunner, M. R. Analysis of the electrochemistry of hemes with E(m)s spanning 800 mV. *Proteins* **75,** 719–734 (2009).
18. Poulos, T. L. & Kraut, J. The stereochemistry of peroxidase catalysis. *J. Biol. Chem.* **255,** 8199–8205 (1980).
19. Erman, J. E., Vitello, L. B., Mauro, J. M. & Kraut, J. Detection of an oxyferryl porphyrin π–cation-radical intermediate in the reaction between hydrogen peroxide and a mutant yeast cytochrome c peroxidase. Evidence

for tryptophan-191 involvement in the radical site of compound I. *Biochemistry* **28,** 7992–7995 (1989).

20. Rittle, J. & Green, M. T. Cytochrome P_{450} compound I: capture, characterization, and C-H bond activation kinetics. *Science* **330,** 933–937 (2010).
21. Rittle, J., Younker, J. M. & Green, M. T. Cytochrome P_{450}: the active oxidant and its spectrum. *Inorg. Chem.* **49,** 3610–3617 (2010).
22. Green, M. T. C-H bond activation in heme proteins: the role of thiolate ligation in cytochrome P_{450}. *Curr. Opin. Chem. Biol.* **13,** 84–88 (2009).
23. Yosca, T. H. *et al.* Iron(IV)hydroxide pK(a) and the role of thiolate ligation in C-H bond activation by cytochrome P_{450}. *Science* **342,** 825–829 (2013).
24. Huang, L., Wojciechowski, G. & Ortiz de Montellano, P. R. Role of heme-protein covalent bonds in mammalian peroxidases. Protection of the heme by a single engineered heme-protein link in horseradish peroxidase. *J. Biol. Chem.* **281,** 18983–18988 (2006).
25. Hoch, U. & Ortiz De Montellano, P. R. Covalently linked heme in cytochrome P_{450} 4a fatty acid hydroxylases. *J. Biol. Chem.* **276,** 11339–11346 (2001).
26. Colas, C., Kuo, J. M. & Ortiz de Montellano, P. R. Asp-225 and Glu-375 in autocatalytic attachment of the prosthetic heme group of lactoperoxidase. *J. Biol. Chem.* **277,** 7191–7200 (2002).
27. Ortiz de Montellano, P. R. Mechanism and role of covalent heme binding in the CYP4 family of P450 enzymes and the mammalian peroxidases. *Drug Metab. Rev.* **40,** 405–426 (2008).

28. Iwata, S., Ostermeier, C., Ludwig, B. & Michel, H. Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature* **376,** 660–669 (1995).
29. Daltrop, O., Allen, J. W. A., Willis, A. C. & Ferguson, S. J. In vitro formation of a *c*-type cytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7872–7876 (2002).
30. Daltrop, O., Smith, K. M. & Ferguson, S. J. Stereoselective in vitro formation of *c*-type cytochrome variants from *Hydrogenobacter thermophilus* containing only a single thioether bond. *J. Biol. Chem.* **278,** 24308–24313 (2003).
31. Daltrop, O. & Ferguson, S. J. In vitro studies on thioether bond formation between *Hydrogenobacter thermophilus* apocytochrome c_{552} with metalloprotoporphyrin derivatives. *J. Biol. Chem.* **279,** 45347–45353 (2004).
32. Barker, P. D. *et al.* Transmutation of a heme protein. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 6542–6546 (1993).
33. Stevens, J. M., Uchida, T., Daltrop, O. & Ferguson, S. J. Covalent cofactor attachment to proteins: cytochrome *c* biogenesis. *Biochem. Soc. Trans.* **33,** 792–795 (2005).
34. Sanders, C., Turkarslan, S., Lee, D.-W. & Daldal, F. Cytochrome *c* biogenesis: the Ccm system. *Trends Microbiol.* **18,** 266–274 (2010).
35. Barker, P. D. & Ferguson, S. J. Still a puzzle: why is haem covalently attached in *c*-type cytochromes? *Struct. Lond. Engl. 1993* **7,** R281–290 (1999).
36. Wood, P. M. Why do *c*-type cytochromes exist? *FEBS Lett.* **164,** 223–226 (1983).

37. Tomlinson, E. J. & Ferguson, S. J. Loss of either of the two heme-binding cysteines from a class I *c*-type cytochrome has a surprisingly small effect on physicochemical properties. *J. Biol. Chem.* **275,** 32530–32534 (2000).
38. Vuletich, D. A., Falzone, C. J. & Lecomte, J. T. J. Structural and dynamic repercussions of heme binding and heme-protein cross-linking in *Synechococcus* sp. PCC 7002 hemoglobin. *Biochemistry* **45,** 14075–14084 (2006).
39. Anderson, J. L. R. *et al.* Constructing a man-made *c*-type cytochrome maquette in vivo: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chem. Sci.* **5,** 507 (2014).
40. Iverson, T. M. *et al.* Heme packing motifs revealed by the crystal structure of the tetra-heme cytochrome c_{554} from *Nitrosomonas europaea*. *Nat. Struct. Mol. Biol.* **5,** 1005–1012 (1998).
41. Leys, D. *et al.* Crystal structures at atomic resolution reveal the novel concept of 'electron-harvesting' as a role for the small tetraheme cytochrome *c*. *J.Biol.Chem.* **277,** 35703–35711 (2002).
42. Mowat, C. G. & Chapman, S. K. Multi-heme cytochromes—new structures, new chemistry. *Dalton Trans.* 3381–3389 (2005).
43. Wyman, J. *Binding and Linkage: Functional Chemistry of Biological Macromolecules*. (University Science Books, 1990).
44. Monod, J., Wyman, J. & Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12,** 88–118 (1965).
45. Lehninger, A. L., Nelson, D. L. & Cox, M. M. *Lehninger Principles of Biochemistry*. (Macmillan, 2005).

46. Perutz, M. F. *et al.* Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by X-ray analysis. *Nature* **185,** 416–422 (1960).
47. Perutz, M. F. Stereochemistry of cooperative effects in haemoglobin. *Nature* **228,** 726–739 (1970).
48. Barrick, D., Ho, N. T., Simplaceanu, V., Dahlquist, F. W. & Ho, C. A test of the role of the proximal histidines in the Perutz model for cooperativity in haemoglobin. *Nat. Struct. Mol. Biol.* **4,** 78–83 (1997).
49. Phillips, S. E. Structure and refinement of oxymyoglobin at 1.6 Å resolution. *J.Mol.Biol.* **142,** 531–554 (1981).
50. Egeberg, K. D. *et al.* The role of Val68(E11) in ligand binding to sperm whale myoglobin. Site-directed mutagenesis of a synthetic gene. *J. Biol. Chem.* **265,** 11788–11795 (1990).
51. Carver, T. E. *et al.* Analysis of the kinetic barriers for ligand binding to sperm whale myoglobin using site-directed mutagenesis and laser photolysis techniques. *J. Biol. Chem.* **265,** 20007–20020 (1990).
52. Moffat, K., Deatherage, J. F. & Seybert, D. W. A structural model for the kinetic behavior of hemoglobin. *Science* **206,** 1035–1042 (1979).
53. Vinogradov, S. N. *et al.* Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 11385–11389 (2005).
54. Vinogradov, S. N. & Moens, L. Diversity of globin function: enzymatic, transport, storage, and sensing. *J. Biol. Chem.* **283,** 8773–8777 (2008).
55. Vinogradov, S. N. *et al.* A model of globin evolution. *Gene* **398,** 132–142 (2007).

56. Vinogradov, S. N., Tinajero-Trejo, M., Poole, R. K. & Hoogewijs, D. Bacterial and archaeal globins - a revised perspective. *Biochim. Biophys. Acta* **1834,** 1789–1800 (2013).

57. Gardner, P. R., Gardner, A. M., Martin, L. A. & Salzman, A. L. Nitric oxide dioxygenase: An enzymic function for flavohemoglobin. *Proc. Natl. Acad. Sci.* **95,** 10378–10383 (1998).

58. Gardner, A. M., Martin, L. A., Gardner, P. R., Dou, Y. & Olson, J. S. Steady-state and transient kinetics of *Escherichia coli* nitric-oxide dioxygenase (flavohemoglobin). The B10 tyrosine hydroxyl is essential for dioxygen binding and catalysis. *J. Biol. Chem.* **275,** 12581–12589 (2000).

59. Ilari, A., Bonamore, A., Farina, A., Johnson, K. A. & Boffi, A. The X-ray structure of ferric *Escherichia coli* flavohemoglobin reveals an unexpected geometry of the distal heme pocket. *J. Biol. Chem.* **277,** 23725–23732 (2002).

60. Smagghe, B. J., Trent, J. T. & Hargrove, M. S. NO dioxygenase activity in hemoglobins is ubiquitous in vitro, but limited by reduction in vivo. *PloS One* **3,** e2039 (2008).

61. Liu, X. *et al.* Diffusion-limited reaction of free nitric oxide with erythrocytes. *J. Biol. Chem.* **273,** 18709–18713 (1998).

62. Shiva, S. *et al.* Deoxymyoglobin Is a Nitrite Reductase That Generates Nitric Oxide and Regulates Mitochondrial Respiration. *Circ. Res.* **100,** 654–661 (2007).

63. Gladwin, M. T. & Kim-Shapiro, D. B. The functional nitrite reductase activity of the heme-globins. *Blood* **112,** 2636–2647 (2008).

64. Tiso, M. *et al.* Human neuroglobin functions as a redox-regulated nitrite reductase. *J. Biol. Chem.* **286,** 18277–18289 (2011).

65. Tiso, M., Tejero, J., Kenney, C., Frizzell, S. & Gladwin, M. T. Nitrite reductase activity of non-symbiotic hemoglobins from *Arabidopsis thaliana*. *Biochemistry* **51,** (2012).

66. Nardini, M. *et al.* Archaeal Protoglobin Structure Indicates New Ligand Diffusion Paths and Modulation of Haem-Reactivity. *Embo Rep* **9,** 157–null (2007).

67. Vuletich, D. A. & Lecomte, J. T. J. A phylogenetic and structural analysis of truncated hemoglobins. *J. Mol. Evol.* **62,** 196–210 (2006).

68. Pesce, A. *et al.* A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO J.* **19,** 2424–2434 (2000).

69. Wittenberg, J. B., Bolognesi, M., Wittenberg, B. A. & Guertin, M. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. *J. Biol. Chem.* **277,** 871–874 (2002).

70. Couture, M., Chamberland, H., St-Pierre, B., Lafontaine, J. & Guertin, M. Nuclear genes encoding chloroplast hemoglobins in the unicellular green alga *Chlamydomonas eugametos*. *Mol. Gen. Genet.* **243,** 185–197 (1994).

71. Ascenzi, P., Bolognesi, M., Milani, M., Guertin, M. & Visca, P. Mycobacterial truncated hemoglobins: from genes to functions. *Gene* **398,** 42–51 (2007).

72. Watts, R. A. *et al.* A hemoglobin from plants homologous to truncated hemoglobins of microorganisms. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10119–10124 (2001).

73. Igarashi, J., Kobayashi, K. & Matsuoka, A. A hydrogen-bonding network formed by the B10-E7-E11 residues of a truncated hemoglobin from *Tetrahymena pyriformis* is critical for stability of bound oxygen and nitric oxide detoxification. *J. Biol. Inorg. Chem.* **16,** 599–609 (2011).
74. Couture, M. *et al.* Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC6803 reveal a unique distal heme pocket. *Eur. J. Biochem. FEBS* **267,** 4770–4780 (2000).
75. Falzone, C. J., Christie Vu, B., Scott, N. L. & Lecomte, J. T. J. The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state. *J. Mol. Biol.* **324,** 1015–1029 (2002).
76. Scott, N. L. *et al.* Truncated hemoglobin from the cyanobacterium *Synechococcus* sp. PCC 7002: evidence for hexacoordination and covalent adduct formation in the ferric recombinant protein. *Biochemistry* **41,** 6902–6910 (2002).
77. Trent, J. T., Kundu, S., Hoy, J. A. & Hargrove, M. S. Crystallographic analysis of *Synechocystis* cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. *J. Mol. Biol.* **341,** 1097–1108 (2004).
78. De Sanctis, D. *et al.* Crystal structure of cytoglobin: the fourth globin type discovered in man displays heme hexa-coordination. *J. Mol. Biol.* **336,** 917–927 (2004).
79. Pesce, A. *et al.* Human brain neuroglobin structure reveals a distinct mode of controlling oxygen affinity. *Struct. Lond. Engl. 1993* **11,** 1087–1095 (2003).

80. Weiland, T. R., Kundu, S., Trent, J. T., Hoy, J. A. & Hargrove, M. S. Bis-histidyl hexacoordination in hemoglobins facilitates heme reduction kinetics. *J. Am. Chem. Soc.* **126,** 11930–11935 (2004).
81. Vu, B. C., Jones, A. D. & Lecomte, J. T. J. Novel histidine-heme covalent linkage in a hemoglobin. *J. Am. Chem. Soc.* **124,** 8544–8545 (2002).
82. Vu, B. C., Vuletich, D. A., Kuriakose, S. A., Falzone, C. J. & Lecomte, J. T. J. Characterization of the heme-histidine cross-link in cyanobacterial hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. *J. Biol. Inorg. Chem.* **9,** 183–194 (2004).
83. Wenke, B. B., Lecomte, J. T. J., Héroux, A. & Schlessman, J. L. 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 (2014).
84. Hoy, J. A., Kundu, S., Trent, J. T., Ramaswamy, S. & Hargrove, M. S. The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. *J. Biol. Chem.* **279,** 16535–16542 (2004).
85. Pond, M. P., Majumdar, A. & Lecomte, J. T. J. Influence of heme post-translational modification and distal ligation on the backbone dynamics of a monomeric hemoglobin. *Biochemistry* **51,** 5733–5747 (2012).
86. Vu, B. C., Nothnagel, H. J., Vuletich, D. A., Falzone, C. J. & Lecomte, J. T. J. Cyanide binding to hexacoordinate cyanobacterial hemoglobins: hydrogen-bonding network and heme pocket rearrangement in ferric H117A *Synechocystis* hemoglobin. *Biochemistry* **43,** 12622–12633 (2004).

87. Scott, N. L. *et al.* Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry* **49,** 7000–7011 (2010).

Chapter 2.

Chemical reactivity of *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 hemoglobins: covalent heme attachment and *bis*-histidine coordination

Henry J. Nothnagel, Matthew R. Preimesberger, Matthew P. Pond, Benjamin Y. Winer, Emily M. Adney, and Juliette T.J. Lecomte

Reproduced with permission from J Biol Inorg Chem (2011) 16:539–552

DOI 10.1007/s00775-011-0754-2.

Author contributions:

HJN: designed research, performed anaerobic UV-visible stopped-flow kinetics of heme-crosslinking, proposed mechanism, analyzed data

MRP: designed research, prepared H46L and H117C GlbN variants, conducted H_2O_2 experiments, performed NMR data acquisition and analysis of variants, analyzed data and wrote the paper

MPP: conducted DTT reduction experiments, analyzed data, wrote the paper

BYW/EMA: protein preparation

JTJL: designed research, analyzed data, wrote the paper

Abbreviations: DT, dithionite; DTT, dithiothreitol; GODCAT, glucose oxidase/D-glucose/catalase; Hb, hemoglobin; LB, Luria-Bertani; rGlbN-R, recombinant GlbN reconstituted with *b* heme; rGlbN-A, recombinant GlbN with covalently attached heme; RNS, reactive nitrogen species; ROS, reactive oxygen species; *Synechococcus* 7002, *Synechococcus* sp. PCC 7002; *Synechocystis* 6803, *Synechocystis* sp. PCC 6803.

Abstract

In the absence of exogenous ligand, the hemoglobins from the cyanobacteria *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 coordinate the heme group with two axial histidines (His46 and His70). These globins also form a covalent linkage between the heme 2-vinyl substituent and His117. The in-vitro mechanism of heme attachment to His117 was examined with a combination of site-directed mutagenesis, NMR spectroscopy, and optical spectroscopy. The results supported an electrophilic addition with vinyl protonation being the rate-determining step. Replacement of His117 with a cysteine demonstrated that the reaction could occur with an alternative nucleophile. His46 (distal histidine) was implicated in the specificity of the reaction for the 2-vinyl group as well as protection of the protein from oxidative damage caused by exposure to exogenous H_2O_2.

Introduction

Hemoglobins (Hbs) are small helical proteins containing one *b* heme group per polypeptide chain. The textbook Hb subunit harbors a ferrous iron coordinated by four pyrrole nitrogens from the porphyrin ring and the Nε2 atom of the proximal histidine. In the oxy state, dioxygen completes the iron octahedral geometry, whereas in the deoxy state, the distal site either remains vacant or is occupied by a loosely bound water molecule [1]. The Hb heme pocket is such that reversible binding of dioxygen occurs with low probability of side reactions, such as iron oxidation, that would impair function. Not all Hbs, however, exhibit these structural and chemical properties. A growing number of characterized members of the superfamily are now known to use a protein side chain as an axial ligand to the ferrous iron on the distal side when the partial pressure of dioxygen is low [2-6]. It has also become clear that many bacterial and plant globins associate with dioxygen and other small molecules for enzymatic purposes rather than delivery [7]. Variations in heme structure, environment, and distal iron ligation are all exemplified in the 2/2 (truncated) lineage of the Hb superfamily, a set of proteins for which there is thus far little functional information, but which holds promises of novel heme chemistry [7].

The 2/2 Hbs of the cyanobacteria *Synechococcus* sp. PCC 7002 (*Synechococcus* 7002 hereafter) and *Synechocystis* sp. PCC 6803 (*Synechocystis* 6803) contain 123 amino acids and differ at 50 positions scattered throughout the sequence. Physiological studies of *Synechococcus* 7002 show that its Hb (called GlbN) participates in protection against reactive oxygen/nitrogen species (ROS/RNS), likely peroxynitrite or nitric oxide or both [8]. Similar in vivo

information is not available for *Synechocystis* 6803 GlbN, but a strain of *E. coli* lacking the essential native flavohemoglobin gene (*hmp*) and complemented with the *glbN* gene is viable. Thus, it appears that *Synechocystis* 6803 GlbN can take on the role of the heme domain of flavohemoglobin in converting NO into nitrate [9].

Both GlbNs exhibit *bis*-histidine coordination of the heme iron in the ferrous and ferric resting states [10-11]. Their fold is otherwise typical of 2/2 Hbs belonging to phylogenetic Group I [12]. Surprisingly, the two cyanobacterial GlbNs also undergo a chemical modification by which the heme is irreversibly attached to the protein. The cross-link involves the Cα of the vinyl group on pyrrole B and the Nε2 atom of His117 [13-14] as depicted in Figure 2.1A.

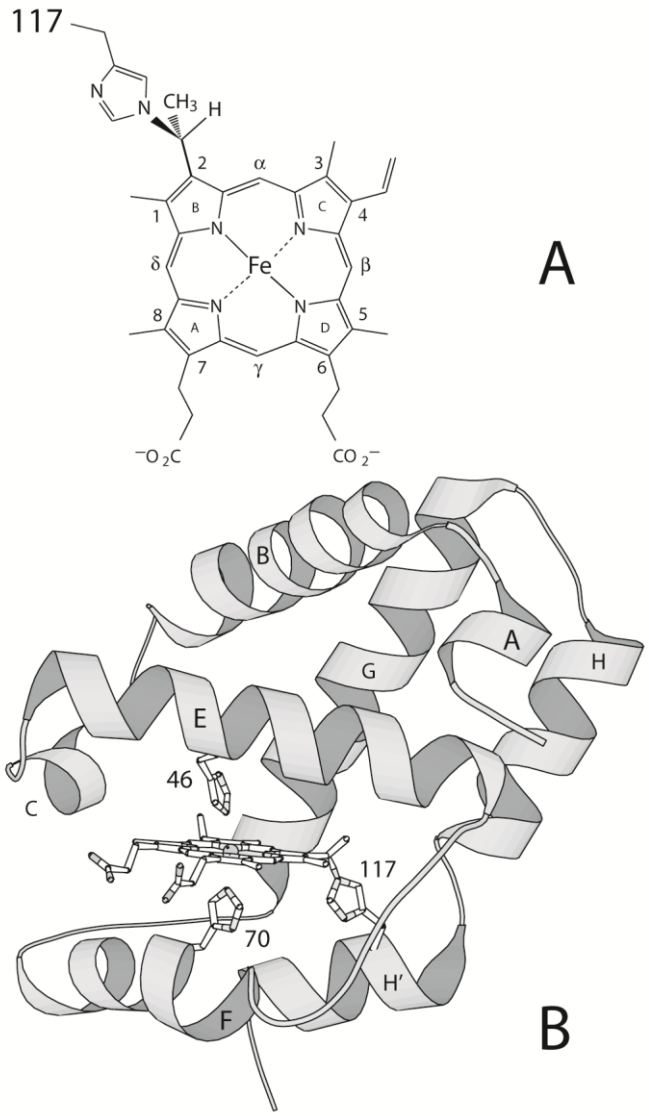


Figure 2.1. (A) The structure of the modified heme with the nomenclature used in the text. The unmodified *b* heme has vinyl substituents at positions 2 (pyrrole B) and 4 (pyrrole C). (B) Ribbon diagram of *Synechococcus* 7002 rGlbN-A (minimized average of NMR ensemble 2KCS). The axial histidines (His46, distal, and His70, proximal) are shown, along with the histidine (His117) forming the covalent adduct to the 2-vinyl. Helices are labeled A, B, C, E, F, G, H and H'.

Wild-type recombinant GlbNs prepared by reconstitution of the apoprotein with Fe(III)-protoporphyrin IX (rGlbN-R hereafter) are efficiently converted to the

adduct (rGlbN-A) by addition of the reducing agent dithionite (DT) [14]. The cross-link appears to have physiological significance as *Synechococcus* 7002 cells contain higher levels of GlbN-A when grown under micro-oxic conditions than when grown under standard oxic conditions [8].

Although post-translational heme vinyl modifications have not been observed in any hemoglobin other than the GlbNs discussed here, heme vinyl groups are moderately reactive and are natural sites for modification in other proteins. The most familiar example is provided by *c*-type cytochromes, which are biosynthesized from apocytochrome and *b* heme and contain one or two thioether bonds to cysteines. Interestingly, the maturation of cytochrome *c* via System I (CcmABCDEFGH) proceeds with transient heme attachment to a histidine or a cysteine of CcmE [15]. The nature of the CcmE histidine adduct has been elucidated: the reversible linkage is between the Nδ1 atom and the heme 2-Cβ [16], unlike in GlbN-A. Non-native attachment of a protein residue at a heme vinyl has also been proposed in the His25SeCys variant of human heme oxygenase-1 (Cβ position [17]), the Ser160Met variant of ascorbate peroxidase under oxidative conditions (Cβ position [18]), and the Ser160Cys variant of the same under reducing conditions [19]. In GlbN, radical and non-radical as well as reductive and oxidative routes are all a priori plausible.

Synechococcus 7002 and *Synechocystis* 6803 GlbNs offer an opportunity to examine the chemistry associated with covalent heme attachment and *bis*-histidyl coordination. These proteins are particularly well suited for analysis because they can be prepared, and are sufficiently stable, in the apoprotein state and the holoprotein state with or without heme covalent attachment.

Here, we investigate the reactivity of GlbNs toward H_2O_2, an ROS generated during aerobic metabolism in cyanobacterial cells, and propose a mechanism for the reduction-driven formation of the heme–protein cross-link. The new data expand our view of the chemistry supported by the globin fold.

Materials and Methods

The procedures for production and purification of recombinant cross-linked and non-cross-linked GlbNs (rGlbN-R and rGlbN-A, respectively) have been published previously [11, 20] and were followed with minor changes. Hemochromogen assay [21-22] and acid–butanone heme extraction [23] were performed as reported in prior work. Details are provided as Supporting Information.

Electronic absorption spectroscopy

Absorbance data were collected on one of three instruments, an AVIV Instruments Model 14-DS spectrophotometer equipped with a Peltier device for temperature control, an Olis RSM-1000 equipped with a stopped-flow rapid mixing apparatus, or a Varian Cary 50. All samples were prepared in degassed buffer, and ranged in concentration from 10 to 70 μM.

rGlbN reduction under micro-oxic conditions

When sensitivity to dioxygen was a concern, the ferrous state was generated in the presence of the glucose oxidase-catalase (GODCAT) system [24]. Glucose

oxidase (*Aspergillus niger*), catalase (bovine liver) and D-(+)-glucose were purchased from Sigma. Typical solutions used for UV-Vis experiments included 300 nM glucose oxidase, 70 nM catalase, and 30 mM D-(+)-glucose in 100mM sodium phosphate pH 7.2. The solutions were allowed to incubate for 1.5 h before the addition of a concentrated protein solution followed by a second hour long incubation period prior to reductant addition. The GODCAT system removes dissolved dioxygen to steady-state concentrations in the 10–20 μM range [25]. Catalase is sensitive to dithiothreitol (DTT) [26], but control experiments did not indicate this to be a concern under our conditions.

H_2O_2 treatment

H_2O_2 (30 % Fischer) was diluted with ddH_2O to produce 1 to 100 mM stock solutions. The concentration of H_2O_2 was determined using the extinction coefficient ε_{240nm} = 39.4 M^{-1} cm^{-1} [27]. Protein solutions contained ~100 mM phosphate, pH ~7.2. For UV-Vis spectrophotometry experiments, protein sample concentrations were 5–10 μM. A five-fold excess of H_2O_2 was added to the protein solution in a stoppered cuvette. After 10–15 s of manual mixing, absorbance spectra were collected on a Varian Cary 50 from 430 nm to 401 nm, with a 0.1 s averaging time, at room temperature, as a function of time. For cyanide protection assays, cyanomet rGlbNs were prepared by incubating proteins with a 5- to 500-fold excess of KCN prior to treatment with H_2O_2.

SDS-PAGE analysis was performed with ~50 μM protein samples. For kinetic experiments, protein samples were treated with 5 eq H_2O_2 and aliquots were extracted as a function of time. At the end of each incubation period, reactions were quenched by addition of SDS sample buffer supplemented with 360 mM β-

mercaptoethanol (final concentration) and boiled for 5 min prior to electrophoresis. In the second type of experiment, samples were exposed to 0.5 eq to 100 eq of H_2O_2 and incubated for 3 h at room temperature prior to quenching, denaturation, and electrophoresis.

Fluorescence emission

A 5- to 100-fold excess of H_2O_2 was added to a ~100 μM protein solution and incubated for 3 h at room temperature. Following incubation, the samples were subjected to an acid-butanone heme extraction. The recovered apoprotein was dialyzed extensively against H_2O to remove any 2-butanone. The protein was diluted to ~2 μM in 100 mM phosphate pH 7.2. Fluorescence data were acquired on an Aviv ATF 105 Automated Titrating Differential/Ratio Spectrofluorometer at 25 °C with excitation at 315 nm. The fluorescence emission spectrum was recorded from 325–500 nm in 1 nm steps with 2 s averaging time.

Formation of the heme–protein cross-link, manual mixing

For manual mixing experiments with DTT and *Synechococcus* 7002 rGlbN, the solid reagent was dissolved in a buffer pre-incubated with GODCAT, as described above, under an N_2 atmosphere and so as to reach a 500x stock. Immediately after the DTT was solubilized, 2 μL was added to a 998 μL solution containing GODCAT-incubated protein in a Quartz cuvette for absorbance measurement. The cuvette was then stoppered, inverted several times, and parafilmed. The cuvette was removed from the nitrogen atmosphere and placed in a Varian Cary 50 spectrophotometer for data collection. Dead times ranged

between 30 and 50 s. The spectra were recorded in the range 700–260 nm in 1-nm steps, with a 0.2 second averaging time, and over a minimum of 24 h.

Manual mixing experiments were also performed with DT and *Synechocystis* 6803 rGlbN-R in the presence and absence of the GODCAT scavenger. Solid DT was dissolved in the appropriate buffer to produce a 1 M stock and was added to yield a 4- to 200-fold excess of DT over protein. DT concentrations were estimated by mass and verified by optical spectroscopy using the extinction coefficient ε_{314nm} = 8 mM^{-1} cm^{-1} [28]. The dead time was 10–30 seconds. The spectra were recorded over the range 500–600 nm with 1-nm steps. Spectra were collected every 20 s. Total acquisition time ranged from 20 min to 2 h.

Formation of the heme-protein cross-link, stopped-flow rapid mixing

To eliminate the possibility that dioxygen was involved in the reaction, kinetic data were collected by stopped-flow rapid mixing with an Olis RSM-1000 UV-Vis spectrometer in an anaerobic chamber. The proteins were wild-type and H117A *Synechocystis* 6803 rGlbN. Dioxygen was removed from 100 mM Tris pH 8.2 through a combination of N_2 sparging and vacuum exposure. Stocks of ~70 μM rGlbN-R were freshly prepared from lyophilized proteins. The initial concentrations were measured with injections of protein stock solutions and buffer containing no DT (ε_{545nm} = 10.67 mM^{-1} cm^{-1}). A ~90 μM DT solution was prepared by dissolving solid DT in deoxygenated buffer. The stock solution was prepared in the anaerobic chamber and transferred to a cuvette with a sealable Teflon valve that was closed before being removed from the chamber. The stocks were quantitated on a HP photodiode array spectrophotometer outside of the anaerobic chamber.

Equal volumes of DT and protein were subjected to rapid stopped-flow mixing. Scans were collected every 0.1 s for 3 min; the wavelength range was 438–662 nm, covered in 1-nm steps. The raw data were subjected to singular value decomposition (SVD) following Henry and Hofrichter [29] using the program Scilab 5.2.2 (http://www.scilab.org/). The abstract kinetic vectors were globally fit with the program Savuka [30].

NMR experiments

NMR samples were prepared in 20 mM phosphate buffer, pH 7.2, at 25 °C (90:10 H_2O:D_2O, or 99.9% D_2O) unless specified otherwise. Protein concentrations were 0.1–5 mM. Samples of CO-bound rGlbN were prepared by saturating the protein solution with CO, then adding a buffered DT solution also saturated with CO. This procedure resulted in samples containing a mixture of rGlbN-R and rGlbN-A.

NMR spectra were collected on Bruker DRX-600 and Avance 600 spectrometers and a Varian Inova 500 spectrometer. ^{1}H chemical shifts were referenced to DSS through the water line with correction for temperature. DQF-COSY, NOESY, WEFT-NOESY, TOCSY, ^{1}H-^{13}C HMQC, and ^{1}H-^{15}N HSQC data were collected as described previously [11, 14, 20].

To establish the nature of the covalent attachment in H46L *Synechococcus* 7002 rGlN-A (see Results section), long-range selective ^{1}H-^{15}N HMQC data were collected on uniformly ^{15}N-labeled ferric protein in the cyanomet state (1 mM protein, pH 7.1 or 6.4, 25 °C) using a 22-ms magnetization transfer delay. Selective 1-ms ^{15}N $\pi/2$ pulses were utilized to minimize backbone excitation. To

determine the structure of the vinyl modification in H117C *Synechocystis* 6803 rGlbN-A, natural abundance ^{1}H-^{13}C non-{^{1}H}-decoupled HSQC data were collected on the ferric protein (3 mM protein, pH* 8.7, 298 K in 99.9% D_2O) using a 1.3-ms transfer time. A 2-ms selective ^{13}C refocusing π pulse was used to achieve band selection in the indirect dimension. ^{1}H-^{13}C coupling was established by omitting the ^{1}H π pulse centered in t_1. Data of the same kind were collected to identify the CH_2D group in the wild-type protein allowed to react in D_2O. In this case, a sensitivity-enhanced version of the pulse sequence was utilized.

NMR data were processed using XWIN-NMR, TOPSPIN (Bruker BioSpin, Rheinstetten, Germany), and NMRPipe [31]; spectra were analyzed with TOPSPIN and Sparky [32].

Results

The NMR structure of *Synechococcus* 7002 rGlbN-A is shown in Figure 2.1B [8]. The 50 amino acid substitutions distinguishing *Synechococcus* 7002 GlbN from *Synechocystis* 6803 GlbN do not change the fold and are responsible for modest variations in thermodynamic stabilities, ligand binding properties, and chemical reactivity. Accordingly, we found that one protein occasionally gave clearer results than the other in selected experiments. However, several conclusions do not depend on the protein choice. These are emphasized throughout.

The role of His46 in the post-translational modification

The sole detectable product of the heme modification caused by DT reduction of ferric wild-type *Synechocystis* 6803 rGlbN-R was identified in prior NMR work as the Markovnikov adduct (R stereochemistry) of His117 onto the heme 2-vinyl [13-14] (Figure 2.1A). When the starting material is ferric H46L rGlbN-R, H46A rGlbN-R, or the cyanomet complex of wild-type rGlbN-R, DT reduction produces the 2-vinyl adduct and a measurable amount of the 4-vinyl adduct [33]. This change in specificity was probed further with *Synechococcus* 7002 rGlbN.

At neutral pH, the electronic absorption spectrum of ferric H46L *Synechococcus* 7002 rGlbN-R (Supporting Information Figure S2.1) was consistent with coordination of a water molecule on the distal side. The ^{1}H NMR spectrum was in agreement as it exhibited the downfield hyperfine-shifted resonances of an aquomet complex (Supporting Information Figure S2.2) and no strongly upfield-shifted signals that could be attributed to the meso protons

of a pentacoordinated species [34]. The pH dependence of the optical spectrum indicated a transition from this state at low pH to a hydroxymet species at high pH. At both extremes of pH, the complexity of the pH titration profile (Supporting Information Figure S2.1) reflected aggregation, heme loss, and denaturation. The thermal denaturation profile monitored by electronic absorption spectroscopy lacked a native state baseline (data not shown). Compared to ferric H46L *Synechocystis* 6803 rGlbN-R, which formed a His–Fe(III)–X complex [33, 35], the *Synechococcus* variant had reduced stability and different propensities for Fe(III) coordination.

Figure 2.2 presents the result of treating various forms of *Synechococcus* 7002 rGlbN-R with DT.

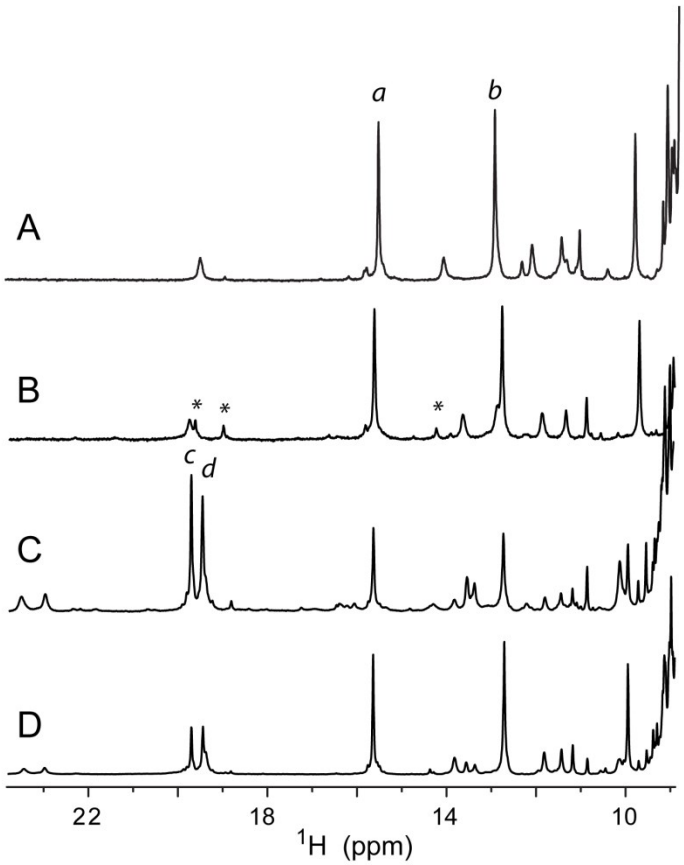


Figure 2.2. ^{1}H NMR spectrum of cyanomet *Synechococcus* 7002 rGlbN-A obtained by reduction of ferric rGlbN-R with DT followed by reoxidation and cyanide binding (downfield hyperfine shifted region). (A) The wild-type protein in its *bis*-histidine resting state reacted at the 2-vinyl position. Peaks *a* and *b* are heme methyl groups. (B) A small amount of an additional product is obtained when the wild-type protein is in the cyanomet state prior to DT treatment (resonances marked with *). (C) The H46L variant reacted at both the 2- and 4-vinyl positions, the adduct at position 4 being the major product. Peaks *c* and *d* are heme methyl groups. (D) The H46L variant in the cyanomet state prior to DT treatment reacted at both the 2- and 4-vinyl positions, the adduct at position 2 being the major product.

In all cases, the iron was re-oxidized with ferricyanide after DT incubation and the products were identified in the cyanomet state so as to compare species with identical His–Fe(III)–CN coordination. Figure 2.2A contains the ^{1}H NMR spectrum of wild-type cyanomet rGlbN-A, which was obtained from starting

material with His–Fe(III)–His coordination. Reaction occurred at the 2-vinyl [14], and this spectrum is used for reference. When the reaction was performed on the cyanomet state of wild-type rGlbN-R (starting material with His–Fe(III)–CN coordination), the resulting spectrum (Figure 2.2B) indicated the formation of a small amount of an additional product. When ferric H46L *Synechococcus* 7002 rGlbN-R was used as the starting material without added exogenous ligand (His–Fe(III)–OH_2/OH^-), the second product was formed in a 2:1 ratio compared to the 2-vinyl adduct (Figure 2.2C). Analysis of homonuclear data and ^{1}H-^{15}N HMQC data (Supporting Information Figure S2.3) confirmed that this second species contained a covalent linkage to His117 Nε2. Cyanomet H46L *Synechococcus* 7002 rGlbN-R (His–Fe(III)–CN) yielded products in a ~1:2.5 ratio in favor of the 2-vinyl adduct (Figure 2.2D).

Homonuclear and ^{1}H-^{13}C HMQC data were used to assign the heme methyls in cyanomet H46L *Synechococcus* 7002 rGlbN-A^2 and rGlbN-A^4 (where the superscript specifies the reacted vinyl). The chemical shifts are provided in Supporting Information Table S2.1 and sample data are shown in Supplementary Information Figure S2.4. The ^{1}H heme methyl order of the A^2 species was found to be 3 > 5 > 8 ~ 1 (from downfield up) and that of the A^4 species, 3 ~ 8 >> 5 > 1. The angle between the N_A–N_B axis (Figure 2.1A) and the projection of the proximal histidine imidazole plane onto the heme plane was estimated for each adduct [36]. The shifts of the A^2 species indicated that the projection of the imidazole plane aligned within a few degrees of the β–δ meso axis as in wild-type rGlbN-R [14]. This is a "staggered" orientation offering minimal steric clash between the porphyrin N atoms and the imidazole ring

CHs. The shifts of the A^4 species indicated that from its rGlbN-R orientation, the heme underwent a 180° rotation about the α–γ axis, followed by an in-plane rotation of ~40°, a rearrangement that brings the 4-vinyl within bonding distance of His117, but also aligns the projection of the proximal histidine imidazole with the porphyrin N_B–N_D axis. This is a strained "eclipsed" orientation [37] expected to impart distinct chemical properties to the two adducts from altered orbital overlap between the histidine nitrogen p_π orbital and the d_π orbital of the heme Fe [38].

As observed in *Synechocystis* 6803 rGlbN [33], the replacement of the distal histidine therefore led to a loss in post-translational modification specificity. In *Synechococcus* 7002 rGlbN, however, the effect was more pronounced, causing a reversal of vinyl preference upon H46L substitution. The 40° in-plane rotation of the heme locates a propionate group at the position occupied by the 8-CH_3 in wild-type rGlbN-R. We hypothesize that the steric hindrance at that site is less severe in *Synechococcus* 7002 rGlbN (Ser65) than in *Synechocystis* 6803 rGlbN (Tyr65), therefore facilitating the addition. Further speculation on the origin of the variable A^2/A^4 ratio is rendered difficult by the complexity of the overall reaction, which involves multiple competing processes, e.g., protein conformational changes, ligand decoordination, and heme reorientation.

Results obtained in the ferrous state in the presence of carbon monoxide provided an additional clue for the influence of the distal ligand on the reactivity of rGlbN-R. When solutions of *Synechocystis* 6803 rGlbN-R were first saturated with CO then reduced with DT, variable amounts of CO-bound rGlbN-A were obtained, which reflected a competition between CO binding and

cross-link formation in the His–Fe(II) or His–Fe(II)–His state. Likewise, *Synechococcus* 7002 rGlbN-R reduced by DT in the presence of CO failed to convert to homogeneous CO-bound rGlbN-A samples. ^{1}H NMR spectra of a *Synechocystis* 6803 rGlbN-R-CO/rGlbN-A-CO mixture are presented in Supporting Information Figure S2.5 with heme assignments in Table S2.2.

His46 coordination and H_2O_2 damage

Many cyanobacteria live under high oxidative stress and possess robust and redundant mechanisms to eliminate ROS and RNS. Heme proteins, unless dedicated to this detoxification, are particularly susceptible to ROS damage. Known mechanisms of resistance involve modes of iron coordination, heme pocket composition, and heme chemical modifications. The extent to which the H_2O_2, a predominant cyanobacterial ROS, destroys rGlbN was therefore inspected. Figure 2.3 displays the results of treatment of ferric wild-type *Synechococcus* 7002 rGlbN-R with a five-fold excess of H_2O_2. Here the optical spectrum was used to detect heme bleaching. Immediately following H_2O_2 addition, there was no evidence for the formation of a hydroxyferryl complex (Compound II) as seen in horse heart myoglobin [39-40]. Over time, the Soret intensity decreased as the heme was degraded. In contrast, when the H46L variant was used, rapid bleaching occurred. Cyanide binding prior to H_2O_2 treatment protected this protein from damage. Similar observations were made with *Synechocystis* 6803 GlbN and its H46L variant (Supporting Information Figure S2.6).

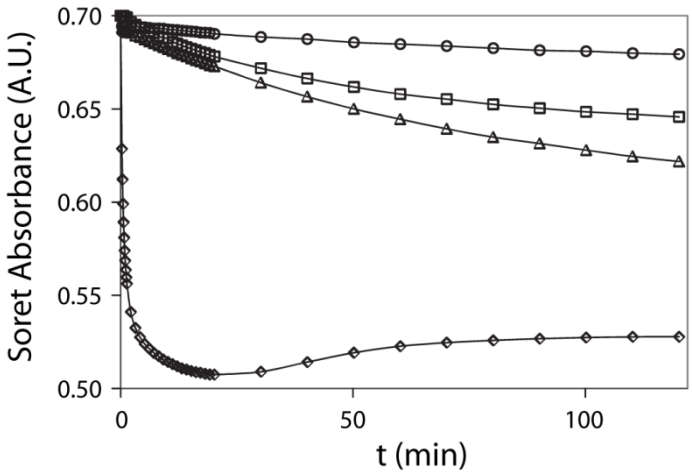


Figure 2.3. Reaction of ferric *Synechococcus* 7002 rGlbN with H_2O_2 monitored by electronic absorption spectroscopy. The absorbance at the Soret maximum is plotted as a function of time after addition of a 5-fold excess of H_2O_2. (□) wild-type rGlbN-R; (△) wild-type rGlbN-A; (◇) H46L rGlbN-R; (○) cyanomet H46L rGlbN-R. The apparent signal recovery of H46L rGlbN-R after 20 min is attributed to scattering. The solid lines are drawn to guide the eye.

Oxidative damage often manifests as intermolecular o,o'-dityrosine cross-links [41]. *Synechocystis* 6803 GlbN and *Synechococcus* 7002 GlbN possess 5 and 3 tyrosines, respectively, and are a priori capable of forming dimers and higher molecular weight cross-linked species. *Synechocystis* 6803 and *Synechococcus* 7002 rGlbN-R samples exposed to H_2O_2 were analyzed by SDS-PAGE. The results are presented in Figure 2.4 for *Synechococcus* 7002 rGlbN and in Supporting Information Figure S2.7 for *Synechocystis* 6803 rGlbN. Regardless of amino acid replacement and cross-linking state, ferric samples treated with H_2O_2 contained species with twice and some thrice the protein molecular weight. Cyanide binding prior to H_2O_2 treatment inhibited oligomer formation presumably by blocking access to the heme iron. Endogenous coordination by

His46 attenuated the damage, but under excess of H_2O_2 was not as effective as bound cyanide.

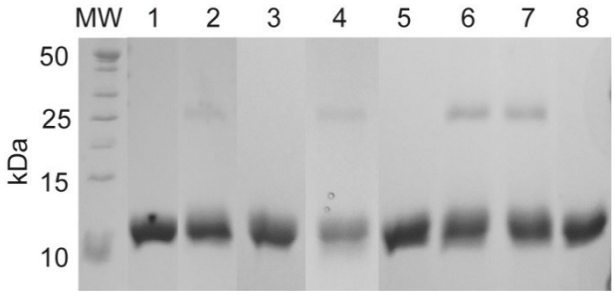


Figure 2.4. Oligomerization caused by H_2O_2 treatment of *Synechococcus* 7002 rGlbN. Protein samples and conditions are: leftmost lane, MW ladder; lane 1, rGlbN-R, no treatment; lane 2, rGlbN-R, 2-h incubation with 5-fold excess H_2O_2; lane 3, rGlbN-A, no treatment; lane 4, rGlbN-A, 2-h incubation with 5-fold excess H_2O_2; lane 5, H46L rGlbN-R, no treatment; lane 6, H46L rGlbN-R, 2-min incubation with 5-fold excess H_2O_2; lane 7, H46L rGlbN-R, 2-h incubation with 5-fold excess H_2O_2; lane 8, cyanomet H46L rGlbN-R, 3-h incubation with 5-fold excess H_2O_2. Note the presence of dimers in lanes 2, 4, 6, and 7.

The o,o'-dityrosine linkage is readily identified by its emission fluorescence spectrum [42], provided that the heme group can be removed from the protein. Following acid-butanone extraction of the heme, the apoprotein samples were analyzed by fluorimetry. The results are particularly clear for *Synechocystis* 6803 GlbN because of higher tyrosine content and higher proportion of oligomeric species with lower concomitant heme degradation. A large increase in fluorescence emission at 410 nm was observed after treatment with H_2O_2, confirming the nature of the product. The spectrum is shown in Supporting Information Figure S2.8.

Protein-based radicals generated in the reaction with H_2O_2 may cause other modifications, for example covalent heme attachment as mentioned above for ascorbate peroxidase [18]. In the preparation of samples for fluorescence analysis, the heme group was successfully extracted from wild-type *Synechocystis* 6803 rGlbN-R after a 3 h treatment with a 5-fold excess H_2O_2 indicating these reactions did not occur appreciably. In the case of wild-type *Synechococcus* 7002 rGlbN and, more obviously, its H46L variant, the aqueous phase retained some coloration after treatment. NMR data collected on the latter after addition of cyanide to the mixture (not shown) confirmed that the colored products were not rGlbN-A^2 or -A^4.

A proposed role for heme adducts in human peroxidases is the prevention of heme degradation by chemistry with reactive intermediates [43]. The ability of the heme–protein cross-link to limit peroxide-induced heme degradation and radical formation was therefore examined. rGlbN-R and rGlbN-A showed comparable responses (Figures 3 and S6), demonstrating that the cross-link did not influence H_2O_2 reactivity significantly.

Reduction with dithiothreitol

Ferric rGlbN-R in solution undergoes the histidine–heme modification unpredictably and over days to months. Addition of DT accelerates the reaction so that wild-type samples are converted on a second to minute time scale at neutral pH. To determine whether the key feature of the DT reaction was the reduction of the heme iron or the involvement of the products of DT oxidation, dithiothreitol (DTT) was used as an alternative reducing agent that produced distinct radical species and had a different mechanism of reaction with

dioxygen [44]. The NMR spectrum of wild-type *Synechocystis* 6803 rGlbN-R incubated with a 20-fold excess of DTT, collected 20–30 min after addition, contained resonances corresponding to ferric rGlbN-A (Figure 2.5). The spectrum also contained resonances from ferric rGlbN-R and additional broad and paramagnetically-shifted peaks. As incubation continued, ferric rGlbN-R signals disappeared, while those of ferric rGlbN-A remained approximately at the same level. Addition of $K_3[Fe(CN)_6]$ to the sample after 12 h converted the sample to the ferric state and only rGlbN-A signals were obtained. The same sequence of events occurred with samples of *Synechococcus* 7002 rGlbN-R. In this case, the broad signals attributed to the binding of DTT or its oxidation byproducts to the ferric protein were stronger and disappeared only when the sample was passed through a Sephadex G-25 desalting column (Supporting Information Figure S2.9).

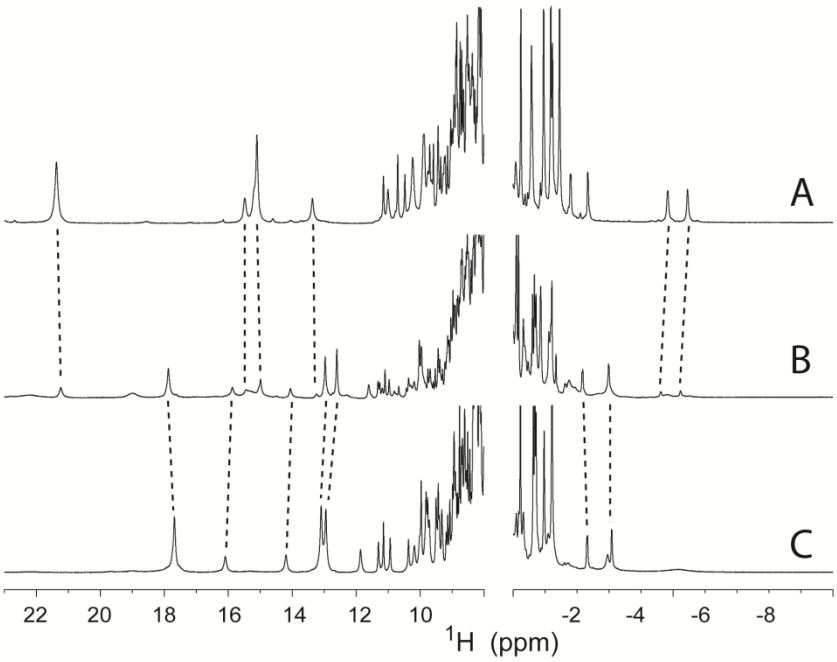


Figure 2.5. 600 MHz ^{1}H 1D NMR spectra of 1 mM wild-type *Synechocystis* 6803 rGlbN (20 mM phosphate buffer pH 7.2, 25 °C, 10 % D_2O). (A) Before DTT addition, the protein is in the ferric rGlbN-R state. (B) 30 min after 20 mM DTT addition, ferric rGlbN-R signals have almost disappeared (indicated by dashed lines from A to B) and the majority of the protein is either in the reduced or the oxidized rGlbN-A state. Note the presence of broad signals attributed to protein containing bound DTT species. (C) After overnight incubation, the protein is entirely in the ferric oxidized rGlbN-A state. Small differences in chemical shifts are due to changing conditions during the reaction.

Requirement for O_2 in the cross-linking reaction

In view of the above results, the need for O_2 in the post-translational modification was inspected by performing the DT reduction under anaerobic conditions. The raw data for a representative experiment (36 μM ferric *Synechocystis* 6803 rGlbN-R, 1.25-fold excess DT, pH 8.2) are shown in Figure 2.6A.

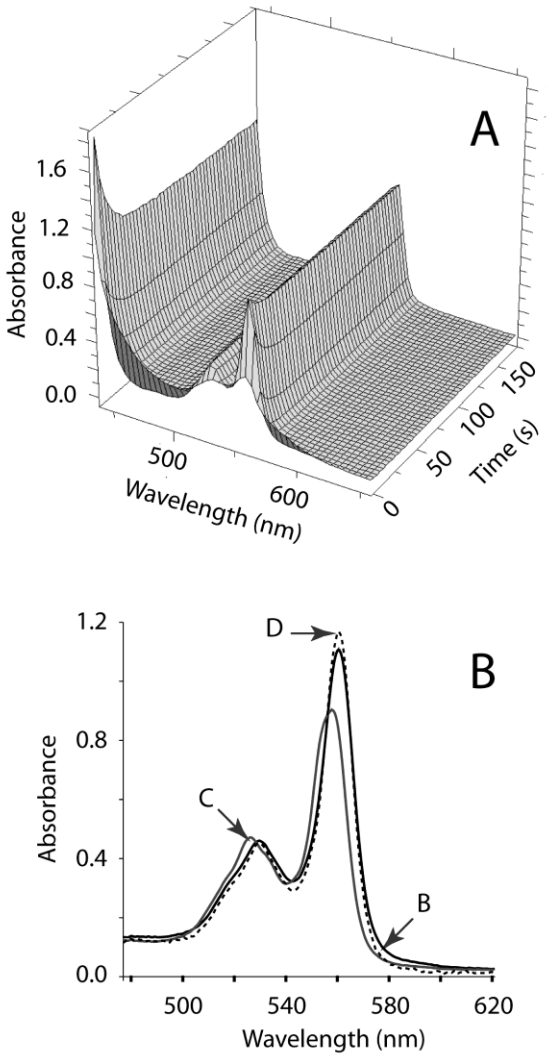


Figure 2.6. (A) UV-Vis spectral results of rapid mixing 36 μM of wild-type *Synechocystis* 6803 rGlbN-R and 45 μM DT under anaerobic conditions. (B) Singular value decomposition of the data retrieved the spectra of ferrous rGlbN-R (B, black, intermediate) and ferrous rGlbN-A (C, grey, final product). The dashed line (D) shows the spectrum of ferrous H117A *Synechocystis* 6803 rGlbN for comparison with GlbN-R.

In the first second of the reaction, the resolved Q bands of a ferrous species grew at 530 nm and 560 nm; these bands then decreased in intensity while bands developed at 526 nm and 558 nm. In this second phase, which was over in ~ 1 min, several isosbestic points were detected. SVD of the time course

followed by global fitting of the data to a sequential A → B → C mechanism returned residuals hinting at a more complex mechanism, likely involving low amplitude rapid phases not well captured by the data collected at short times (200 ms – 1 s). Spectral changes occurring at longer times (t > 1 s) were consistent with a well-defined B → C step proceeding with k_{obs} ~ 0.1 s^{-1}. The spectra of the two reduced species are shown in Figure 2.6B (solid lines). Species C was readily identified as ferrous rGlbN-A by comparison with known spectra.

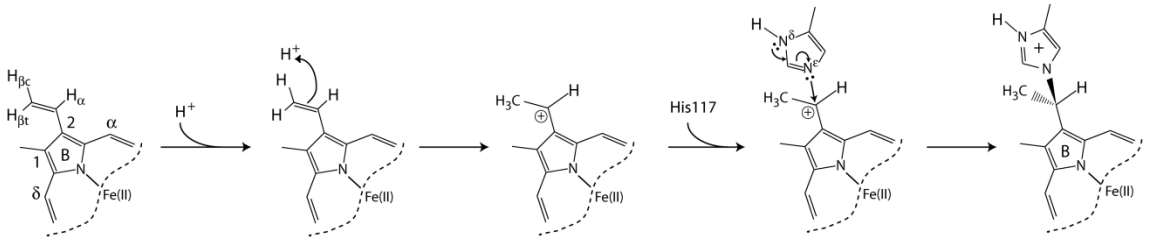
The most noticeable differences between the spectrum of the reduced intermediate and that of ferrous rGlbN-A were a hypsochromic shift and the hypochromicity of the latter. The shift was consistent with saturation of one heme vinyl group. When ferric H117A *Synechocystis* 6803 rGlbN-R was used as the starting material in the stopped-flow experiments, reduction was achieved within 1 s and no further spectral change was observed. The spectrum of ferrous H117A *Synechocystis* 6803 rGlbN-R (Figure 2.6B, dashed line) resembled closely that of the intermediate species and confirmed that the ferrous intermediate (B) was rGlbN-R.

The conversion of *Synechocystis* 6803 rGlbN-R to rGlbN-A was performed by manual mixing at pH 9.5 with various concentrations of DT and in the presence or absence of the GODCAT dioxygen scavenger. Reduction (A → B step) occurred during the dead time of the experiment (10–30 s). Several isosbestic points corresponding to the B → C step were observed (Supporting Information Figure S2.11), and global analysis of the time traces returned a k_{obs} of ~3 × 10^{-3} s^{-1}. Preliminary experiments show this rate constant to be independent of

dithionite or rGlbN concentration and ~40-fold slower than at pH 8.2. Performing the reaction in ^2H_2O at a pD of 9.5 indicated a moderate solvent kinetic isotope effect of ~2.

Proposed mechanism of His-heme cross-link formation

Collectively, the data suggested a cross-linking mechanism (Scheme 2.1) by which the 2-vinyl group is protonated at the Cβ position to generate a Cα carbocation that subsequently undergoes nucleophilic attack by a neutral His117. Reduction of the heme iron would accelerate the reaction by increasing the electron density on the 2-vinyl, and coordination of electron-withdrawing ligands, such as CO, would decelerate or inhibit the reaction. Two aspects of Scheme 2.1 were tested by inspecting the product obtained when the reaction was performed in D_2O or when a cysteine replaced the histidine at position 117.



Scheme 2.1.

Product obtained in D_2O

The carbocation mechanism includes the protonation of the 2-vinyl $C\beta H_2$. Performing the reaction in D_2O was therefore expected to produce a singly or multiply deuterated methyl at the Cβ position. For the purpose of identification, two-dimensional ^{1}H-^{13}C correlation data were collected without ^{1}H decoupling in the indirect dimension. In this experiment, a CH_3 group gives rise to a 3:1:1:3 quartet, whereas a CH_2D group yields a 2:0:2 triplet and a CHD_2 yields a doublet, both further split by deuterium coupling. Figure 2.7 shows the spectral region containing the 2-$^{13}C\beta^{1}H_3$ cross-peak in wild-type *Synechocystis* 6803 rGlbN-A [14]. A triplet, too broad for the deuterium coupling to be resolved, was obtained when the reaction was performed in D_2O. No quartet corresponding to a CH_3 group was detected. Under all chosen conditions, R stereospecificity was maintained at the Cα, and multiple exchange steps resulting in further deuteration of the vinyl group did not occur. In addition, the cross-peak from the 2-CαH was present with intensity comparable to that observed when the reaction was performed in H_2O. Thus, mechanisms involving non-labile hydrogen abstraction were eliminated.

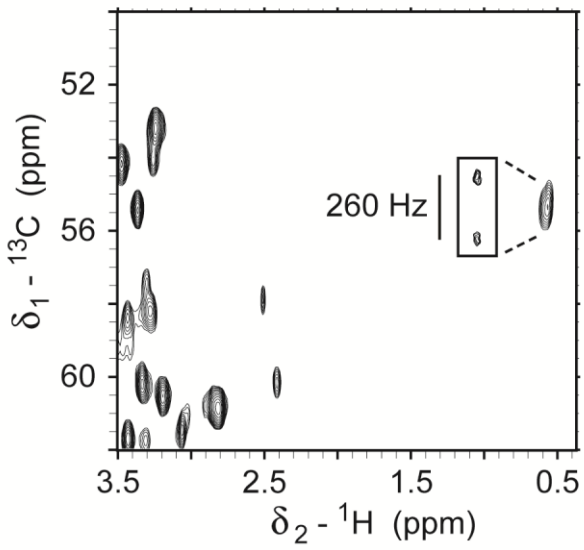


Figure 2.7. ^{1}H-^{13}C correlated data collected on wild-type *Synechocystis* 6803 GlbN-A produced by DT reduction of GlbN-R in D_2O. A region of the HSQC is presented showing the location of the 2-Cβ cross-peak. The inset contains the same in the absence of ^{1}H or ^{2}H decoupling in the ^{13}C dimension. The splitting pattern indicates the formation of a CH_2D group. Data were collected in 99% D_2O, 20 mM phosphate, pH 7.2 and 40 °C at 600 MHz.

Reactivity of H117C rGlbN

His117 was replaced with a cysteine to test the ability of an alternative nucleophile to form the Markovnikov adduct to the 2-vinyl. The ^{1}H spectrum of ferric H117C *Synechocystis* 6803 rGlbN-R was consistent with a minimally perturbed *bis*-histidine complex (Figure 2.8A), as observed for the H117A variant [14]. Incubation of the ferric protein with DT followed by reoxidation resulted in an altered spectrum that indicated chemical modification of the heme (Figure 2.8B).

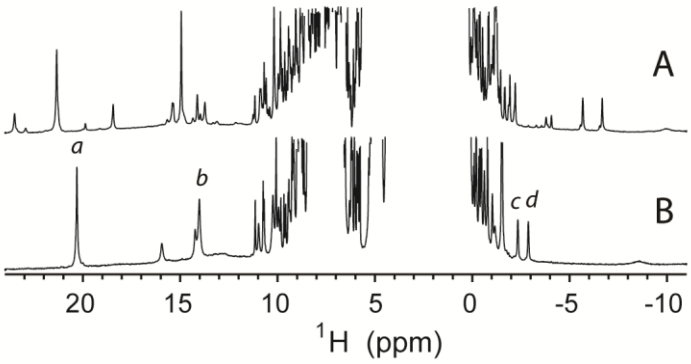


Figure 2.8. 600 MHz ^{1}H 1D spectra of H117C *Synechocystis* 6803 rGlbN in the ferric *bis*-histidine state. (A) rGlbN-R, which resembles the wild-type spectrum and (B), rGlbN-A, obtained after treatment of rGlbN-R with DT and re-oxidation of the iron. The changes in the spectrum indicated that a modification had occurred. Data were collected in 90% H_2O/10% D_2O, 20 mM phosphate, pH 8.7 and 25 °C. Labeled peaks are a, 5-CH_3; b, 1-CH_3; c, 4-vinyl Hβ *trans* and d, 4-vinyl Hβ *cis*. The spectrum in (A) contains a small amount of the minor heme orientational isomer, which nearly disappeared with further equilibration.

The acid–butanone extraction procedure yielded a colored aqueous phase characteristic of a heme group covalently attached to the protein. The pyridine hemochrome spectrum of the product (Supporting Information Figure S2.12) was within 1 nm of that of wild-type rGlbN-A and had an α band with λ_{max} at 554 nm, 3 nm lower than the expected 557 nm of a *b*-heme [47] and comparable to that of the species obtained from a cysteine variant of cytochrome b_5 that contains a single thioether linkage [48].

^{1}H and ^{1}H-^{13}C correlated NMR data (Supporting Information Figure S2.13) were used to identify the signals of the modified heme. They confirmed that the four methyl groups were intact and that the 2-vinyl group had been eliminated. They also identified J-coupled signals at ^{1}H frequencies of –0.1 ppm and 3.2 ppm arising from a porphyrin 2-substituent. A natural

abundance ^{1}H-^{13}C HSQC spectrum collected without decoupling in the indirect dimension revealed the 3:1:1:3 signature of a CH_3 group at –0.1 ppm (Figure 2.9; heme chemical shifts are listed in Supporting Information Table S2.3). Thus, the modified heme contained a –CH–CH_3 moiety at the 2 position of the porphyrin ring.

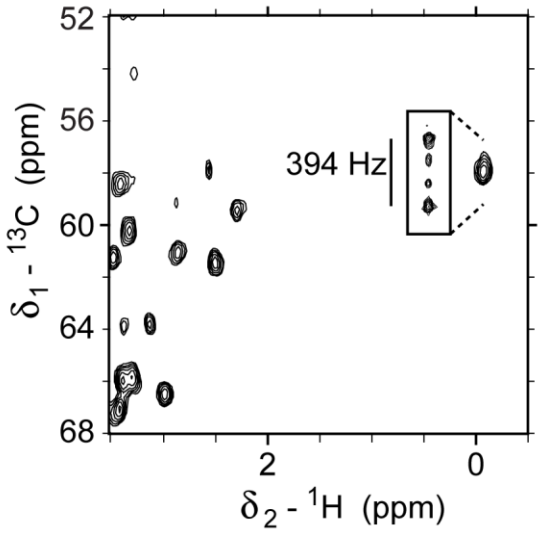


Figure 2.9. ^{1}H-^{13}C correlated data collected on H117C *Synechocystis* 6803 rGlbN-A in the ferric *bis*-histidine state. A portion of the HSQC is shown, which contains the cross-peak of the heme 2-Cβ. The inset contains the same in the absence of decoupling in the ^{13}C dimension. The multiplet structure identifies a methyl group. Data were collected in 99% D_2O, 20 mM phosphate, pH 8.7 and 25 °C at 500 MHz (decoupled data) and 600 MHz (inset).

Modeling Cys117 in the wild-type rGlbN-R structure [49] places the Sγ atom at > 3 Å from the heme 2-Cβ atom and at a greater average distance yet from the heme 2-Cα atom. Because a structural rearrangement is required for a cysteine to reach the heme 2-vinyl Cα, and cysteines are more prone to radical

formation than histidines, the absence of product containing a linkage at Cβ lent further support to Scheme 2.1.

When H117C *Synechococcus* 7002 rGlbN-R was used as the starting material for DT reduction, heme modification occurred as well (Supporting Information Figure S2.14). However, the reactivity appeared to be lower than in H117C *Synechocystis* 6803 rGlbN-R, and samples of modified H117C *Synechococcus* 7002 rGlbN suitable for complete spectral analysis were not obtained. For this protein as well, the distance of closest approach between a modeled Cys117 Sγ and the heme 2-vinyl Cα exceeds 3 Å in the GlbN-R and -A structures. It is possible that the structural fluctuations that allowed Cys117 to reach the 2-vinyl group in *Synechocystis* 6803 rGlbN were less effective in *Synechococcus* 7002 rGlbN. Alternatively, the thiol group may be involved in different interactions in the two proteins.

As an additional test for a radical reaction involving Compound I (an oxoferryl porphyrin π-cation radical) and a thiolate group [48], ferric H117C and H117A rGlbN-Rs were treated with H_2O_2 and their absorbance spectrum examined. As in the wild-type proteins, bleaching was the principal consequence of incubation with 5-fold excess of the peroxide. The spectral changes associated with the formation of rGlbN-A were not observed.

Propionates and vinyls in rGlbN-R

The stereo-electronic properties of the attacked vinyl depend on its orientation with respect to the porphyrin plane [50]. In the reduced state with

CO bound, the wild-type rGlbNs have intra-heme NOEs in agreement with the "trans" orientation of both vinyls: strong effect between the 2-α-H and the α-meso (4-α-H and β-meso), and medium-to-small effect between the 2-β-H_t and 1-CH_3 (4-β-H_t and 3-CH_3). The same holds for the ferric *bis*-histidine states. The temperature response of the chemical shift provides additional information in these low-spin complexes [51]. A plot of 2-vinyl ^1H chemical shifts vs. inverse temperature (Curie plot) is presented for both *Synechococcus* 7002 rGlbN-R and *Synechocystis* 6803 rGlbN-R in Supporting Information Figure S2.15. The trends were linear over the narrow range of temperatures allowed in the experiment and indicated little mobility of the groups.

The electron density of the heme group can also be modulated by the interactions of the propionates with the protein [52-54]. In *Synechocystis* 6803 rGlbN [55-56] and presumably in *Synechococcus* 7002 rGlbN as well, the 6-propionate is exposed to solvent, whereas the 7-propionate is engaged in electrostatic interactions and H-bonds. These features are sensitive to the nature of the distal ligand to the heme [56]. Although a complete analysis of propionate interactions is beyond the scope of this report, it should be noted that in all complexes that have been inspected thus far by NMR [14], including cyanomet H46L *Synechococcus* 7002 rGlbN-A[2] (Figure S2.4), the 6-propionate $^{13}C\alpha$ signal is systematically ~12 ppm upfield of the 7-propionate $^{13}C\alpha$ signal. Considering that the orientation of the proximal histidine is along the β–δ meso axis, this disparity in contact shift may be linked to a biased electronic distribution.

Discussion

Hemoglobins with *bis*-histidine coordination have been the subject of studies seeking to establish their standing in the evolution of the hemoglobin superfamily [57-59] and to determine the implications of their coordination scheme [5-6, 35, 60-62]. Human neuroglobin [63] and possibly α hemoglobin in complex with the chaperone α-hemoglobin-stabilizing protein (AHSP) [64] illustrate that blocking the distal site with a histidine can limit the reactivity of the heme group, protecting it from damage and eliminating side reactions generating ROS. Replacement of the distal histidine of *Synechocystis* 6803 GlbN and *Synechococcus* 7002 GlbN, whether -R or -A, with a weaker ligand resulted in increased H_2O_2 damage and confirmed that some measure of protection was conveyed by the *bis*-histidine scheme.

Physiological data obtained on *Synechococcus* 7002 show that GlbN imparts substantial resistance to the oxidative/nitrosative stress caused by high nitrate concentrations as the source of nitrogen in the growth medium [8]. The attenuated susceptibility to damage provided by distal histidine coordination may therefore be advantageous to GlbN in the cyanobacterial cell. However, at least one other important consequence of this distal coordination was noted in vitro: without it, the post-translational modification of the heme group yielded not one but two products that differed in the reacted vinyl group, the orientation of the heme in its cavity, and presumably chemical reactivity.

Cross-linking in *Synechocystis* 7002 and *Synechococcus* 6803 GlbNs may be necessary if, during their function, these proteins visit states of low heme

affinity. This is expected of the ferrous protein [46] and in complexes with certain exogenous ligands (e.g., NO· [65]). Interestingly, the levels of heme oxygenase 2 increase significantly under microaerobic conditions in *Synechocystis* 6803 [66-67] and heme attachment would prevent heme degradation. The fact that Cys117 can add to the 2-vinyl efficiently, at least in *Synechocystis* 6803 GlbN, confirmed that a histidine is not the only solution to the irreversible attachment of the heme to GlbN. Histidine is generally less reactive than cysteine and may be better suited if the protein must reside in a cellular environment where S-nitrosylation and other reactions could block the thiol functionality.

Regarding the mechanistic aspects of the post-translational modification, the results obtained with H46L *Synechococcus* 7002 rGlbN-R indicated that a pentacoordinate ferrous state (or a ferrous state with loosely associated water) was reactive. The nature of the wild-type species undergoing the addition, however, remains undetermined. In the ferric state, we have found no evidence for the population of pentacoordinate wild-type GlbN complexes, but reduction of the heme iron will weaken the distal histidine ligation bond [68-69], and decoordination may occur. The optical spectra of ferrous H117A *Synechococcus* 7002 rGlbN-R and *Synechocystis* 6803 rGlbN-R were inspected using the empirical correlation between the fractional population of *bis*-histidine complex and the ratio of the absorbance at 555 nm and 540 nm reported by Hargrove and colleagues [69]. The ratio was ~2.3 for ferrous H117A *Synechocystis* 6803 rGlbN-R and only ~1.6 for ferrous H117A *Synechococcus* 7002 rGlbN-R. These values correspond to ~100% and ~40% *bis*-histidine hexacoordination, respectively, and it is possible that the same trend holds in wild-type rGlbN-R.

The rates of His46 decoordination (14 s^{-1}, [70]) and recoordination (4,200 s^{-1}, [35]) in ferrous *Synechocystis* 6803 rGlbN-A are relatively fast. If similar rates apply in GlbN-R, the mechanism of cross-linking could include fast pre-equilibrium between hexa- and pentacoordinated species. A pentacoordinate, His–Fe(II), reactive species may also be formed when cyanomet GlbN is reduced.

The reductive electrophilic addition of Scheme 2.1 has been proposed for the biosynthesis of the histidine–heme linkage in the cytochrome *c* maturation protein CcmE [71] before it was determined that this linkage was of the Nδ1–Cβ type [16] rather than the Nε2–Cα type as in GlbN. In vitro formation of thioether bonds in *c*-type cytochromes is also thought to be initiated by protonation of the vinyls [48, 72-73], the reactive cysteines serving as proton donors. The evidence presented here for GlbN supports the non-radical electrophilic route for several reasons. rGlbN-A^2 (addition at Cα) is produced at the exclusion of other species under a large number of conditions, which include 1) anaerobic reduction with DT, 2) microaerobic (i.e., substoichiometric O_2) reduction with DTT, and 3) His117 replacement with cysteine. In addition, rGlbN-R did not yield rGlbN-A when H_2O_2 was added to the ferric state. Preliminary kinetic analysis demonstrated that proton transfer contributed to the rate determining step. His117 has a pK_a of 6.9 in ferric *Synechocystis* 6803 rGlbN-R [14] and a pK_a of 6.1 in ferric *Synechococcus* 7002 rGlbN-R [8]. Assuming that a similar pK_a difference persists in the ferrous state, comparative rate measurements as a function of pH will aid in the identification of the proton donor. These experiments are in progress.

What are the factors that control the cross-linking reaction? A sine-qua-non condition is that nucleophile and vinyl group must be able to reach each other. The solution structures of *Synechocystis* 6803 rGlbN-R [49] and *Synechococcus* 7002 rGlbN-R [8] show that His117 points toward solvent most of the time, and that a change in rotameric state suffices to direct the imidazole ring to the 2-vinyl. Although the differences between the two rGlbNs were exacerbated with the shorter Cys117 side chain and suggested that a key structural or dynamic property of the nucleophile's local environment was important, the geometrical constraints on the nucleophile and vinyl Cα did not appear particularly difficult to satisfy. Paoli and coworkers have inspected the three-dimensional structure of multiple *b* heme proteins and found variability in the composition of the heme edge [74], which emphasizes that cysteines and histidines can a priori be accommodated near vinyl groups.

The production of rGlbN-A^2 and rGlbN-A^4 with identical Nε2–Cα linkages starting from what must be states with different vinyl packing and heme environment argues for broad applicability of Scheme 2.1 to other heme proteins. Yet, this type of post-translational modification is rarely observed, and additional determinants of reactivity are likely at work. The vinyl groups of rGlbN were ordinary in geometry (*trans* orientation) and mobility (linear Curie behavior) and did not stand out as candidates. However, subtle influences, such as heme distortions [75] and heme propionate interactions [53-54], and less subtle influences, such as axial ligand nature, affect the heme electron density and can restrict the number of proteins and states in which the reaction is possible.

Not surprisingly, we have obtained evidence that oxidation and ligation state alter the reactivity of rGlbN. These two factors may explain the low levels of cross-linked GlbN in *Synechococcus* cells grown under aerobic conditions [8] either because of oxygen binding or iron oxidation. It is interesting that the heme group in the X-ray structures of *Synechocystis* 6803 rGlbN-A and H117A rGlbN shows out-of-plane distortions that differ from the doming typically observed in globins. Ruffling is pronounced in cyanomet and azidomet rGlbN-A, with magnitudes comparable to those in *c*-type cytochromes [75], and, as mentioned in the results, the interactions of the heme propionates with the protein are asymmetric. The consequences of these various features are not readily deconvoluted from the rGlbN data, but will be worth investigating experimentally and computationally with other rGlbN variants and modified hemes.

The two thioether bonds in *c*-type cytochromes are formed from a CXXCH motif, the role of which has been discussed extensively [73, 76-77]. An XXXCH signature, leading to the formation of a single thioether bond, is recognizable in mitochondrial cytochromes *c* and c_1 from trypanosomatids and euglenids [78]. Single thioether bonds to heme are also found in certain b_6f cytochromes [79], but their sequence motif is less distinctive. In prior work [8], we have provided a list of five additional Group I 2/2 Hbs with high sequence homology to the GlbNs discussed here. These proteins are expected to exhibit *bis*-histidine coordination and to undergo the His–heme modification. The latter inference is based on the presence of a histidine at an appropriate topological position and not on local H-helix sequence, which is quite variable. Whether a labile 6th axial ligand is an essential feature of proteins capable of the post-translational

modification remains to be explored, but in any case, no simple primary structure criterion is likely to emerge for the formation of His–heme bonds. Our results caution that chemical characterization will be required to ascertain bond formation even for proteins with high structural homology near heme vinyl groups, but they also suggest that the modification may be more common than currently thought.

In conclusion, we presented evidence that the His–heme covalent bond in reduced GlbN is formed by an electrophilic addition. We propose that a similar addition occurs in vitro in *c*-type cytochromes and in vivo in *Synechocystis* 6803 and *Synechococcus* 7002 cells. The ease with which various forms of rGlbN can be prepared will allow for a further examination of the reaction and the definition of the most important determinants of reactivity for this post-translational modification.

Acknowledgement

This study was supported by National Science Foundation grants MCB-0349409 and MCB-0843439. EMA and BYW were supported by NIH grant 2T32 GM008403-21 (Draper, D., PI). Nancy L. Scott, Dr. B. Christie Vu, Dr. Christopher J. Falzone, and Dr. David A. Vuletich participated in the early stages of the study. Additional help is gratefully acknowledged: Dr. Ananya Majumdar for NMR data collection, Dr. Valeryi Smirnov for anaerobic experiments, Aaron Robinson for fluorescence data collection, and Dr. Justine Roth for the use of her laboratory and helpful discussions.

References

1. Fermi G, Perutz MF, Shaanan B, Fourme R (1984) J. Mol. Biol. 175:159-174

2. Das TK, Couture M, Lee HC, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB, Guertin M (1999) Biochemistry 38:15360-15368

3. Fago A, Hundahl C, Malte H, Weber RE (2004) IUBMB Life 56:689-696

4. de Sanctis D, Dewilde S, Vonrhein C, Pesce A, Moens L, Ascenzi P, Hankeln T, Burmester T, Ponassi M, Nardini M, Bolognesi M (2005) J. Biol. Chem. 280:27222-27229

5. Hoy JA, Hargrove MS (2008) Plant Physiol. Biochem. 46:371-379

6. Pesce A, Thijs L, Nardini M, Desmet F, Sisinni L, Gourlay L, Bolli A, Coletta M, Van Doorslaer S, Wan X, Alam M, Ascenzi P, Moens L, Bolognesi M, Dewilde S (2009) J. Mol. Biol. 386:246-260

7. Vinogradov SN, Moens L (2008) J. Biol. Chem. 283:8773-8777

8. Scott NL, Xu Y, Shen G, Vuletich DA, Falzone CJ, Li Z, Ludwig M, Pond MP, Preimesberger MR, Bryant DA, Lecomte JTJ (2010) Biochemistry 49:7000-7011

9. Smagghe BJ, Trent JT, 3rd, Hargrove MS (2008) PLoS ONE 3:e2039

10. Couture M, Das TK, Savard PY, Ouellet Y, Wittenberg JB, Wittenberg BA, Rousseau DL, Guertin M (2000) Eur. J. Biochem. 267:4770-4780

11. Scott NL, Falzone CJ, Vuletich DA, Zhao J, Bryant DA, Lecomte JTJ (2002) Biochemistry 41:6902-6910

12. Pesce A, Couture M, Dewilde S, Guertin M, Yamauchi K, Ascenzi P, Moens L, Bolognesi M (2000) EMBO J. 19:2424-2434

13. Vu BC, Jones AD, Lecomte JTJ (2002) J. Am. Chem. Soc. 124:8544-8545

14. Vu BC, Vuletich DA, Kuriakose SA, Falzone CJ, Lecomte JTJ (2004) J. Biol. Inorg. Chem. 9:183-194

15. Goddard AD, Stevens JM, Rao F, Mavridou DA, Chan W, Richardson DJ, Allen JW, Ferguson SJ (2010) J. Biol. Chem. 285:22882-22889

16. Lee D, Pervushin K, Bischof D, Braun M, Thöny-Meyer L (2005) J. Am. Chem. Soc. 127:3716-3717

17. Jiang Y, Trnka MJ, Medzihradszky KF, Ouellet H, Wang Y, Ortiz de Montellano PR (2009) J. Inorg. Biochem. 103:316-325

18. Metcalfe CL, Ott M, Patel N, Singh K, Mistry SC, Goff HM, Raven EL (2004) J. Am. Chem. Soc. 126:16242-16248

19. Metcalfe CL, Daltrop O, Ferguson SJ, Raven EL (2007) Biochem. J. 408:355-361

20. Scott NL, Lecomte JTJ (2000) Protein Sci. 9:587-597

21. de Duve C (1948) Acta Chem. Scan. 2:264-289

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

23. Teale FWJ (1959) Biochim. Biophys. Acta 35:543

24. Englander SW, Calhoun DB, Englander JJ (1987) Anal. Biochem. 161:300-306

25. Aitken CE, Marshall RA, Puglisi JD (2008) Biophys. J. 94:1826-1835

26. Sun Y, Oberley LW (1989) Free Radic. Biol. Med. 7:595-602

27. Nelson DP, Kiesow LA (1972) Anal. Biochem. 49:474-478

28. Di Iorio EE (1981) Methods Enzymol 76:57-72

29. Henry ER, Hofrichter J (1992) Methods Enzymol. 210:129-192

30. Bilsel O, Zitzewitz JA, Bowers KE, Matthews CR (1999) Biochemistry 38:1018-1029

31. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) J. Biomol. NMR 6:277-293

32. Goddard TD, Kneller DG (2006) University of California, San Francisco

33. Nothnagel HJ, Love N, Lecomte JT (2009) J. Inorg. Biochem. 103:107-116

34. La Mar GN, Satterlee JD, de Ropp JS (2000) In: Smith KM, Kadish K, Guilard R (eds) The Porphyrin Handbook. Academic Press, Burlington, MA, pp. 185-298

35. Hvitved AN, Trent JT, 3rd, Premer SA, Hargrove MS (2001) J. Biol. Chem. 276:34714-34721

36. Bertini I, Luchinat C, Parigi G, Walker FA (1999) J. Biol. Inorg. Chem. 4:515-519

37. Samuni U, Ouellet Y, Guertin M, Friedman JM, Yeh SR (2004) J. Am. Chem. Soc. 126:2682-2683

38. Nonaka D, Wariishi H, Fujii H (2009) Biochemistry 48:898-905

39. Egawa T, Shimada H, Ishimura Y (2000) J. Biol. Chem. 275:34858-34866

40. Cooper CE, Jurd M, Nicholls P, Wankasi MM, Svistunenko DA, Reeder BJ, Wilson MT (2005) Dalton Trans.:3483-3488

41. Lardinois OM, de Montellano PR (2001) J. Biol. Chem. 276:23186-23191

42. Malencik DA, Anderson SR (2003) Amino Acids 25:233-247

43. Huang L, Wojciechowski G, Ortiz de Montellano PR (2006) J Biol Chem 281:18983-18988

44. Usha Devi S, Ramasarma T (1987) Mol. Cell. Biochem. 77:111-120

45. Netto LE, Stadtman ER (1996) Arch. Biochem. Biophys. 333:233-242

46. Hoy JA, Smagghe BJ, Halder P, Hargrove MS (2007) Protein Sci. 16:250-260

47. Paul KG, Theorell H, Akeson A (1953) Acta Chem. Scan. 7:1284-1287

48. Barker PD, Ferrer JC, Mylrajan M, Loehr TM, Feng R, Konishi Y, Funk WD, MacGillivray RT, Mauk AG (1993) Proc Natl Acad Sci U S A 90:6542-6546

49. Falzone CJ, Vu BC, Scott NL, Lecomte JTJ (2002) J. Mol. Biol. 324:1015-1029

50. Marzocchi MP, Smulevich G (2003) J. Raman Spect. 34:725-736

51. Satterlee JD, Erman JE (1983) J. Biol. Chem. 258:1050-1056

52. Barrows TP, Poulos TL (2005) Biochemistry 44:14062-14068

53. Guallar V, Olsen B (2006) J. Inorg. Biochem. 100:755-760

54. Xu C, Ibrahim M, Spiro TG (2008) Biochemistry 47:2379-2387

55. Hoy JA, Kundu S, Trent JT, 3rd, Ramaswamy S, Hargrove MS (2004) J. Biol. Chem. 279:16535-16542

56. Trent JT, 3rd, Kundu S, Hoy JA, Hargrove MS (2004) J. Mol. Biol. 341:1097-1108

57. Trevaskis B, Watts RA, Andersson CR, Llewellyn DJ, Hargrove MS, Olson JS, Dennis ES, Peacock WJ (1997) Proc. Natl. Acad. Sci. U. S. A. 94:12230-12234

58. Burmester T, Ebner B, Weich B, Hankeln T (2002) Mol. Biol. Evol. 19:416-421

59. Sturms R, Kakar S, Trent J, 3rd, Hargrove MS (2010) Biochemistry 49:4085-4093

60. Hargrove MS (2000) Biophys. J. 79:2733-2738

61. Weiland TR, Kundu S, Trent JT, 3rd, Hoy JA, Hargrove MS (2004) J. Am. Chem. Soc. 126:11930-11935

62. Yoon J, Herzik MA, Winter MB, Tran R, Olea C, Marletta MA (2010) Biochemistry

63. Lardinois OM, Tomer KB, Mason RP, Deterding LJ (2008) Biochemistry 47:10440-10448

64. Hamdane D, Vasseur-Godbillon C, Baudin-Creuza V, Hoa GH, Marden MC (2007) J. Biol. Chem. 282:6398-6404

65. Goodrich LE, Paulat F, Praneeth VK, Lehnert N (2010) Inorg. Chem. 49:6293-6316

66. Minamizaki K, Mizoguchi T, Goto T, Tamiaki H, Fujita Y (2008) J. Biol. Chem. 283:2684-2692

67. Yilmaz M, Kang I, Beale SI (2010) Photosynth Res 103:47-59

68. Cowley AB, Kennedy ML, Silchenko S, Lukat-Rodgers GS, Rodgers KR, Benson DR (2006) Inorg. Chem. 45:9985-10001

69. Smagghe BJ, Halder P, Hargrove MS (2008) Methods Enzymol. 436:359-378

70. Smagghe BJ, Sarath G, Ross E, Hilbert JL, Hargrove MS (2006) Biochemistry 45:561-570

71. Stevens JM, Daltrop O, Higham CW, Ferguson SJ (2003) J. Biol. Chem. 278:20500-20506

72. Daltrop O, Allen JW, Willis AC, Ferguson SJ (2002) Proc Natl Acad Sci U S A 99:7872-7876

73. Kranz RG, Richard-Fogal C, Taylor JS, Frawley ER (2009) Microbiol. Mol. Biol. Rev. 73:510-528

74. Schneider S, Marles-Wright J, Sharp KH, Paoli M (2007) Nat. Prod. Rep. 24:621-630

75. Shelnutt JA, Song XZ, Ma JG, Jia SL, Jentzen W, Medforth CJ (1998) Chem. Soc. Rev. 27:31-41

76. Cowley AB, Lukat-Rodgers GS, Rodgers KR, Benson DR (2004) Biochemistry 43:1656-1666

77. Bowman SE, Bren KL (2010) Inorg. Chem. 49:7890-7897

78. Allen JW, Ginger ML, Ferguson SJ (2004) Biochem. J. 383:537-542

79. Kurisu G, Zhang H, Smith JL, Cramer WA (2003) Science 302:1009-1014

Supporting Information for

Chemical reactivity of *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 hemoglobins: covalent heme attachment and *bis*-histidine coordination

Henry J. Nothnagel, Matthew R. Preimesberger, Matthew P. Pond, Benjamin Y. Winer, Emily M. Adney, and Juliette T.J. Lecomte

Contents

Figure S2.2: ^{1}H NMR spectrum of ferric H46L *Synechococcus* 7002 and *Synechocystis* 6803 rGlbN-R without added exogenous ligand.

Figure S2.3: ^{1}H-^{15}N HMQC spectra of cyanomet H46L *Synechococcus* 7002 rGlbN-A^{2} and -A^{4}.

Figure S2.4: ^{1}H-^{13}C HMQC data collected on a mixture of cyanomet H46L *Synechococcus* 7002 rGlbN-A^{2} and -A^{4}.

Table S2.1: Heme ^{1}H chemical shifts of H46L *Synechococcus* 7002 rGlbN-A^{2}-CN and rGlbN-A^{4}-CN.

Figure S2.5: Sections of the TOCSY and NOESY spectra of a mixture of *Synechocystis* 6803 rGlbN-R-CO and rGlbN-A-CO.

Table S2.2: Heme ^{1}H chemical shifts of *Synechocystis* 6803 rGlbN-R-CO and rGlbN-A-CO.

Figure S2.6: Reaction of ferric rGlbN with H_2O_2 monitored by electronic absorption spectroscopy.

Figure S2.7: Reaction of *Synechocystis* 6803 rGlbN-R with H_2O_2 monitored by SDS-PAGE.

Figure S2.8: Fluorescence emission spectra of wild-type *Synechocystis* 6803 rGlbN in the apoprotein state.

Figure S2.9: Reaction of *Synechococcus* 7002 rGlbN with dithiothreitol monitored by NMR spectroscopy.

Figure S2.10: Reaction of *Synechococcus* 7002 rGlbN with dithiothreitol monitored by absorbance spectroscopy.

Figure S2.11: Isotope effect on the cross-link formation in *Synechocystis* 6803 rGlbN at pL 9.5.

Figure S2.12: Hemochromogen assay on *Synechocystis* 6803 H117C rGlbN-R and -A.

Figure S2.13: Sections of DQF-COSY, NOESY, and ^{1}H-^{13}C HSQC data collected on ferric H117C *Synechocystis* 6803 rGlbN-A.

Table S2.3: Heme ^{1}H chemical shifts of ferric H117C *Synechocystis* 6803 rGlbN-R and rGlbN-A.

Figure S2.14: ^{1}H NMR spectra of H117C *Synechococcus* 7002 rGlbN.

Figure S2.15: Curie plot for the ^{1}H signals of selected heme substituents in *Synechocystis* 6803 and *Synechococcus* 7002 rGlbN-R.

References

Materials and methods

Non-crosslinked protein expression and purification

Protein expression was performed as published previously for *Synechococcus* 7002 rGlbN [1] and *Synechocystis* 6803 rGlbN [2], where "r" indicates the recombinant origin of the material. Wild-type *Synechococcus* 7002 apo rGlbN and *Synechocystis* 6803 apo rGlbN partitioned primarily into inclusion bodies. These were dissolved in 8 M urea, and refolding was achieved along partial purification by passage over a Sephadex G-50 gel filtration column. Holoprotein (rGlbN-R) was obtained by incubation at 4 °C with a solution of hemin chloride (Sigma-Aldrich). Excess heme was removed by centrifugation (48,000 × *g*, 45 min), filtration, and passage through a DEAE Sephacel (Pharmacia, Uppsala Sweden) anion exchange column. Elution of the pure holoprotein was achieved with a 0 to 0.5 M NaCl gradient. Protein solutions were exchanged with water or phosphate buffer by ultrafiltration and lyophilized for long term storage if needed. Between 10 and 30 mg of pure holoprotein were typically obtained per L of medium. The previously published extinction coefficients were used for concentration determination: ε_{410nm} = 100,000 M^{-1} cm^{-1} for wild-type ferric *Synechocystis* 6803 rGlbN-R [2] and ε_{411nm} = 96,000 M^{-1} cm^{-1} for wild-type ferric *Synechococcus* 7002 rGlbN-R [1]. Apoprotein concentrations were evaluated using the extinction coefficients of ε_{280nm} = 4.47 $mM^{-1}cm^{-1}$ (*Synechococcus* 7002 apo rGlbN, [3]) and ε_{280nm} = 7.36 $mM^{-1}cm^{-1}$ (*Synechocystis* 6803 apo rGlbN, [4]).

The new variants studied in this work were H117C *Synechocystis* 6803 rGlbN, H117C *Synechococcus* 7002 rGlbN, and H46L *Synechococcus* 7002 rGlbN. Mutations were introduced using the Stratagene QuikChange® Site-Directed Mutagenesis method (La Jolla, CA). The mutations were confirmed by sequencing of the genes (DNA Analysis Facility of the Johns Hopkins Medical Institute). Holoprotein preparation and purification was as for other variants and without added reducing agent. Only minor quantities of dimeric material were detected by SDS-PAGE in the purified proteins (reducing agent omitted).

Concentrations of apoprotein and holoprotein were estimated using wild-type extinction coefficients.

Cross-linked protein preparation

The proteins with covalently attached heme (rGlbN-A) were prepared from rGlbN-R by addition of a solution of sodium dithionite (DT, Riedel-de Haën 86% or Sigma > 86%) prepared immediately before use. DT concentrations were estimated by optical spectroscopy using the extinction coefficient ε_{314nm} = 8 mM^{-1} cm^{-1} [5]. Lyophilized protein stocks were solubilized in N_2-sparged buffer or water. Proteins were reduced under N_2, with five or higher fold excess DT, and incubation times ranged between 15 min to 48 h (H117C variants) depending on the assumed reactivity of the specific sample. After the incubation period, the reaction mixture was either (1) desalted over a short Sephadex G-25 sizing column, concentrated, and treated with five or greater equivalents of potassium ferricyanide for 15–30 min to return all heme iron to the ferric oxidation state, or (2) directly treated with potassium ferricyanide. Potassium ferri- and ferrocyanide were eliminated by gel filtration with Sephadex G-25. The formation of the cross-link was ascertained by NMR spectroscopy or unsuccessful acid-butanone heme extraction (see below).

Pyridine hemochrome assay

The nature of heme modifications was inspected with the hemochromogen assay [6-7]. DT-treated and untreated H117C *Synechocystis* 6803 rGlbN samples were examined in parallel. At least three reduced pyridine (Fisher Certified ACS >99%)–heme absorbance spectra were collected to yield average absorbance spectra.

Heme extraction

As a test for covalent heme attachment, the acid–butanone method of Teale [8] for *b* heme extraction was applied to holoproteins. rGlbN solutions were acidified in a glass vial by slow drop-wise addition of 0.1 or 1 M HCl to a pH of ~2–2.5 and an equal volume of 2-butanone (99%, Alfa Aesar) was added. Aqueous and organic phases were mixed gently and allowed to separate. In this test, a colorless aqueous phase and a colored organic phase indicated that the heme was not covalently attached to the protein and therefore successfully extracted. Conversely, a colorless organic phase and a colored aqueous phase suggested the formation of rGlbN-A with the caveat that other water-soluble pigmented products could also be present.

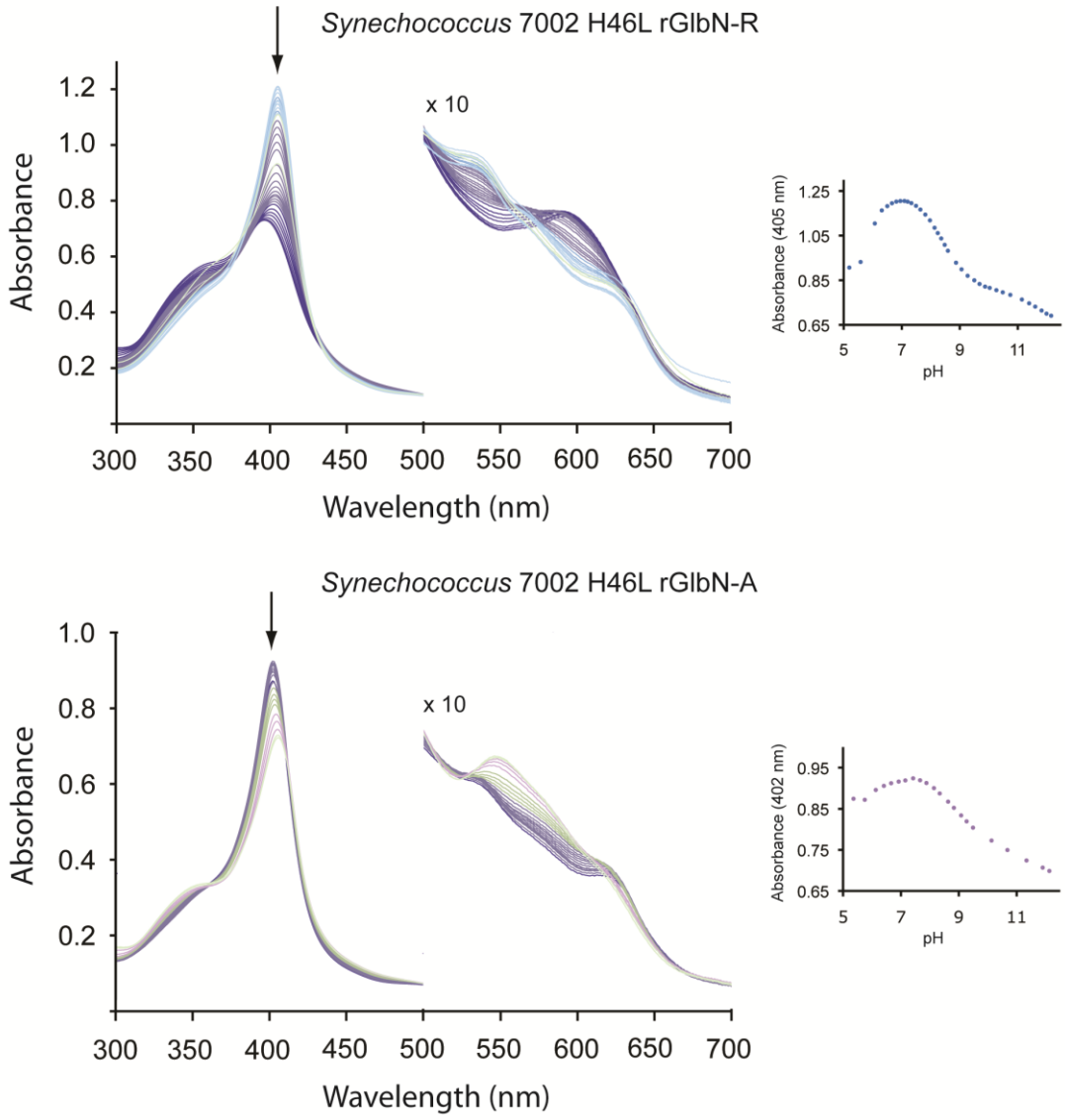


Figure S2.1. Electronic absorption spectra of ferric H46L *Synechococcus* 7002 rGlbN as a function of pH. Top: rGlbN-R; darker colors correspond to higher pH values. Bottom: rGlbN-A; lighter colors correspond to higher pH values. Because the transitions were incompletely defined, the data were not fitted. Approximate pK_a are 8.5 (rGlbN-R) and 9.1 (rGlbN-A).

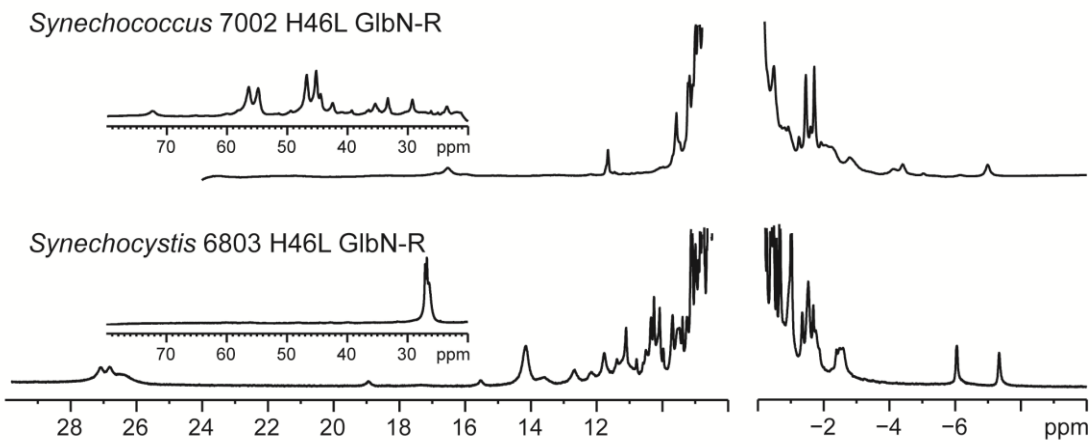


Figure S2.2. ^{1}H NMR spectrum of ferric H46L *Synechococcus* 7002 and *Synechocystis* 6803 rGlbN-R without added exogenous ligand. No signal was detected upfield from the displayed region.

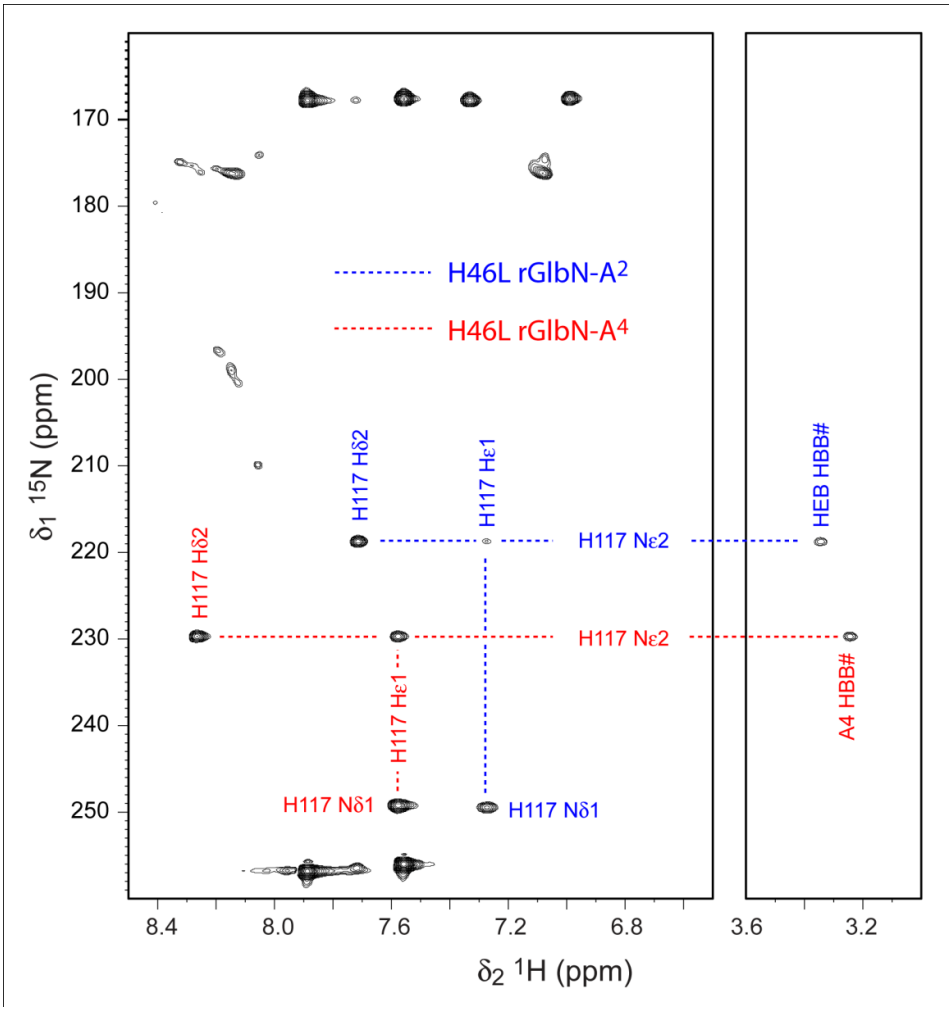


Figure S2.3. ^{1}H-^{15}N HMQC spectra of cyanomet H46L *Synechococcus* 7002 rGlbN-A^{2} (blue) and -A^{4} (red). HEB indicates the heme modified at the 2 position and A4 the heme modified at the 4 position. HBB# denotes the CβH_3. Connectivities are indicated for His117 only. Sample conditions were 20 mM phosphate buffer, pH ~6.4, at 298 K. Proton frequency: 600 MHz.

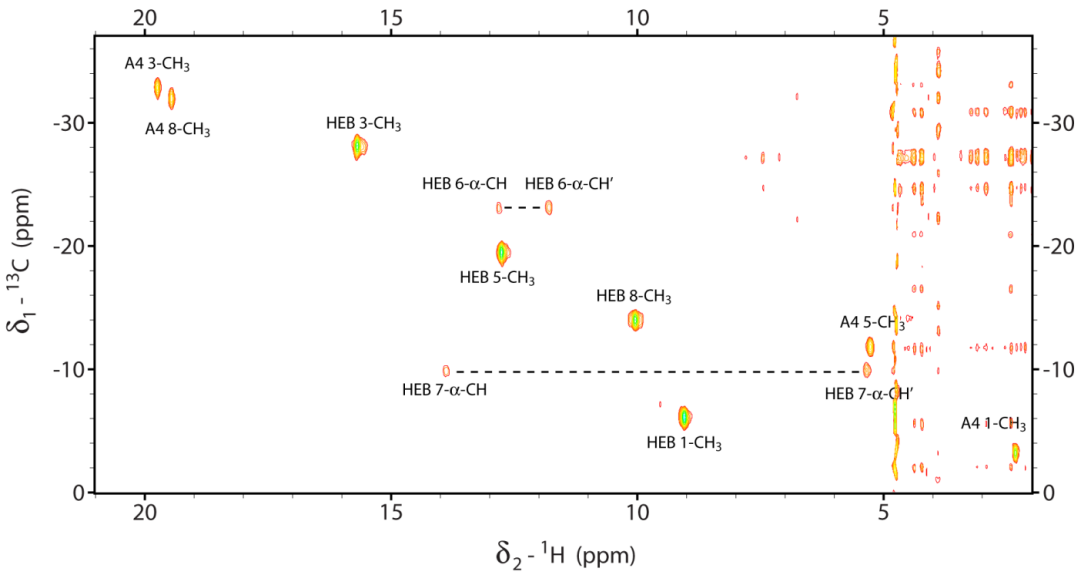


Figure S2.4. ^{1}H-^{13}C HMQC data collected on a mixture of cyanomet H46L *Synechococcus* 7002 rGlbN-A^{2} and -A^{4}. The region containing the heme methyl groups and α-propionates is shown. HEB identifies the A^{2} heme and A4, the heme with His117 linked to the 4-vinyl. Sample conditions were 20 mM phosphate buffer, pH 7.1, 298 K. Proton frequency: 600 MHz.

Table S2.1: Heme ^{1}H chemical shifts of H46L *Synechococcus* 7002 rGlbN-A^{2}-CN and rGlbN-A^{4}-CN.

	rGlbN-A^{2}-CN	rGlbN-A^{4}-CN
1-CH_3	9.05	2.32
2-vinyl α	–	13.60
2-vinyl βc	–	–2.30
2-vinyl βt	–	–4.33
2-α-CH	4.10	–
2-β-CH_3	3.40	–
α meso	–2.65	–1.77
3-CH_3	15.69	19.77
4-vinyl α	8.10	–
4-vinyl βc	1.13	–
4-vinyl βt	2.17	–
4-α-CH	–	
4-β-CH_3	–	3.30
β meso	0.64	
5-CH_3	12.79	5.27
6-α-prop.	12.81	
6-α′-prop.	11.87	
6-β-prop.	1.46	
6-β′-prop.	1.16	
γ meso	–0.55	0.89
8-CH_3	10.01	19.51
7-α-prop.	13.90	23.57
7-α′-prop.	5.34	10.18
7-β-prop.	1.32	1.59
7-β′-prop.	0.10	0.90
δ meso	2.41	6.47

Chemical shifts are listed at pH 7.2, 25 °C in 100% ^2H_2O. These were obtained by using a combination of DQF-COSY, TOCSY, and NOESY data in 90% H_2O/10% ^2H_2O and in 100% ^2H_2O.

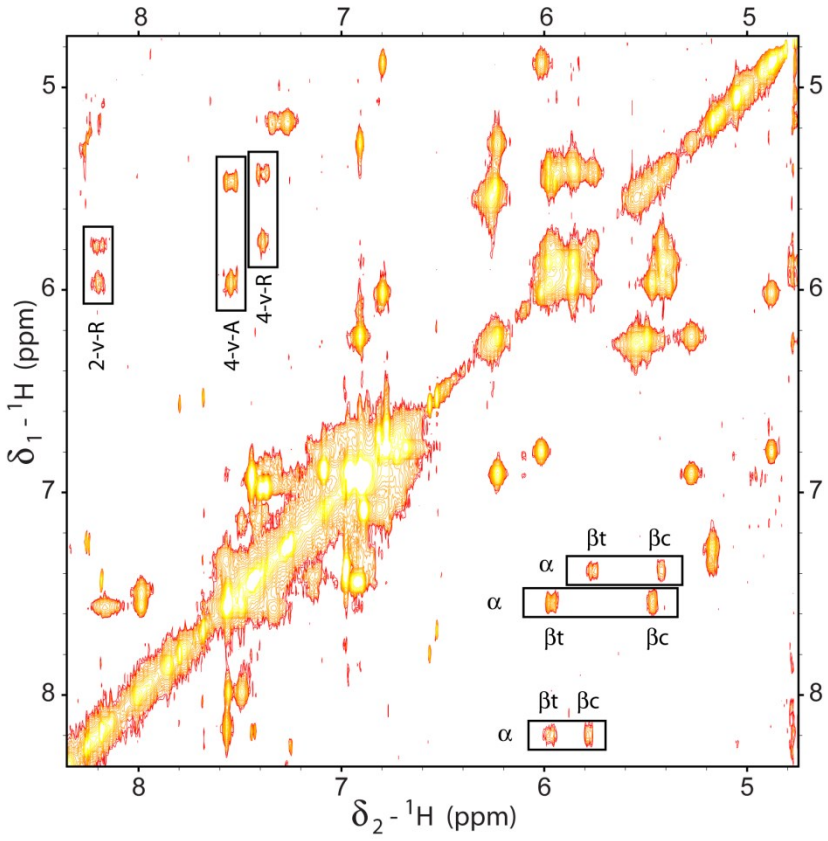


Figure S2.5A. Aromatic region of the TOCSY spectrum of a mixture of *Synechocystis* 6803 rGlbN-R-CO and rGlbN-A-CO. The boxes frame the cross-peaks for the 2-vinyl in rGlbN-R (2-v-R), the 4-vinyl in rGlbN-R (4-v-R), and the 4-vinyl in rGlbN-A (4-v-A). Chemical shifts are listed in Table S2.

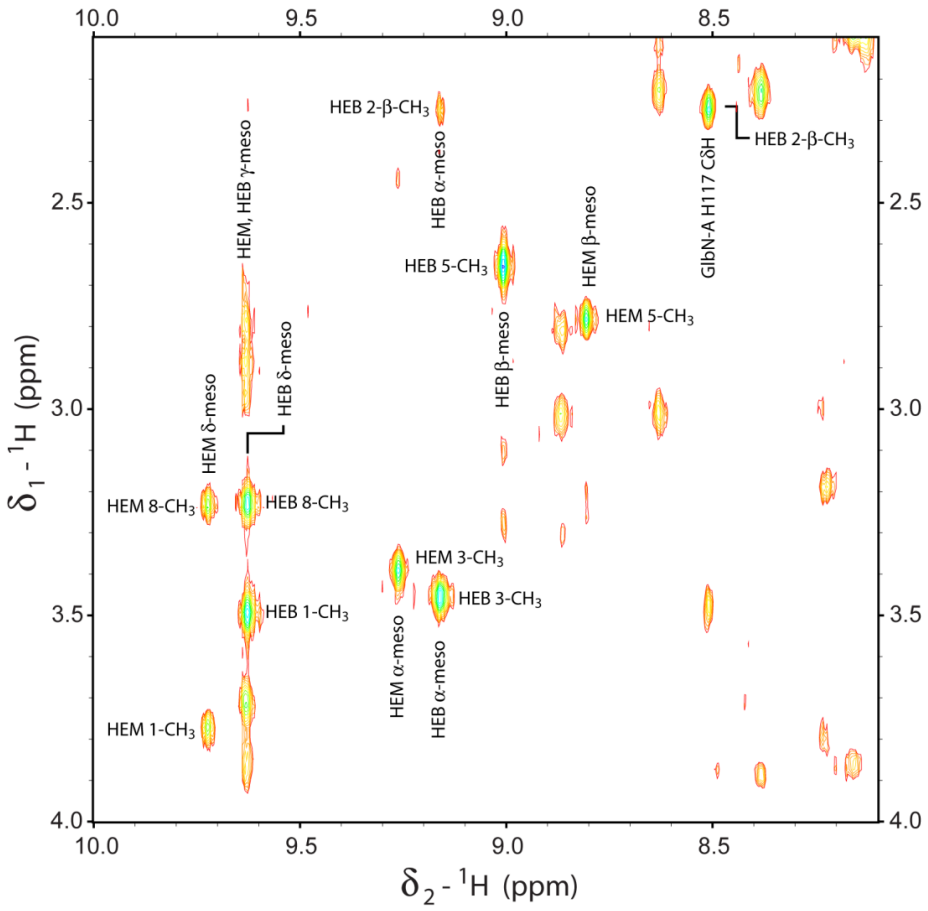


Figure S2.5B. Region of the NOESY data of a mixture of *Synechocystis* 6803 rGlbN-R-CO and rGlbN-A-CO. HEM refers to the unreacted heme; HEB refers to the hybrid b/c heme of rGlbN-A.

Table S2.2: Heme ^{1}H chemical shifts of *Synechocystis* 6803 rGlbN-R-CO and rGlbN-A-CO.

	rGlbN-R-CO	rGlbN-A-CO
1-CH_3	3.77	3.49
2-vinyl α	8.19	–
2-vinyl βc	5.77	–
2-vinyl βt	5.96	–
2-α-CH	–	6.72
2-β-CH_3	–	2.27*
α meso	9.27	9.16
3-CH_3	3.39	3.46
4-vinyl α	7.39	7.55
4-vinyl βc	5.42	5.46
4-vinyl βt	5.75	5.95
β meso	8.81	9.01
5-CH_3	2.78	2.65
γ meso	9.64	9.64
8-CH_3	3.23	3.23
δ meso	9.72	9.63

Chemical shifts are listed at pH 7.2, 25 °C in 100% $^{2}H_2O$. These were obtained by using a combination of DQF-COSY, TOCSY, and NOESY data in $^{2}H_2O$.

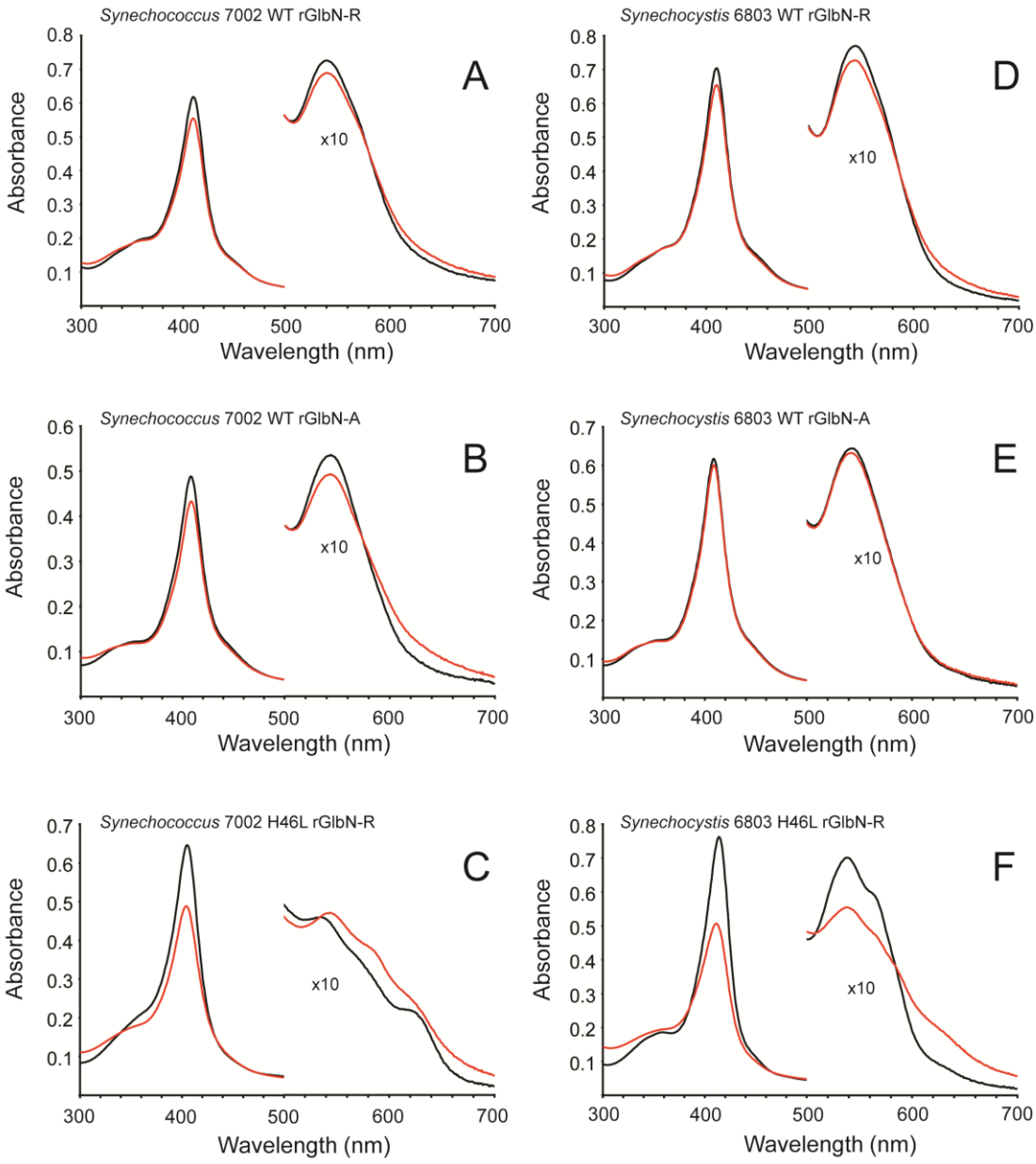


Figure S2.6A–F. Reaction of ferric rGlbN with H_2O_2 monitored by electronic absorption spectroscopy. The spectra are shown before (black) and 2 h after addition of a 5-fold excess of H_2O_2 (red). (A–C) *Synechococcus* 7002, (D–F) *Synechocystis* 6803; (A) and (D) wild-type rGlbN-R; (B) and (E) wild-type rGlbN-A; (C) and (F) H46L rGlbN-R.

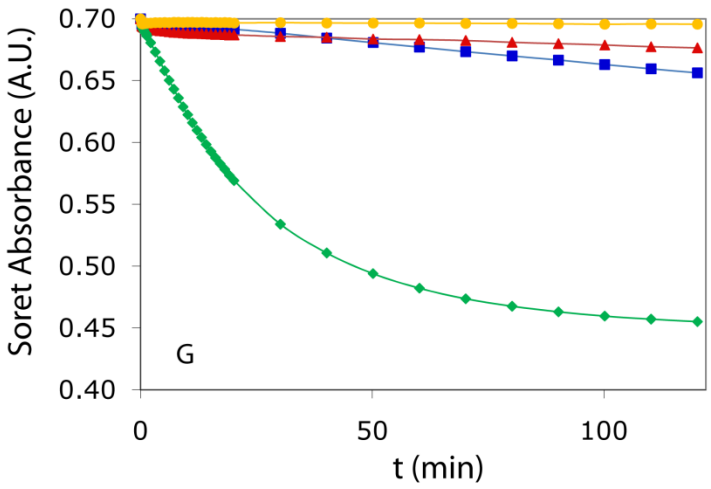


Figure S2.6G. Reaction of ferric rGlbN with H_2O_2 monitored by electronic absorption spectroscopy. (G) Time course obtained by following the Soret intensity of the *Synechocystis* 6803 rGlbN proteins. (□) wild-type rGlbN-R; (△) wild-type rGlbN-A; (◇) H46L rGlbN-R; (○) cyanomet H46L rGlbN-R. Solid lines are drawn to guide the eye. The difference in H46L behavior observed between *Synechococcus* 7002 (Figure 2.3) and *Synechocystis* 6803 may be related to the His–Fe(III)–X ligation of the latter.

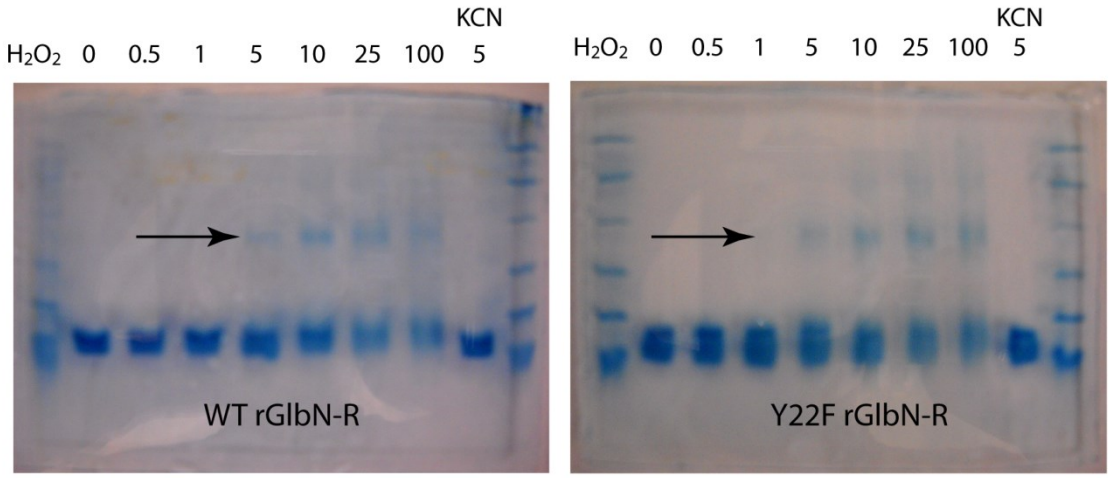


Figure S2.7. Reaction of *Synechocystis* 6803 rGlbN-R with H_2O_2 monitored by SDS-PAGE. The leftmost and rightmost lanes contain the molecular weight markers. Samples were incubated with the specified fold of H_2O_2 for 3 h. KCN above the right lane of each gel indicates overnight incubation with excess KCN prior to H_2O_2 treatment. The arrows indicate the position of the dimer.

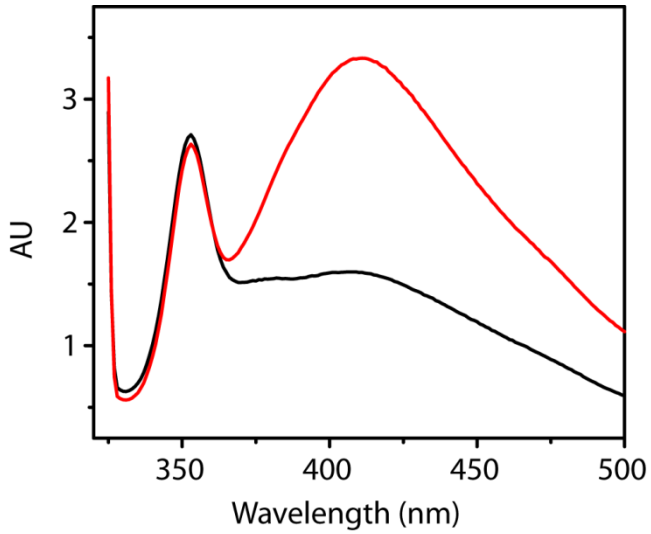


Figure S2.8. Fluorescence emission spectra of wild-type *Synechocystis* 6803 rGlbN in the apoprotein state. The ferric holoprotein protein (rGlbN-R) was either incubated for 3 h with a 5-fold excess of H_2O_2 treatment (red) or not (black). The excitation wavelength was 315 nm. The increased emission at 410 nm is indicative of the formation of o-o'-dityrosine linkage(s).

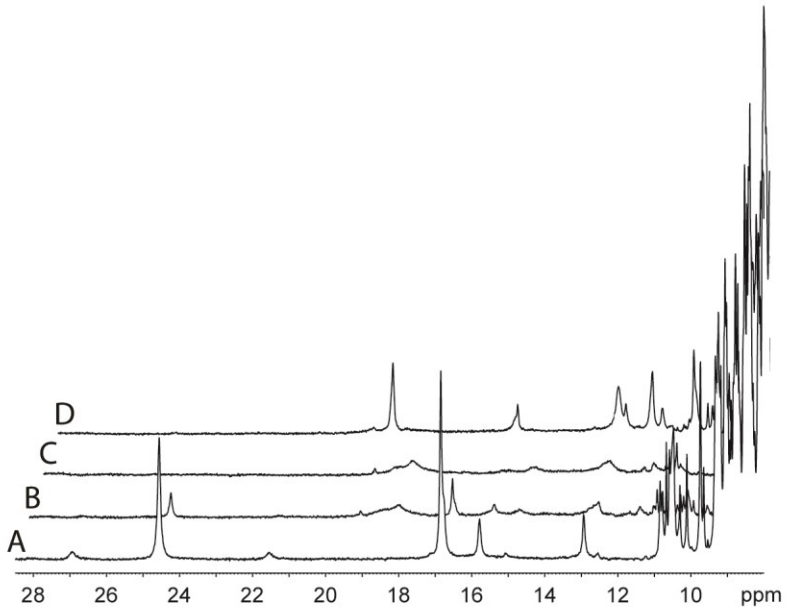


Figure S2.9. 600 MHz ^{1}H 1D NMR spectra of 600 μM wild-type *Synechococcus* 7002 rGlbN (20 mM phosphate buffer pH 7.3, 25 °C, 10 % D_2O). (A) Before DTT addition, the protein is in the ferric rGlbN-R state. (B) 40 min after 20 mM DTT addition, ferric rGlbN-R signals have almost disappeared and the majority of the protein is in the reduced or the oxidized rGlbN-A state. Note the presence of broad signals attributed to protein containing bound DTT species. (C) After extended incubation all peaks that correspond to oxidized rGlbN-R are no longer present (D) After treatment with $K_3[Fe(CN)_6]$ and passage over a Sephadex G-25 desalting column, the protein is entirely in the ferric oxidized rGlbN-A state. The traces were offset horizontally for clarity.

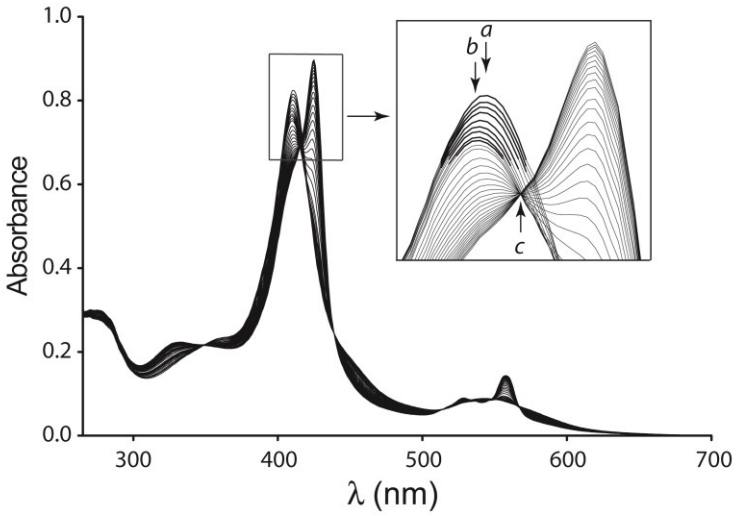


Figure S2.10. Reaction of wild-type ferric *Synechococcus* 7002 rGlbN-R and 11-fold excess DTT in the presence of the GODCAT dioxygen scavenger. The initial absorbance spectrum has a Soret maximum at 411 nm (marked by *a* in the inset). After ~ 2h, the maximum had shifted to 409 nm (marked by *b* and emphasized with thicker lines). This hypsochromic shift and decrease in intensity of the Soret maximum occurred with an isosbestic point at 422 nm and corresponded to the formation of ferric rGlbN-A. Further spectral changes occurred with an isosbestic point at 415 nm (marked by *c*) corresponding to the reduction of rGlbN-A to the ferrous state (maxima at 425, 528, and 558 nm). Other isosbestic points are visible in the figure and supported the ferric rGlbN-R → ferric rGlbN-A → ferrous rGlbN-A sequence described in the text.

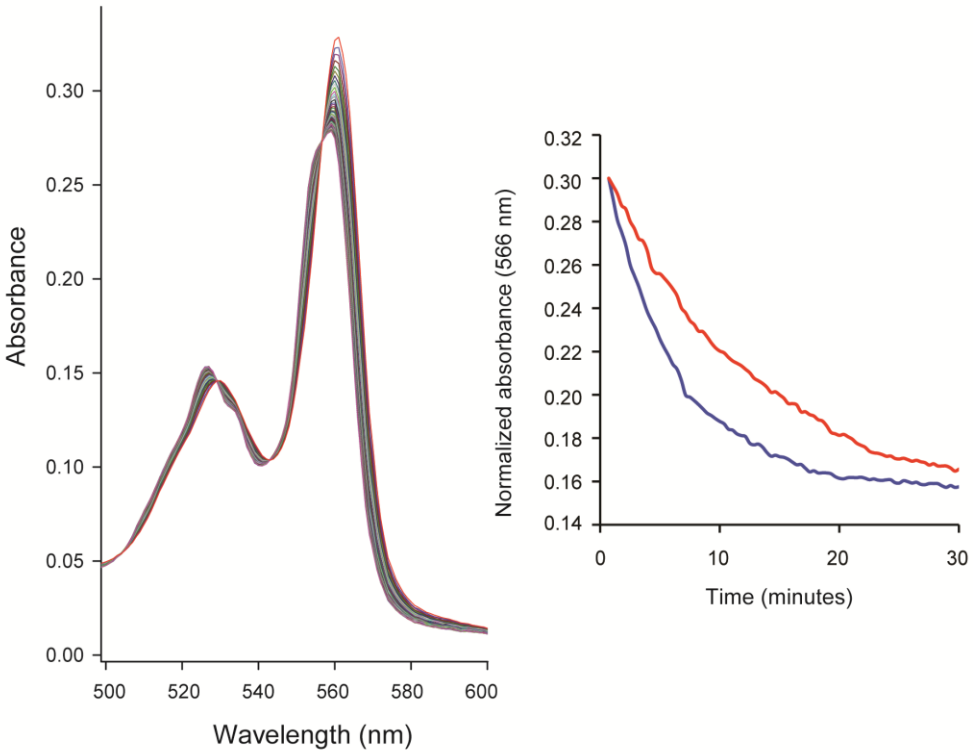


Figure S2.11: Formation of the heme-protein cross-link in *Synechocystis* 6803 rGlbN at pL 9.5 (400 mM glycine buffer) in the presence of the GODCAT oxygen scavenger. Left: evolution of the Q bands as a function of time after DT addition in H_2O solution. The red spectrum was the earliest that could be collected in this manual mixing experiment and corresponded to reduced rGlbN-R. Right: comparison of the reaction kinetics by monitoring absorbance at 566 nm as a function of time. The blue trace corresponds to the reaction carried out in H_2O, pH = 9.5; the red trace corresponds to the reaction carried out in ^2H_2O, pD = 9.5. The traces could be fit with a single exponential decay and indicated a solvent isotope effect of 2 at this pL.

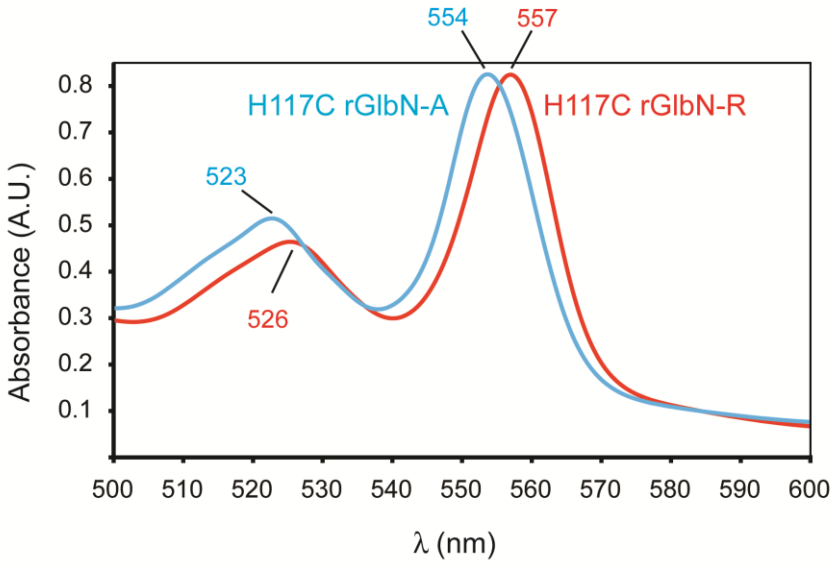


Figure S2.12: Hemochromogen assay performed on *Synechocystis* 6803 H117C rGlbN-R and -A. The spectra were scaled to yield the same α band intensity (0.8 A.U.). The expected maxima for *b*-heme are 557 nm (α band) and 526 nm (β band) [9]. Cytochrome c with two thioether linkages has maxima at 550 nm and 521 nm, whereas cytochrome *c* with a single thioether linkage has maxima at 552–555 and 525 nm [10-11]. Additional values are given in [12].

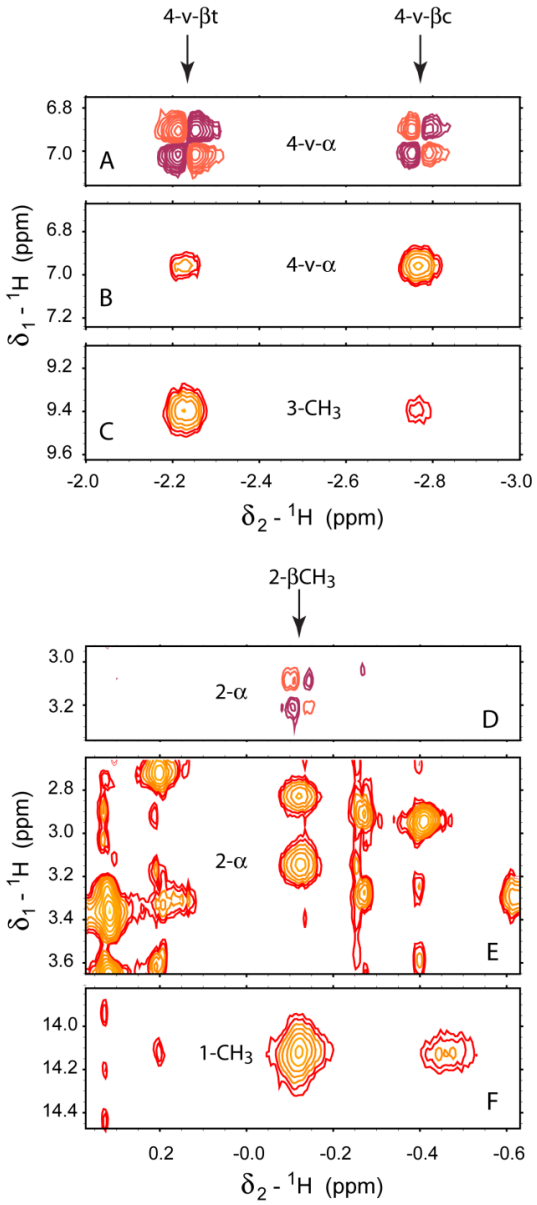


Figure S2.13 A–F. Sections of DQF-COSY (A & D) and NOESY (B, C, E & F) data collected on ferric H117C *Synechocystis* 6803 rGlbN-A illustrating the connectivities for the 2 and 4 heme substituents. Sample conditions were 20 mM phosphate buffer, pH ~8.6, at 298 K. Proton frequency: 600 MHz. Panels A & B and D & E correspond to similar regions of the DQF-COSY and NOESY spectra for the heme substituents at positions 4 and 2, respectively. The panels highlight through-bond and through-space connectivities between the protons on the α and β carbons. Panels E and F display NOE cross peaks between the 4 and 2 substituents and the protons of the nearest heme methyl group.

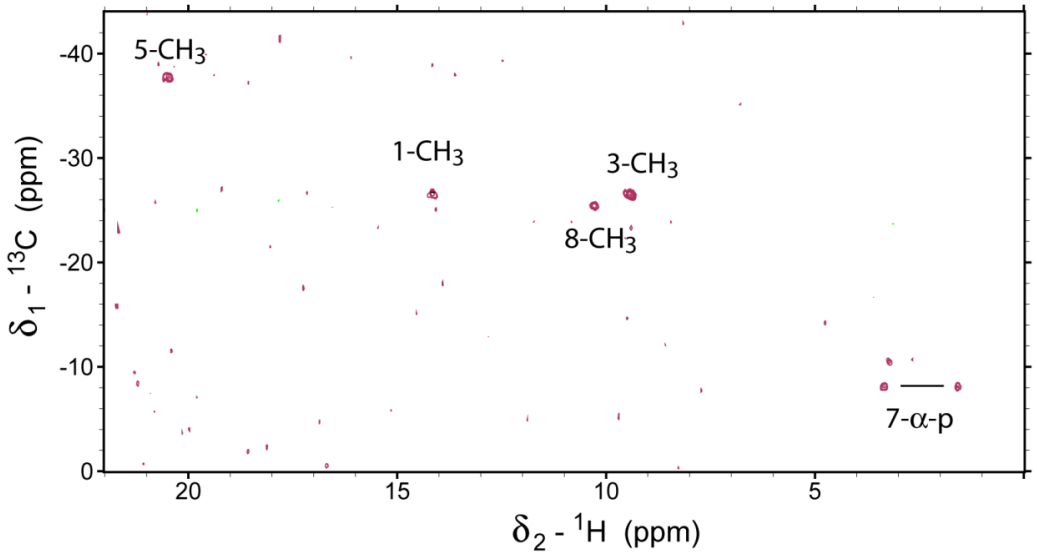


Figure S2.13G. ^{1}H-^{13}C HSQC data collected on H117C *Synechocystis* 6803 rGlbN-A. Sample conditions were 20 mM phosphate buffer, pH ~6.4, at 298 K. Proton frequency: 500 MHz.

Table S2.3: Heme ^{1}H chemical shifts of ferric H117C *Synechocystis* 6803 rGlbN-R and rGlbN-A.

	rGlbN-R	rGlbN-A
1-CH_3	14.97	14.12
2-vinyl α	14.85	–
2-vinyl βc	–5.86	–
2-vinyl βt	–4.98	–
2-α-CH	–	3.15
2-β-CH_3	–	–0.13
α meso		
3-CH_3	9.39	9.40
4-vinyl α	6.64	6.96
4-vinyl βc	–1.59	–2.77
4-vinyl βt	–1.10	–2.23
β meso		
5-CH_3	20.58	20.54
6-α-prop.	9.57	9.26
6-α′-prop.	8.69	8.51
6-β-prop.	0.63	0.24
6-β′-prop.	1.26	0.93
γ meso		
7-α-prop.	3.76	3.30
7-α′-prop.	2.38	1.51
7-β-prop.	–0.85	–1.05
7-β′-prop.	–0.56	–0.62
8-CH_3	9.36	10.21
δ meso		–0.46

Chemical shifts are listed at pH 7.2 (rGlbN-R) and 8.7 (borate buffer, rGlbN-A), 25 °C in 100% $^{2}H_2O$. Obtained by using a combination of DQF-COSY, TOCSY, and NOESY data in $^{2}H_2O$.

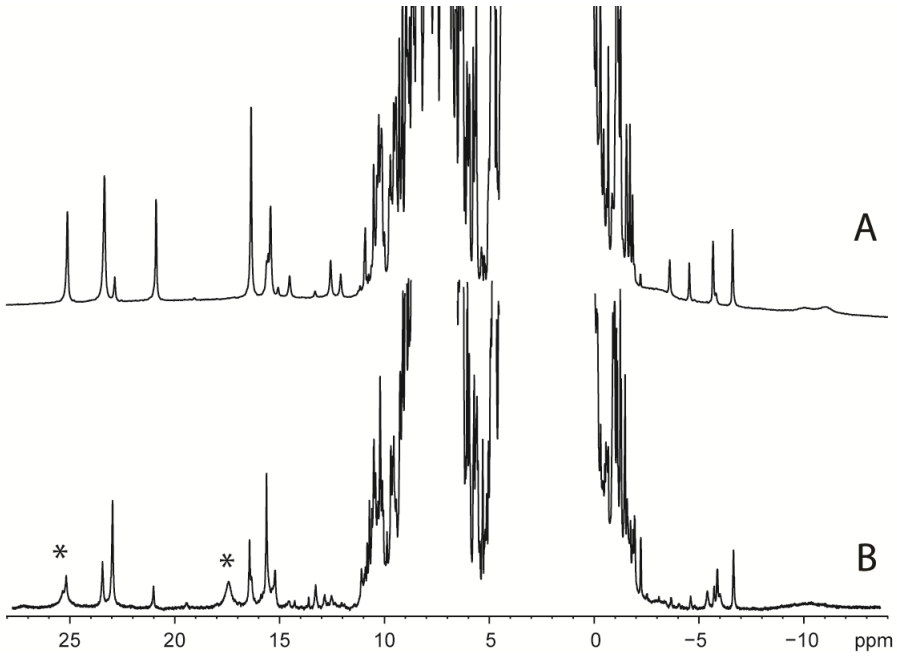


Figure S2.14: 600 MHz ^{1}H NMR spectra of H117C *Synechococcus* 7002 rGlbN in the ferric *bis*-histidine state. (A) rGlbN-R, which resembles the wild-type spectrum; this sample contains about 40% of the minor heme orientational isomer. (B) Mixture of rGlbN-R and rGlbN-A, obtained after treatment of rGlbN-R with DT and re-oxidation of the iron. New lines are marked by * and indicate that a protein modification has occurred. Data were collected in 90% H_2O/10% D_2O, 20 mM phosphate, pH 8.3 and 25 °C.

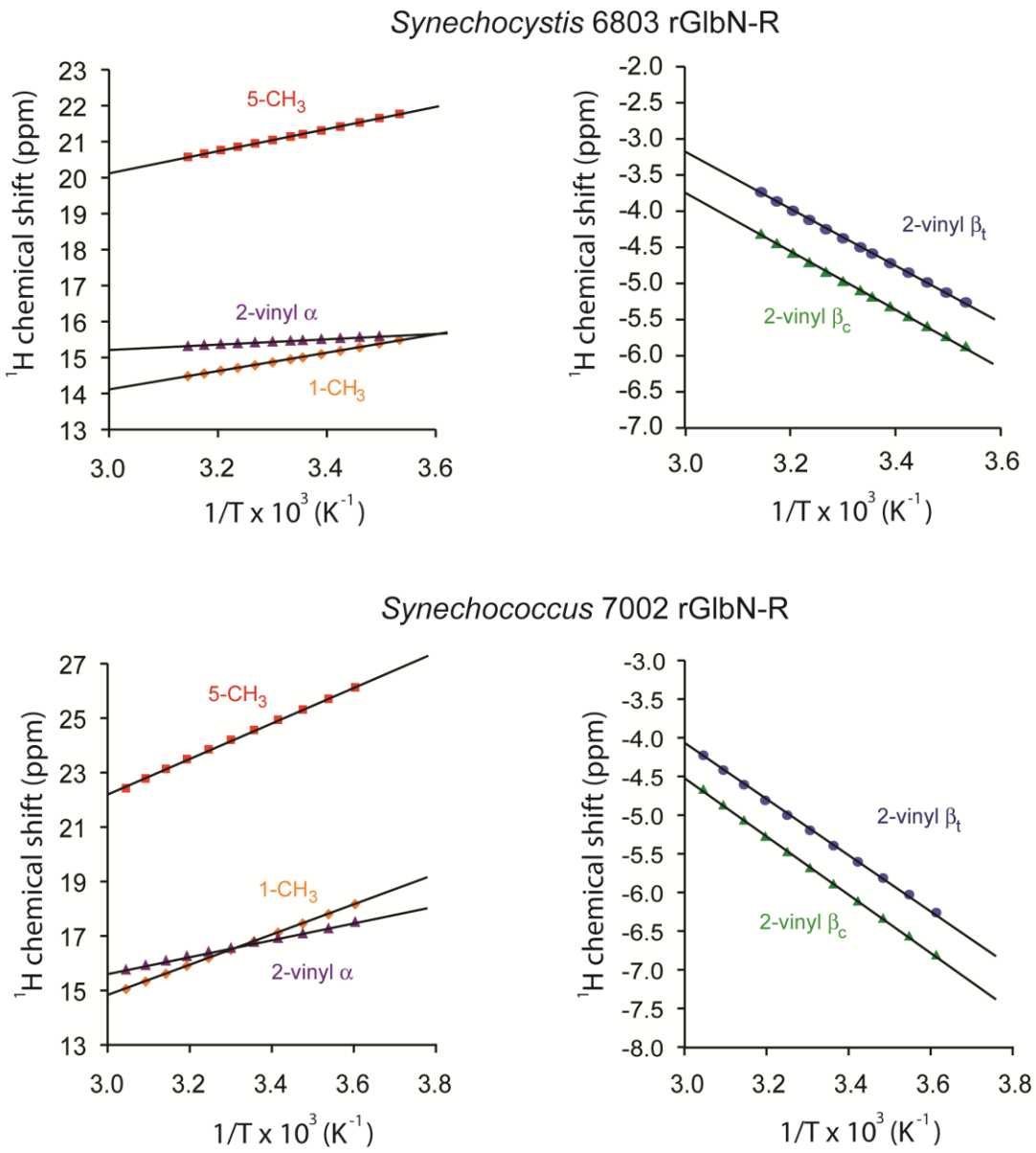


Figure S2.15: Curie plot for the ^{1}H signals of selected heme substituents in *Synechocystis* 6803 (600 MHz) and *Synechococcus* 7002 (500 MHz) rGlbN-R at pH 7.2 (10% ^2H_2O).

Supporting Information References

1. Scott NL, Falzone CJ, Vuletich DA, Zhao J, Bryant DA, Lecomte JTJ (2002) Biochemistry 41:6902-6910

2. Scott NL, Lecomte JTJ (2000) Protein Sci. 9:587-597

3. Vuletich DA, Falzone CJ, Lecomte JTJ (2006) Biochemistry 45:14075-14084

4. Lecomte JTJ, Scott NL, Vu BC, Falzone CJ (2001) Biochemistry 40:6541-6552

5. Di Iorio EE (1981) Methods Enzymol 76:57-72

6. de Duve C (1948) Acta Chem. Scan. 2:264-289

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

8. Teale FWJ (1959) Biochim. Biophys. Acta 35:543

9. Paul KG, Theorell H, Akeson A (1953) Acta Chem. Scan. 7:1284-1287

10. Barker PD, Ferrer JC, Mylrajan M, Loehr TM, Feng R, Konishi Y, Funk WD, MacGillivray RT, Mauk AG (1993) Proc Natl Acad Sci U S A 90:6542-6546

11. Allen JW, Ginger ML, Ferguson SJ (2004) Biochem. J. 383:537-542

12. Antonini E, Brunori M, Caputo A, Chiancone E, Rossi-Fanelli A, Wyman J (1964) Biochim. Biophys. Acta 79:284-292

Chapter 3.

Electron self-exchange and self-amplified posttranslational modification in the hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002

Matthew R. Preimesberger, Matthew P. Pond, Ananya Majumdar, and Juliette T.J. Lecomte

Reproduced with permission from J Biol Inorg Chem (2012) 17(4):599-609 doi: 10.1007/s00775-012-0880-5.

Author contributions:

MRP: designed research, prepared samples, conducted UV-visible experiments and analysis, performed NMR experiments and analyzed data (ZZ exchange, band-selective HSQC-monitored kinetics, triple resonance backbone assignments), wrote data fitting programs, conducted ET/ESE error analysis, and wrote the paper

MPP: performed NMR T_1 relaxation experiments

AM: wrote ZZ-exchange pulse sequences, assisted with NMR data acquisition, wrote the paper

JTJL: designed research, analyzed data, performed magnetic susceptibility structure calculations, wrote the paper

Abbreviations: ESE, electron self-exchange; ET, electron transfer; GlbN, hemoglobin product of the *glbN* gene; GlbN-A, GlbN with covalently attached heme; GlbN-R, GlbN with non-covalently attached heme; Hb, hemoglobin; PTM, post-translational modification; WT, wild-type.

Abstract

Many heme proteins undergo covalent attachment of the heme group to a protein side chain. Such post-translational modifications adjust the thermodynamic and chemical properties of the holoprotein. Their importance in biological processes makes them attractive targets for mechanistic studies. We have proposed a reductively-driven mechanism for the covalent heme attachment in the monomeric hemoglobins produced by the cyanobacteria *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 (GlbNs) [Nothnagel HJ, Preimesberger MR, Pond MP, Winer BY, Adney EM, Lecomte JTJ (2011) J. Biol. Inorg. Chem. 16:539-552]. These GlbNs coordinate the heme iron with two axial histidines, a feature that distinguishes them from most hemoglobins and conditions their redox properties. Here, we uncovered evidence for an electron exchange chain reaction leading to complete heme modification upon substoichiometric reduction of GlbN prepared in the ferric state. GlbN electron self-exchange rate constants measured by NMR spectroscopy were on the order of 10^2–10^3 M^{-1} s^{-1} and consistent with the proposed autocatalytic process. NMR data on ferrous and ferric *Synechococcus* GlbN in solution indicated little dependence of the structure on the redox state of the iron or cross-link status of the heme group. This allowed for the determination of lower bounds to the cross-exchange rate constants according to Marcus theory. The observations illustrated the ability of bishistidine hemoglobins to undergo facile interprotein electron transfer and the chemical relevance of such transfer for covalent heme attachment.

Introduction

Hemoglobins (Hbs) are well known for their ability to bind and release the dioxygen required for vertebrate metabolism. Genomic initiatives, however, have brought to light a rich collection of Hb sequences from all three kingdoms of life. They show that the Hb superfamily has existed for billions of years and that its members are capable of performing a range of physiological tasks in diverse environments [1]. Modern non-vertebrate Hbs contain clues for understanding the modulation of heme prosthetic group chemistry and the molecular evolution of Hb function [2].

GlbN is the monomeric "truncated" Hb of the oxygenic cyanobacteria *Synechococcus* sp. PCC 7002 [3, 4] and *Synechocystis* sp. PCC 6803 [5, 6]. The two proteins are 59% identical in sequence and belong to a growing group of Hbs that ligate the heme iron with two histidines (proximal and distal) in the reduced (ferrous) and oxidized (ferric) states. *Synechococcus* and *Synechocystis* GlbNs are nevertheless capable of exogenous ligand binding by displacement of the distal histidine, but they differ from all other Hbs in that they undergo covalent heme attachment [7, 8]. The post-translational modification (PTM), depicted in Figure 3.1, occurs rapidly in the reduced state in the absence of exogenous ligand. The function of *Synechocystis* GlbN is unknown; *Synechococcus* GlbN protects the cyanobacterial cell against damage from nitrate metabolism [4], likely through enzymatic processes requiring redox cycling of the heme iron. Both forms of *Synechococcus* GlbN, with and without the PTM, have been extracted from the cell [4].

Various heme–protein cross-links have been documented in cytochromes [9, 10] and peroxidases [11]. In these proteins, covalent linkages are proposed to alter holoprotein stability, reactivity, and redox properties. Covalent linkages are expected to contribute to the functional diversification of the Hb superfamily as well; therefore the mechanism and consequences of the heme PTM in GlbN merits further investigation.

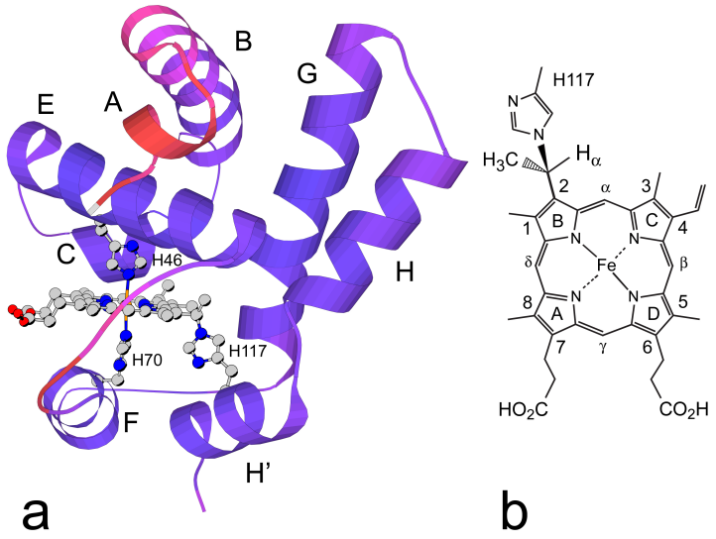


Figure 3.1. (a) Solution structure of *Synechococcus* GlbN with cross-linked heme (GlbN-A) in the ferric state [4]. A low energy member of the 2KSC family is shown with redder colors indicating larger ensemble rmsd from the average. Helices, axial histidines (His46 or E10, distal, and His70 or F8, proximal) and reactive histidine (His117) are labeled. (b) The structure of the His117–heme covalent linkage [7].

The PTM of Fe(III)GlbN via Fe(II)GlbN does not require an enzyme or dioxygen [7, 12]. Observations thus far are consistent with an electrophilic addition involving the protonation of the 2-vinyl Cβ [12], as shown in Figure 3.2.

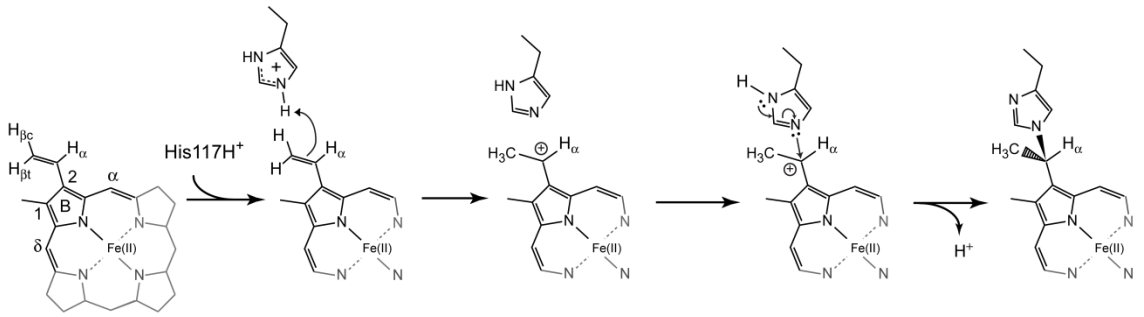


Figure 3.2. Proposed GlbN reductive cross-linking mechanism. Only a portion of the Fe(II) protoporphyrin IX molecule is shown.

An intriguing aspect of this particular PTM is that, despite the apparent simplicity of the alkylation reaction, the histidine–heme bond has been found only in the GlbN of the two cyanobacteria studied here. It is not clear whether the rarity hints at subtle molecular determinants or results mainly from the absence of histidines at appropriate locations. In pursuing the description of the reductive mechanism, we investigated the intrinsic ability of *Synechococcus* and *Synechocystis* GlbNs to undergo electron transfer and we characterized the structure of the reduced state of post-translationally modified *Synechococcus* GlbN in solution, a form for which no prior data were available. Our results demonstrate that inter-GlbN electron exchange is readily detected by NMR spectroscopy, not only within a redox pair (self-exchange), but also between cross-linked and non-crosslinked GlbN (cross-exchange). This latter process illustrates how a heme modification is amplified and propagated through a GlbN solution.

Materials and methods

Protein expression and purification

The procedures used to produce recombinant GlbNs (WT and H117A *Synechococcus* GlbN, *Synechocystis* WT GlbN) have been published previously [3, 6, 12, 13] and were followed with minor adaptations. A description is provided in the Supporting Information, along with the reported extinction coefficients used for protein concentration determination. When required for clarity, we refer to the non-crosslinked protein as GlbN-R and to the cross-linked protein as GlbN-A [12].

Generation of Fe(II)GlbN from Fe(III)GlbN under microoxic conditions

The glucose oxidase/D-glucose/catalase (GODCAT) enzymatic O_2 elimination system [14] was used to generate microoxic conditions in solution. The concentration of GODCAT components depended on the nature of the experiment. Final reagent concentrations were: 0.04–0.7 mg/mL glucose oxidase, 0.04–0.15 mg/mL catalase, and 0.1–0.2% D-glucose (mass/volume). High buffer concentrations (200–250 mM) were used to minimize pH changes caused by GODCAT enzymatic turnover and the oxidation of dithionite (DT). Under these conditions, negligible changes in pH occurred over time.

DT stock solutions were prepared immediately before use. The concentration was estimated by mass or by absorbance measurement with an extinction coefficient of 8 mM^{-1} cm^{-1} at 314 nm [15]. DT oxidation eliminates the 314-nm absorbance, which provided a convenient means of assessing the depletion of the reductant.

Fe(II)GlbN UV/vis samples

Fe(III)GlbN and GODCAT solutions were allowed to incubate together for approximately 20 min prior to the addition of 200-fold DT. During this time, the concentration of the samples was measured and their stability assessed by the constancy of their spectrum in the 250–700 nm region. Upon addition of reductant, the cuvette was stopper-sealed and the content mixed by inversion prior to data acquisition. For kinetic conversion experiments, the dead time was about 15 s. Absorbance data were collected every 1 nm over the range 600–500 nm, using a 0.2-s averaging time.

Fe(II)GlbN NMR samples

Fe(III)GlbN and GODCAT solutions were allowed to incubate for approximately 1 h prior to treatment with DT. During this time, the protein concentration was determined by absorbance measurement, and ^{1}H 1D and ^{1}H-^{15}N 2D HSQC NMR data were collected to confirm sample homogeneity. Fe(III)GlbNs were stable in the presence of GODCAT. Fe(III)GlbN (+ GODCAT) solutions were placed in a glove box continuously purged with 99.9% Ar. GlbN reduction (30–100%) was achieved by adding a small volume of DT stock solution also prepared under Ar. The protein solution was mixed by pipette and transferred to a Shigemi NMR tube, which was sealed with Parafilm. GlbN redox mixtures prepared in this manner underwent some initial oxidation but eventually reached a stable Fe(II)/Fe(III) ratio, which was maintained from days to weeks. This enabled the acquisition of multidimensional NMR spectra. For kinetic conversion experiments, the dead time between the addition of DT and the initiation of NMR data acquisition was ~ 12 min.

Kinetic experiments (optical)

Electronic absorption spectroscopy, manual mixing. Electronic absorption spectra were collected at room temperature on a Varian Cary 50-Bio spectrophotometer. Protein concentrations were 5–11 μM. Kinetic experiments monitoring the conversion of *Synechococcus* Fe(III)GlbN-R to Fe(II)GlbN-A upon treatment with 200-fold excess DT were conducted in 250 mM borate, pH 9.2 in the presence of GODCAT. Data collection was performed in triplicate.

Electronic absorption spectroscopy, rapid mixing stopped-flow. Electronic absorption spectra were collected at 24 °C on a Hi-Tech Scientific stopped-flow spectrophotometer equipped with a photodiode array detector. Syringes contained either *Synechocystis* GlbN-R (+ GODCAT) in buffer or 200-fold excess DT. Solutions were mixed by stopped-flow (200 μL reaction volume). Final reaction concentrations were 5 μM GlbN, 1 mM DT, and 200 mM buffer. Reaction mixing times (dead times) varied between 100 ms at low pH and 750 ms at pH 8.9, but in all cases were such that reliable rate constants could be obtained from the data. During the mixing time, complete GlbN-R reduction occurred. Reactions were carried out at pH 5.3 (acetate), pH 5.7 (acetate/phosphate), pH 6.0 (phosphate), p^1H or p^2H 6.7 (phosphate), pH 7.5 (phosphate), pH 8.3 (tris), pH 8.9 (tris), and pH 9.5 (glycine). A total of 200 spectra were recorded per run in the wavelength range 320–704 nm.

Determination of post-translational modification rate (optical data). Singular value decomposition (SVD) and global nonlinear least-squares fitting were performed on optical spectra using the program KinTek Explorer (KinTek corporation) [16, 17] or a combination of SciLab 5.2 for SVD and Savuka [18]

for global fitting. Two or three abstract kinetic vectors were determined to be significant on the basis of their large singular values and high autocorrelation coefficients.

The pH dependence of the *Synechocystis* GlbN cross-linking reaction was modeled assuming the existence of two Fe(II)GlbN states in rapid equilibrium: a protonated, "reactive" form that can serve as a H^+ donor in the Markovnikov addition reaction, and a deprotonated, "unreactive" form that cannot act as a general acid. Using this model, the pH dependence data were adequately fit using a single apparent pK_a and Hill coefficient in Equation (3.1) with the program Kaleidagraph 3.6 (Figure S3.1).

$$k_{obs} = k_{max}\left(\frac{10^{n(pKa-pH)}}{1+10^{n(pKa-pH)}}\right) \quad (3.1)$$

NMR spectroscopy

Samples and acquisition. *Synechococcus* and *Synechocystis* GlbN samples were prepared in 225–250 mM phosphate (pH 7.1) or borate (pH 9.2) buffer in 95:5 $^1H_2O/^2H_2O$. Protein concentration ranged from 0.6 mM to 4.9 mM. NMR spectra were collected at 298 K with Bruker DRX-600, Bruker Avance-600, Bruker Avance II-600, or Varian Inova 800 spectrometers each equipped with a cryoprobe. ^1H chemical shifts were referenced to DSS through the ^1H_2O line (4.76 ppm). Water-presaturation 1D data and ^1H-^{15}N 2D HSQC data were collected as described previously [19]. To confirm backbone HN assignments determined using ZZ exchange methods, triple resonance HNCACB and CBCA(CO)NH experiments were collected on *Synechococcus* WT Fe(II)GlbN-A prepared by treatment with five-fold excess DT as described above.

Processing. NMR data were processed using either Bruker TopSpin 2.1 (Bruker BioSpin, Rheinstetten, Germany) or NMRPipe 3.0 [20]. Spectra were analyzed using the programs Sparky [21] and TopSpin 2.1. Further information is provided in our prior NMR work on the same proteins [22, 23].

Band-selective 1D and 2D ^{15}N-^{1}H HSQC spectroscopy. We used an ^{15}N-edited 1D experiment (band-selective 1D ^{15}N-^{1}H HSQC) to collect high S/N data rapidly. The sequence was derived from a WATERGATE HSQC with water flip-back. ^{15}N band selectivity was achieved by addition of a 20-ms RE-BURP (refocusing band-selective pulse with uniform response and phase) [24] centered at 130.5 ppm (*Synechococcus* GlbN) or 131.5 ppm (*Synechocystis* GlbN) at the end of the minimal evolution period. Excited signals were retained by phase cycling of the selective ^{15}N pulse. The excitation bandwidth (~900 Hz) was determined using a 2D version of the experiment as shown in Figure 3.3 (a complete explanation is provided in the Supporting Information). Each 1D data point was collected in 1.5 min (*Synechocystis* GlbN) or 2.5 min (*Synechococcus* GlbN).

2D ^{15}N-^{1}H ZZ exchange NMR spectroscopy. Electron self-exchange (ESE) rates were measured in *Synechococcus* WT GlbN-A and H117A GlbN and *Synechocystis* GlbN-A at 600 MHz with an ^{15}N ZZ exchange experiment [25]. To eliminate the effects of heteronuclear cross-correlated relaxation, an even number of 1.3 ms G^{3} ^{1}H selective inversion pulses [26] were applied every 10 ms during the mixing period (τ_{mix}), with the ^{1}H carrier positioned 4 ppm downfield of the H_2O resonance (4.76 ppm). The duration and positioning of these pulses

ensured minimal perturbation of the water resonance, which was thereby maintained close to thermal equilibrium for all values of τ_{mix}.

The mixing time was varied in the following order (ms): 497, 994, 45, 633, 249, 339, 1288, 791, 136, 497, 1582 (*Synechococcus* WT GlbN-A); and 497, 994, 294, 746, 136, 1582, 68, 1288, 610, 384, 203, 497 (*Synechococcus* H117A GlbN). For *Synechocystis* WT GlbN-A, the mixing time was varied as follows (ms): 497, 294, 136, 1288, 633, 1582, 1084, 384, 836, 68, and 497. Duplicate τ_{mix} (497 ms) data points were used to estimate peak volume uncertainty (~10%). The *Synechococcus* WT GlbN-A sample contained approximately 40–45% Fe(II) and 60–55% Fe(III) protein (2.0 mM total GlbN). H117A GlbN ESE kinetics were measured on a ~1:3 Fe(II)/Fe(III) protein sample (1.7 mM total GlbN). The Fe(II)/Fe(III) redox mixture of *Synechocystis* WT GlbN-A was prepared at a total protein concentration of approximately 4.9 mM with about 15% and 85% Fe(II) and Fe(III) protein.

Resolved cross-peaks were integrated to generate evolution curves. These were globally fitted and provided forward and reverse apparent first order rate constants for each residue. The sum of the two constants was divided by the total protein concentration to obtain the second order rate constant for the self-exchange reaction. Additional information is provided in the Supporting Information.

3D ^{1}H-^{15}N-^{1}H SE ZZ-exchange NMR spectroscopy

A 3D sensitivity-enhanced version of the ^{15}N ZZ-exchange experiment [25] was used to transfer the previously determined backbone ^{15}N-^{1}H assignments of

Synechococcus WT Fe(III)GlbN-A [23] to Fe(II)GlbN-A through longitudinal magnetization exchange. The additional ^{1}H frequency-labeling period considerably increased the number of resolved resonances since many NHs exhibit a degenerate ^{15}N chemical shift in ferrous and ferric GlbN-A. The ZZ mixing time was set to 484 ms to maximize cross-peak intensity. Data were collected at 800 MHz to enhance chemical shift dispersion. The pulse sequence is shown in Figure S3.2.

Heteronuclear relaxation measurements. ^{15}N T_1 measurements were carried out on *Synechococcus* WT GlbN-A in the Fe(II) and Fe(III) oxidation states using a standard pulse sequence [27] with a WATERGATE modification [28]. The experiments were performed at 600 MHz on the Bruker spectrometers mentioned above. The sample conditions were ~1.5 mM GlbN-A at pH 7.2 in 20 mM (Fe(III)GlbN-A) or 100 mM Fe(II)GlbN-A phosphate buffer containing 10% $^{2}H_2O$. Fe(II)GlbN-A was generated by the addition of a 5-fold molar excess DT in the presence of GODCAT. The T_1 delays were set randomly to 14, 105*, 203, 301, 448, 595, 742*, 1183, 2408 ms for Fe(III)GlbN-A and 14, 630*, 1400, 1975* ms for Fe(II) GlbN-A (where the asterisks indicate duplicate time points for error analysis); the delay between transients was 3 s. The data were analyzed using the jackknife method of error analysis with the program Curvefit (A. G. Palmer, Columbia University) through the use of the script Sparky2rate (J. P. Loria, Yale University).

Structural calculations

Backbone ^{1}H, ^{15}N, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts (pH 7.1) for *Synechococcus* WT Fe(II)GlbN-A at pH 7.1–7.2 were used as input for the program TALOS+ [29]. In

addition, the Fe(II)GlbN-A chemical shifts were subtracted from those of Fe(III)GlbN-A [23] to obtain $\Delta\delta$ for comparison with the pseudocontact contribution, δ_{pc}, at 114 HN and 18 CαH sites. The pseudocontact shift is expressed as

$$\delta_{pc} = \frac{1}{12\pi r_i^3}\left[\Delta\chi_{ax}\left(3\cos^2\theta_i - 1\right) + \frac{3}{2}\Delta\chi_{rh}\sin^2\theta_i\cos 2\varphi_i\right] \qquad (3.2)$$

in the frame of reference that diagonalizes χ, the magnetic susceptibility tensor. In Equation 3.2, r_i represents the distance between the ^{1}H of interest (i) and the Fe center; θ_i and φ_i are the angular coordinates; and $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components of χ [30]. The $\Delta\delta$ values were tabulated for use as restraints in refinement with XPLOR-NIH [31] (version 2.23) and the PARArestraint module [32]. All other previously deposited 2KSC restraints (NOEs, dihedral angles, and H-bonds) [4] were also applied. Model 2 of the Fe(III)GlbN-A ensemble 2KSC was chosen and subjected to a Gaussian/DELPHIC torsional potential [33] as previously reported [4]. The program Numbat [34] was used to obtain starting $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ values from the experimental shifts and 2KSC model 2.

Results and Discussion

Kinetics of the cross-linking reaction as a function of pH

The reactive histidine of *Synechocystis* GlbN-R, His117, has a pK_a ~ 6.9 in Fe(III)GlbN-R [8]. Preliminary pH dependence data [12] led us to propose that in addition to serving as the nucleophile, His117 also served as the proton

donor to form the 2-vinyl α carbocation (Figure 3.2). The change in the optical spectrum on going from Fe(II)GlbN-R to Fe(II)GlbN-A is small but sufficient to follow the cross-linking reaction after complete reduction with DT [12]. Above pH 9, the reaction proceeded with an observed rate constant, k_{xl}, below 10^{-2} s^{-1} (5×10^{-3} s^{-1} at pH 9.2) and could be monitored by manual mixing experiments. Between pH 8.9 and pH 5.7, stopped-flow rapid mixing was required as the observed rate constant increased to 0.9 s^{-1}. Below pH 5.7, the ferrous heme dissociated from the apoprotein, as demonstrated by scavenging with horse apomyoglobin (data not shown). This additional equilibrium complicated the conversion mechanism and rendered the fits unreliable. The observed reaction rate constant, plotted as a function of pH, is shown in Figure S3.1 of the Supporting Information. The curve displayed an inflection at pH ~6.8 arguing against water acting as the proton donor and supporting the hypothesis that His117 acted as the general acid in the rate-determining protonation of Fe(II) heme 2-vinyl Cβ as shown in Figure 3.2.

The pK_a of His117 in *Synechococcus* Fe(III)GlbN-R is ~6.1 [4]. *Synechococcus* GlbN-R is not as stable as *Synechocystis* GlbN-R at low pH (data not shown), and therefore a full pH dependence of the cross-linking rate was not obtained. Given the similarity between the two proteins [8], the reaction mechanism was assumed to be conserved. Conversion to GlbN-A following reduction was expected to be observable by manual mixing methods at high pH. We chose a pH of 9.2 as a compromise between rate of reaction and alkaline stability and we determined the kinetics as above. Addition of excess DT to Fe(III)GlbN-R reduced the entire protein population within the dead time of the experiment (<

1 s). Conversion to GlbN-A occurred in an apparent first order process as described by Equation 3.3:

$$\text{Fe(II)GlbN-R} \xrightarrow{k_{xl}} \text{Fe(II)GlbN-A} \qquad (3.3).$$

Isosbestic points were detected during the first 40 min of reaction (~8 half-lives), before heme bleaching caused by extended incubation with excess reducing agent became significant. A typical time course is shown in Figure S3.3 of the Supporting Information. At pH 9.2, the observed rate constant extracted from the visible spectra was $k_{xl} = (2.3 \pm 0.1) \times 10^{-3}$ s^{-1}. This led to the expectation that better than 90% conversion to Fe(II)GlbN-A would be achieved within the time required to prepare an NMR sample and set up the experiment (~12 min).

Kinetics of reaction upon substoichiometric Fe(III)GlbN-R reduction

A simple scenario predicts that if a substoichiometric amount of reducing agent is added to Fe(III)GlbN-R at pH 9.2, the reaction mixture will in short order contain Fe(III)GlbN-R and Fe(II)GlbN-A. Heme proteins, however, can undergo electron exchange, and the composition and stability of the mixture will depend on the rate of interprotein electron transfer (ET). Thus, it is possible that Fe(III)GlbN-R, Fe(III)GlbN-A, Fe(II)GlbN-R, and Fe(II)GlbN-A coexist at various times in solution. NMR spectroscopy was used to investigate this aspect of the post-translational modification.

In mixtures of GlbN-R and GlbN-A, several heme signals from Fe(III)GlbN-R and Fe(III)GlbN-A (both paramagnetic, $S = 1/2$ complexes) are resolved in the downfield region of the ^{1}H spectrum [8]. These signals, which report on relative

concentrations, can be followed as a function of reaction progress. In contrast, no signal from Fe(II)GlbN-R or Fe(II)GlbN-A (diamagnetic, S = 0 complexes) could be reliably monitored by 1D observation. 2D ^{1}H-^{15}N HSQC data offered the desired spectral resolution, but insufficient temporal resolution. An ^{15}N band-selective 1D HSQC experiment was used to circumvent this problem. In both GlbNs, the downfield-shifted amide of residue 80 (Thr in *Synechococcus* GlbN and Asn in *Synechocystis* GlbN) was chosen as a probe because oxidation state and cross-link status perturbed the ^{1}H shift without significantly affecting the ^{15}N shift. This is shown in Figure 3.3 for *Synechococcus* GlbN Thr80. Band selection eliminated overlapping signals and provided 1D ^{1}H spectra suitable for kinetic analysis (Figure 3.4a).

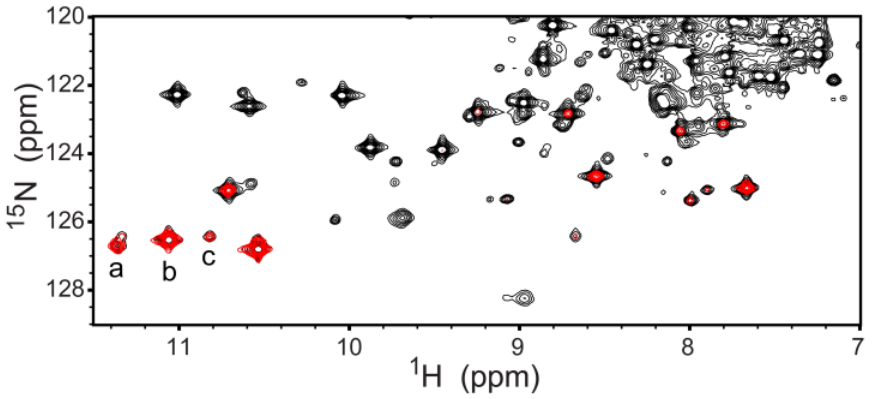


Figure 3.3. Portion of ^{1}H-^{15}N HSQC spectra collected 5 h after substoichiometric reduction of a *Synechococcus* Fe(III)GlbN-R sample. Black, conventional WATERGATE water flip-back spectrum; red, band-selected spectrum. Cross peaks a, b, and c are from Thr80 HN in Fe(II)GlbN-A, Fe(III)GlbN-A, and Fe(III)GlbN-R, respectively.

In the experiment shown in Figure 3.4a, the reaction was initiated by addition of DT such that approximately 40% of a 2 mM *Synechococcus* Fe(III)GlbN-R NMR sample was reduced. As predicted by the kinetics of cross-link formation (Equation 3.3) determined optically, a burst of Fe(II)GlbN-A production occurred during the dead time. But, this initial production of Fe(II)GlbN-A was followed by slow formation of Fe(III)GlbN-A at the expense of Fe(III)GlbN-R. A slight intensity decrease was observed in Fe(II)GlbN-A, which also occurred with the inert H117A variant (Figure S3.5). This minor phase was attributed to reoxidation unrelated to cross-link propagation. Data collected after 17 h confirmed the complete disappearance of Fe(III)GlbN.

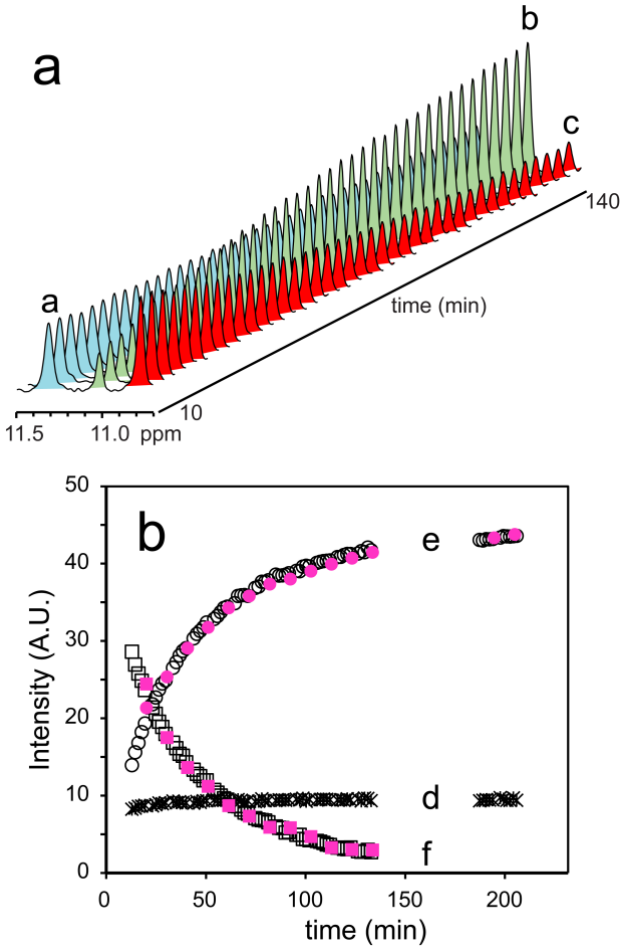


Figure 3.4. (a) Stacked plot of the 1D spectra obtained with ^{15}N band selection after substoichiometric reduction of *Synechococcus* Fe(III)GlbN-R. The time course of the GlbN-R → GlbN-A conversion is illustrated with the signals from Thr80 HN in Fe(II)GlbN-A (light blue, a), Fe(III)-GlbN-A (light green, b), and Fe(III)GlbN-R (red, c). The signal of Fe(II)GlbN-A initially contained intensity from unresolved Thr80 HN in Fe(II)GlbN-R (Figure S3.4). (b) The same experiment with *Synechocystis* GlbN. The signals are from Asn80 HN in Fe(II)GlbN-A (∗, d), Fe(III)-GlbN-A (○, e), and Fe(III)GlbN-R (□, f), and heme 5-CH_3 in Fe(III)GlbN-R (■) and Fe(III)GlbN-A (●). The heme signals were scaled arbitrarily to coincide with the Asn80 signals. Simulation and fits of the data are shown in Supporting Information.

The same experiment was performed with *Synechocystis* GlbN (Figure 3.4b). The sample was reduced with DT to obtain ~15% Fe(II)GlbN-R at the beginning of the reaction. As observed with *Synechococcus* GlbN, cross-link propagation continued to completion. Additionally, *Synechocystis* GlbN appeared to be less susceptible to autoxidation, and the level of reduced protein (Figure 3.4b, trace d) was nearly constant throughout the conversion.

A parsimonious model of the results invokes electron exchange between Fe(II)GlbN-A and Fe(III)GlbN-R (Scheme 3.1) combined with Fe(II)GlbN-R cross-linking (Equation 3.3). The bimolecular exchange step effectively recycles electrons to form reactive Fe(II)GlbN-R until all Fe(III)GlbN-R is exhausted. The kinetic data were fit accordingly to yield $K_{12} = k_{12}/k_{21}$ and a lower bound for k_{12} (Figures S6–S9). For *Synechococcus* GlbN (Figure 3.4a), K_{12} was ~1/10 and k_{12} > 4 M^{-1} s^{-1}. For *Synechocystis* GlbN (Figure 3.4b), K_{12} was ~ 1/4 and k_{12} > 10 M^{-1} s^{-1}. In both cases, the left-hand side of Scheme 3.1 was thermodynamically favored.

$$\text{Fe(III)GlbN-R} + \text{Fe(II)GlbN-A} \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} \text{Fe(II)GlbN-R} + \text{Fe(III)GlbN-A}$$

Scheme 3.1. Electron transfer in GlbN-R/GlbN-A

Electron self-exchange in GlbN

To test the plausibility of Scheme 3.1, we sought to determine the time scale of GlbN electron self-exchange (ESE). We reasoned that extremely slow rates would not be consistent with the observed progress of the conversion and that extremely fast rates would lead to a different NMR spectral appearance. We used two versions of *Synechococcus* GlbN: WT GlbN-A and H117A GlbN, which is incapable of cross-linking and serves as a GlbN-R surrogate. The ESE reactions are shown in Scheme 3.2, where * marks individual molecules and ^{117}GlbN represents the H117A variant. The ESE rate constants (k_{11} and k_{22}) were determined at pH 9.2 for consistency with the conversion conditions. In both GlbNs, the ^{1}H-^{15}N HSQC spectrum of a mixture of reduced and oxidized states was a superposition of unperturbed parent resonances, indicating that ESE was slow on the chemical shift time scale and that the putative encounter complexes were not populated to an appreciable extent. An ^{15}N ZZ exchange experiment [25] was therefore used for rate measurement.

$$\text{Fe(II)GlbN-A} + \text{Fe(III)GlbN-A*} \underset{k_{11}}{\overset{k_{11}}{\rightleftarrows}} \text{Fe(III)GlbN-A} + \text{Fe(II)GlbN-A*}$$

$$\text{Fe(II)}^{117}\text{GlbN} + \text{Fe(III)}^{117}\text{GlbN*} \underset{k_{22}}{\overset{k_{22}}{\rightleftarrows}} \text{Fe(III)}^{117}\text{GlbN} + \text{Fe(II)}^{117}\text{GlbN*}$$

Scheme 3.2. Electron self-exchange in WT GlbN-A and H117A GlbN

The inset of Figure 3.5 shows a portion of a *Synechococcus* GlbN-A ^{15}N ZZ spectrum (the full spectrum is shown in Figure S3.10). The pulse sequence generates Nz magnetization, which is allowed to exchange between the two redox partners during the mixing time. Four types of peaks forming a quartet are detected per residue: auto peaks have ^{15}N and ^{1}H chemical shifts corresponding to the oxidized or reduced species as one would observe in the ^{1}H-^{15}N HSQC spectra of the pure species, and exchange peaks have the ^{15}N shift of one species and the ^{1}H shift of the other. As the mixing time is varied, the intensity of the peaks changes according to exchange rates and ^{15}N longitudinal relaxation rates.

In Figure 3.5, the curves represent fits to Gln71 data using established expressions for magnetization exchange in 2D experiments [25, 35]. ^{15}N longitudinal relaxation rates (^{15}N T_1) extracted via exchange data fitting were consistent with the values measured independently (Figure S3.11). Global analysis of several quartets corresponding to residues in various secondary structure elements (e.g., Figure S3.12) confirmed that *Synechococcus* GlbN-A and H117A GlbN each registered a single exchange process. The rate constants were $k_{11} = (4.5 \pm 0.4) \times 10^2$ M^{-1} s^{-1} for WT GlbN-A and $k_{22} = (3.7 \pm 0.5) \times 10^2$ M^{-1} s^{-1} for H117A GlbN, with confidence intervals determined as described in

Figures S13 and S14. The ESE rate constant was also measured for *Synechocystis* GlbN-A and found to be (1.4 ± 0.4) × 10^3 M^{-1} s^{-1} (Figure S3.15), 3 times faster than for *Synechococcus* GlbN-A under the same conditions.

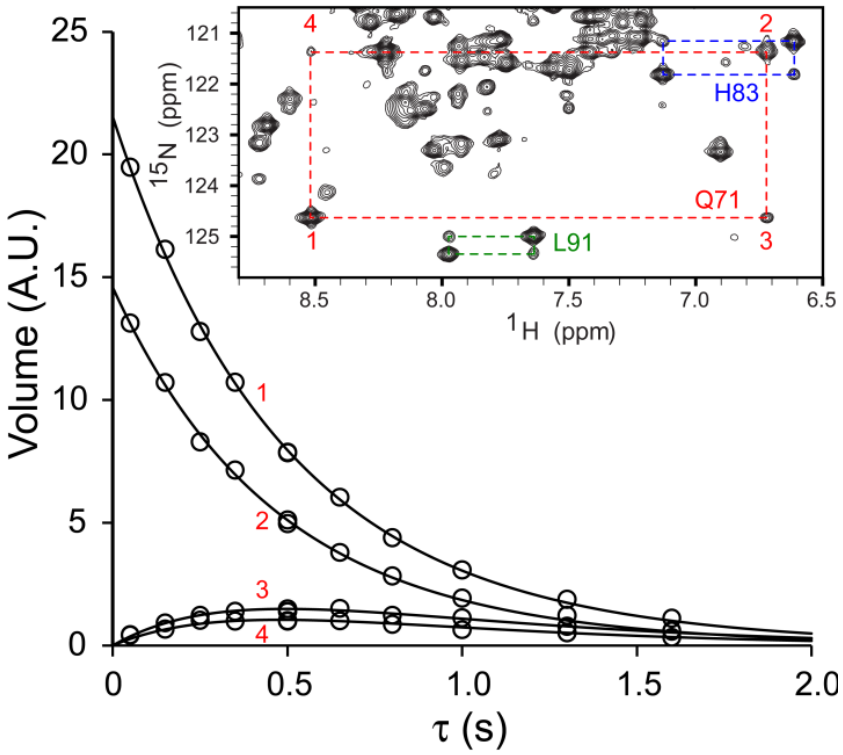


Figure 3.5. Electron self-exchange rate determination in *Synechococcus* GlbN-A using the time dependence of the auto (1, 2) and exchange (3, 4) peaks of Gln71 in the ^{15}N ZZ experiment. The inset shows a small portion of the ^{15}N ZZ exchange spectrum acquired at 600 MHz with τ ~ 0.5 s (pH 9.2, 298 K).

Although *Synechococcus* H117A GlbN differs from GlbN-R in some respects [36], the k_{22} value (Figure S3.12) argued for a minimal influence of the heme–protein

cross-link on the features controlling ESE. The experimental k_{ESE} values were compatible with the proposed inter-GlbN ET-mediated post-translational modification (Scheme 3.1 and Equation 3.3) and NMR spectral appearance.

The kinetics of ESE have been determined by NMR spectroscopy in many heme proteins under a broad range of conditions [37]. Measured ESE rate constants vary between 10^2 and 10^7 M^{-1} s^{-1}. Values for cytochrome b_5 range around 10^4 M^{-1} s^{-1} and those for cytochrome *c* around 10^5 M^{-1} s^{-1}. Rate constants on the order of 10^3 M^{-1} s^{-1} have been obtained for myoglobin and hemoglobin with bound trimethylphosphine [37]. The GlbN values were on par with these unnatural complexes and 10- to 100-fold slower than those for cytochromes b_5 and *c*. We note that the ability to measure ESE in a *native* form of hemoglobin is directly related to the stable bishistidine coordination of the iron in both the ferric and ferrous states. The relatively fast GlbN rate constants caution that kinetic analysis of bishistidine globin reactions involving redox processes may have to take into account interGlbN electron exchange and its concentration dependence.

The cellular biosynthetic machinery is expected to produce ferrous GlbN-R from apoprotein and heme. *Synechococcus* GlbN-R, however, can survive the cellular environment without undergoing the PTM and interestingly, when *Synechococcus* cells are grown under microoxic conditions, the relative levels of GlbN-A are higher than under normal conditions [4]. We have reported that CO, when bound to Fe(II)GlbN-R, retards significantly the conversion to GlbN-A [12]. This behavior is consistent with the proposed mechanism, which calls for stabilization of the carbocation (carbenium ion) with increased electron density

at the heme 2-vinyl (Figure 3.2). Similarly, O_2 binding is expected to be inhibitory. Whether or not the autocatalytic modification takes place in the cell is unclear as it depends on the rates of a number of processes (e.g., exogenous ligand off rate) and local protein concentration. The ESE rates, however, do suggest that GlbN-R and GlbN-A are capable of efficient electron transfer with each other and other proteins.

Estimation of the ET rate constants

The electron transfer reaction (Scheme 3.1), by nature of the partners, resembles a self-exchange reaction (Scheme 3.2). The parallel offers a favorable situation for further analysis of rate constants using Marcus theory. We therefore inspected the prediction of the cross-relation [38]

$$k_{12} = (k_{11}k_{22}K_{12}f_{12})^{1/2} \qquad (3.4)$$

which estimates the ET rate constant for the cross-reaction (k_{12}, Scheme 3.1) from the ESE rate constants of individual redox couples (k_{11} and k_{22}, Scheme 3.2). In Equation 3.4, f_{12} is an efficiency factor assumed to be unity and K_{12} is typically calculated with the known midpoint reduction potentials of individual couples [37]. The reactivity of bishistidine Fe(II)GlbN-R, however, complicates direct potential measurement, and instead we used K_{12} derived from Scheme 3.1 and Equation 3.3. For *Synechococcus* GlbN, the ET rate constant obtained from Equation 3.4 with K_{12} = 1/10, k_{11} = 450 M^{-1} s^{-1}, and k_{22} = 370 M^{-1} s^{-1} was k_{12} = 130 M^{-1} s^{-1}, well within the range allowed by kinetic data fitting (k_{12} > 4 M^{-1} s^{-1}) and in agreement with slow exchange on the chemical shift time scale. Likewise, using K_{12} = 1/4 and k_{11} = k_{22} = 1400 M^{-1} s^{-1} for *Synechocystis* GlbN

returned k_{12} = 700 M^{-1} s^{-1}, also allowed by the conversion data (k_{12} > 10 M^{-1} s^{-1}) and consistent with slow exchange. The K_{12} values corresponded to –60 mV (*Synechococcus*) and –35 mV (*Synechocystis*) differences in potential between the GlbN-R and GlbN-A couples (GlbN-R having the lower E°) and suggested a small effect from heme–protein cross-linking.

Structural comparison of the four complexes involved in the conversion

The cross-relation given above assumes identical work terms for GlbN-R and GlbN-A self-exchange reactions [38]. These terms capture the energy required to bring the reactants into an ET-competent configuration and depend on the structure of the proteins involved. As a first level of verification for the use of Equation 3.4, we gathered structural data on all participating species with the expectation that large differences would cast doubt on the validity of the approach and that highly similar partners would transfer electrons via the same interface and mechanism.

Structural similarity has been documented for *Synechocystis* WT GlbN-A and H117A GlbN in the crystalline state, leading to the conclusion that neither reduction nor cross-linking has a dramatic effect on this particular protein [39]. Little difference has been noted between *Synechococcus* WT Fe(III)GlbN-R and Fe(III)GlbN-A in solution [4], or between H117A Fe(III)GlbN and WT Fe(III)GlbN-R [36], also in solution. However, no information is available on reduced synechococcal GlbN. We therefore initiated the characterization of Fe(II)GlbN-A. This complex is diamagnetic and amenable to standard NMR experiments.

We have assigned the backbone and side chain resonances of *Synechococcus* Fe(III)GlbN-A [23]. Because exchange between the ferric and

ferrous states is slow on the chemical shift time scale, the ZZ experiment can in principle be used to transfer assignments from Fe(III)GlbN-A to Fe(II)GlbN-A. The spectral congestion displayed by the redox mixture (Figures 5 and S10) made it clear that a 2D approach would yield little information. The same data revealed that the difference in chemical shift (in Hz) in the ^{1}H dimension was generally much larger than that in the ^{15}N dimension. A 3D H-Nz-N-H exchange experiment was therefore designed to take advantage of the ^{1}H dispersion. Figure 3.6 illustrates the quality of the data with the ^{1}H–^{1}H projection of the 3D set. Each cross peak connects a known Fe(III)GlbN-A resonance to its counterpart in Fe(II)GlbN-A. A total of 60% of Fe(II)GlbN-A HN assignments were immediately obtained with this approach. Triple resonance backbone-based experiments resolved remaining ambiguities.

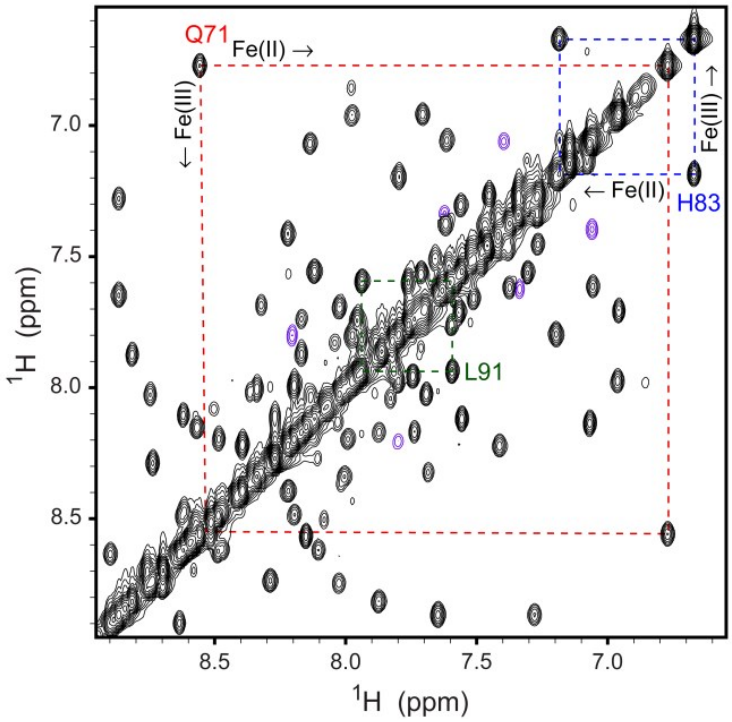


Figure 3.6. Portion of the ^{1}H-^{1}H projection of the H-Nz-N-H 3D experiment. Data were collected at 800 MHz (pH 9.2, 298 K) on a ~4:6 mixture of *Synechococcus* Fe(II) and Fe(III)GlbN-A. The mixing time was ~0.5 s. Peaks in magenta were folded in the ^{15}N dimension.

The backbone chemical shifts of *Synechococcus* Fe(II)GlbN-A were used to map the secondary structure with the program TALOS+ [29]. The helices of GlbN did not appear to be affected by reduction of the iron (Figure S3.17). Tertiary structure was assessed indirectly with the pseudocontact shift experienced by backbone protons in Fe(III)GlbN-A. If Fe(III)GlbN-A and Fe(II)GlbN-A are isostructural, $\Delta\delta$ of individual ^{1}Hs is equal to the hyperfine shift, composed of a contact contribution extending over a few bonds from the paramagnetic center and a pseudocontact contribution (δ_{pc}) effective through space [30]. Because δ_{pc} is sensitive to the position of the nucleus relative to the paramagnetic center (Equation 3.2, [30]), a comparison of experimental $\Delta\delta$ with δ_{pc} calculated from the coordinates of one of the structures tests the extent of structural similarity.

A total of 114 backbone amide ^{1}H and 18 ^{1}Hα $\Delta\delta$ values were used as restraints for XPLOR-NIH refinement of a low-energy member of the 2KSC family and determination of the magnetic susceptibility tensor [32, 33]. The refined *Synechococcus* Fe(III)GlbN-A coordinates were within the bundle of the original ensemble (Figure S3.18), and calculated δ_{pc} values corresponded well to the experimental $\Delta\delta$ values (Figure 3.7). For example, Gln71 HN, which is adjacent to the proximal histidine and was used in the evaluation of the self-exchange rate (Figure 3.5), experienced a δ_{pc} of 1.76 ppm, calculated to be 1.84

ppm. The generally good correlation shown in Figure 3.7 supported that the *average* structures of Fe(II)GlbN-A and Fe(III)GlbN-A were similar.

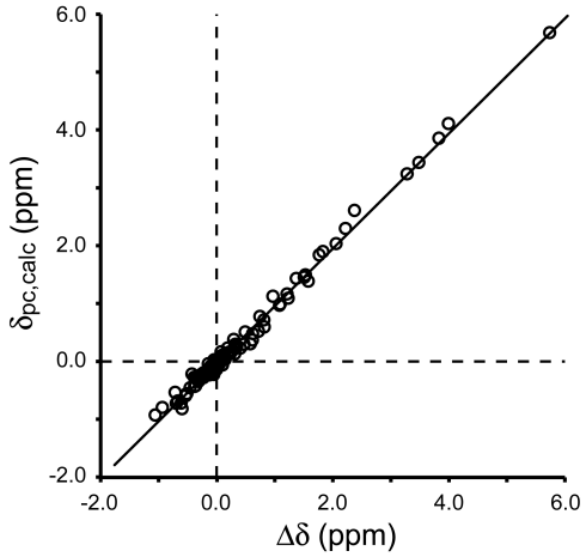


Figure 3.7. Correlation between observed and calculated pseudocontact shifts for *Synechococcus* Fe(III)GlbN-A after structural refinement using $\Delta\delta$ as restraints in addition to NOEs, H-bonds, and dihedral angles ($\delta_{pc,calc}$ = 0.997 $\Delta\delta$ – 0.019, R^2 = 0.9924). χ Tensor information is provided in the Supporting Information, Figures S19–S21.

Exchange NMR data collected at pH 9.2 on partially reduced H117A GlbN samples (not shown) indicated that the variant maintained its fold in both oxidation states. We therefore concluded that, as for *Synechocystis* GlbN, neither reduction nor cross-linking perturbed the structure significantly. In that respect, Scheme 3.1 (ET) and Scheme 3.2 (ESE) were considered identical.

The aggregated observations support our application of the cross-relation (Equation 3.4) to test the consistency of the kinetic measurements. GlbN emerges from our studies as an excellent system for ET investigations under a

complementary array of conditions (e.g., temperature and ionic strength). Thus far, we have found no NMR evidence for a chemically gated mechanism. Characterization of the encounter complex may offer insight into how similar the ET mechanism and intrinsic ET capabilities of GlbN are to those of established electron transfer proteins. The midpoint potential difference and k_{12} values will remain tentative until additional evidence confirms that the observed rates are governed by ET [40].

Globin function, globin properties, and bishistidine coordination

The best-known globin function is dioxygen transport and storage. Nitric oxide dioxygenation is a second, well-documented role performed by flavohemoglobins [41], certain single domain canonical (3/3) globins [42, 43], and 2/2 globins [44]. It is now recognized that a large number of Hbs exist as bishistidine complexes. Because distal histidine coordination can interfere with exogenous ligand binding and identical coordination in the ferric and ferrous states facilitates electron transfer, attention has recently turned to this type of process as yet another function represented in the Hb superfamily [45, 46]. Among bishistidine Hbs, the strength of the distal histidine ligation as measured by the equilibrium constant between a "pentacoordinate" His–Fe species and a hexacoordinate His–Fe–His species [47] varies greatly from protein to protein. Two bishistidine hemoglobins in *C. elegans* (GLB-6 and GLB-26) constitute extreme examples of stable ligation as they do not bind O_2 or CO [46]. These Hbs have been proposed to serve as ET proteins.

The two GlbNs of interest here bind exogenous ligands. The full complement of distal histidine coordination effects has yet to be explored. Thus far, we have

shown that His46 ligation protects the heme from damage by H_2O_2 [12], which is likely an advantage as GlbNs function under cellular conditions generating peroxides. We also have shown that cyanide binding and the H46A or H46L replacement alter the regiospecificity of the post-translational modification [48]. *Synechococcus* GlbN has been implicated in the detoxification of reactive nitrogen species [4] and if, as suspected, it is directly involved in NO or peroxynitrite processing, the reaction mechanism is anticipated to require redox cycling of the iron. Further studies will reveal whether access to a relatively stable bishistidine state in both ferric and ferrous states is an important feature of GlbN related to electron transfer in the context of enzymatic turnover. In any case, comparative studies of hexacordination stability and ET in other Hbs will provide molecular insight into the determinants of ligand binding and electron transfer functions.

Conclusion

We have shown that reducing any amount of bishistidine GlbN in a pool of oxidized bishistidine GlbN will trigger a chain reaction converting the entire sample to GlbN-A. To our knowledge, this is the first example of a redox chain reaction leading to a physiologically pertinent heme modification. We have related this amplification of the modification to the ease with which GlbN exchanges electrons. Our observations highlight the ability of bishistidine hemoglobins to participate in redox reactions and more generally illustrate the ability of a heme protein to catalyze its own post-translational modification.

Acknowledgement

This study was supported by National Science Foundation grant MCB-0349409. NMR facilities and resources at Johns Hopkins University were provided by the Biomolecular NMR Center. The authors thank Selena Rice for assistance with the optical measurements, Richard Himes, Ryan Peterson and Dr. Kenneth Karlin for the use of their stopped-flow equipment, and Dr. Christopher Falzone for useful discussions and careful reading of the manuscript. Dr. Henry Nothnagel's insight into heme chemistry was essential in the initial phases of the work. Figure 3.1a was prepared with Molscript [49].

References

1. Vinogradov SN, Moens L (2008) J. Biol. Chem. 283:8773-8777

2. Kakar S, Sturms R, Tiffany A, Nix JC, DiSpirito AA, Hargrove MS (2011) Biochemistry 50:4273-4280

3. Scott NL, Falzone CJ, Vuletich DA, Zhao J, Bryant DA, Lecomte JTJ (2002) Biochemistry 41:6902-6910

4. Scott NL, Xu Y, Shen G, Vuletich DA, Falzone CJ, Li Z, Ludwig M, Pond MP, Preimesberger MR, Bryant DA, Lecomte JTJ (2010) Biochemistry 49:7000-7011

5. Couture M, Das TK, Savard PY, Ouellet Y, Wittenberg JB, Wittenberg BA, Rousseau DL, Guertin M (2000) Eur. J. Biochem. 267:4770-4780

6. Scott NL, Lecomte JTJ (2000) Protein Sci. 9:587-597

7. Vu BC, Jones AD, Lecomte JTJ (2002) J. Am. Chem. Soc. 124:8544-8545

8. Vu BC, Vuletich DA, Kuriakose SA, Falzone CJ, Lecomte JTJ (2004) J. Biol. Inorg. Chem. 9:183-194

9. Bowman SE, Bren KL (2008) Nat. Prod. Rep. 25:1118-1130

10. Pearson AR, Elmore BO, Yang C, Ferrara JD, Hooper AB, Wilmot CM (2007) Biochemistry 46:8340-8349

11. Huang L, Wojciechowski G, Ortiz de Montellano PR (2006) J. Biol. Chem. 281:18983-18988

12. Nothnagel HJ, Preimesberger MR, Pond MP, Winer BY, Adney EM, Lecomte JTJ (2011) J. Biol. Inorg. Chem. 16:539-552

13. Vuletich DA, Falzone CJ, Lecomte JTJ (2006) Biochemistry 45:14075-14084

14. Englander SW, Calhoun DB, Englander JJ (1987) Anal. Biochem. 161:300-306

15. Di Iorio EE (1981) Methods Enzymol 76:57-72

16. Johnson KA, Simpson ZB, Blom T (2009) Anal. Biochem. 387:30-41

17. Johnson KA (2009) Methods Enzymol. 467:601-626

18. Bilsel O, Zitzewitz JA, Bowers KE, Matthews CR (1999) Biochemistry 38:1018-1029

19. Falzone CJ, Vu BC, Scott NL, Lecomte JTJ (2002) J. Mol. Biol. 324:1015-1029

20. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) J. Biomol. NMR 6:277-293

21. Goddard TD, Kneller DG (2006) University of California, San Francisco

22. Falzone CJ, Lecomte JTJ (2002) J. Biomol. NMR 23:71-72

23. Pond MP, Vuletich DA, Falzone CJ, Majumdar A, Lecomte JTJ (2009) Biomol. NMR Assign. 3:211-214

24. Geen H, Freeman R (1991) J. Magn. Reson. 93:93-141

25. Farrow NA, Zhang O, Forman-Kay JD, Kay LE (1994) J. Biomol. NMR 4:727-734

26. Emsley L, Bodenhausen G (1990) Chem. Phys. Lett. 165:469-476

27. Kay LE, Torchia DA, Bax A (1989) Biochemistry 28:8972-8979

28. Piotto M, Saudek V, Sklenár V (1992) J. Biomol. NMR 2:661-665

29. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) J. Biomol. NMR 44:213-223

30. Emerson SD, La Mar GN (1990) Biochemistry 29:1556-1566

31. Brünger AT (1992) X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR. Yale University Press, New Haven

32. Banci L, Bertini I, Cavallaro G, Giachetti A, Luchinat C, Parigi G (2004) J. Biomol. NMR 28:249-261

33. Schwieters CD, Kuszewski JJ, Clore GM (2006) Progr. NMR Spectroscopy 48:47-62

34. Schmitz C, Stanton-Cook MJ, Su XC, Otting G, Huber T (2008) J. Biomol. NMR 41:179-189

35. Jeener J, Meier BH, Bachmann P, Ernst RR (1979) J. Chem. Phys. 71:4546-4553

36. Vu BC, Nothnagel HJ, Vuletich DA, Falzone CJ, Lecomte JTJ (2004) Biochemistry 43:12622-12633

37. Simonneaux G, Bondon A (2005) Chem. Rev. 105:2627-2646

38. Marcus RA, Sutin N (1985) Biochim. Biophys. Acta 811:265-322

39. Hoy JA, Smagghe BJ, Halder P, Hargrove MS (2007) Protein Sci. 16:250-260

40. Davidson VL (2000) Acc. Chem. Res. 33:87-93

41. Bonamore A, Boffi A (2008) IUBMB Life 60:19-28

42. Gardner PR (2005) J. Inorg. Biochem. 99:247-266

43. Gardner PR, Gardner AM, Brashear WT, Suzuki T, Hvitved AN, Setchell KD, Olson JS (2006) J. Inorg. Biochem. 100:542-550

44. Ouellet H, Ouellet Y, Richard C, Labarre M, Wittenberg B, Wittenberg J, Guertin M (2002) Proc. Natl. Acad. Sci. U. S. A. 99:5902-5907

45. Fago A, Mathews AJ, Moens L, Dewilde S, Brittain T (2006) FEBS Lett. 580:4884-4888

46. Kiger L, Tilleman L, Geuens E, Hoogewijs D, Lechauve C, Moens L, Dewilde S, Marden MC (2011) PLoS ONE 6:e20478

47. Kakar S, Hoffman FG, Storz JF, Fabian M, Hargrove MS (2010) Biophys. Chem. 152:1-14

48. Nothnagel HJ, Love N, Lecomte JT (2009) J. Inorg. Biochem. 103:107-116

49. Kraulis P (1991) J. Appl. Crystallogr. 24:946-950

Supporting information

Electron self-exchange and self-amplified post-translational modification in the hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002

1. Materials

Chemical	Source
Argon	Airgas
bovine hemin chloride	Sigma
bovine liver catalase	Sigma
99.9 % D_2O	Sigma
D-(+)-glucose	Sigma
A. niger glucose oxidase	Sigma
^{15}N ammonium chloride	Sigma
sodium dithionite 85% (DT)	Sigma
urea	J.T. Baker
all other chemicals	Sigma

2. Protein expression and purification: WT and H117A *Synechococcus* GlbN, WT *Synechocystis* GlbN

The following is based on the procedures described in [1-4]. Briefly, wild-type (WT) and variant GlbNs were overexpressed in BL21(DE3) *E. coli* cells. The proteins partitioned primarily into the insoluble fraction. Purified inclusion bodies were solubilized in a 50 mM Tris 1 mM EDTA buffer (pH ~8) containing 8 M urea and then loaded onto a G-50 size exclusion column for concurrent

apoGlbN refolding and purification. Fractions containing apoGlbN were identified by SDS-PAGE. Holoproteins were generated by coarse titration with hemin chloride dissolved in 0.1 M NaOH. Excess hemin was removed by passage of the GlbN solution over a DEAE anion exchange column after filtering and centrifugation. GlbN elution was achieved with a 0–0.5 M NaCl gradient. Fractions deemed pure by SDS-PAGE were pooled, concentrated by ultrafiltration, and exchanged into water by dialysis. The purified GlbNs were lyophilized and stored at –20 °C for the long term. The preparation method yielded Fe(III)GlbN (oxidized holoprotein, without post-translational modification). Yields were typically 25 mg /L medium.

Protein concentration determination

Protein concentrations were determined optically. ApoGlbN extinction coefficients were obtained from the ExPASY ProtParam server. HoloGlbN extinction coefficients were determined on a per-heme basis using the pyridine hemochromogen method.[4, 5] Values are listed below. The H117A GlbN variants were assumed to have the same extinction coefficient as the WT protein.

Protein	ε (mM^{-1} cm^{-1})	λ (nm)
apoGlbN		
Synechococcus WT	4.47	280
Synechocystis WT	7.36	280
holoGlbN (Fe(III), bishistidine)		
Synechococcus WT GlbN-R	96	411
Synechococcus WT GlbN-A	87	409
Synechocystis WT GlbN-R	100	411

3. NMR spectroscopy (Methods)

Band-selective 1D and 2D ^{15}N-^{1}H HSQC spectroscopy

Standard 2D ^{15}N-^{1}H HSQC experiments were adequate to resolve a small number of NMR signals from Fe(III)GlbN-R, Fe(II)GlbN-A, and Fe(III)GlbN-A. To monitor the conversion of GlbN-R to GlbN-A, we used substoichiometric reduction at high pH (9.2) in order to slow down the reaction. However, significant changes in the concentration of Fe(III)GlbN-R/-A still occurred during 2D acquisition (12 min per experiment), in particular at the beginning when accurate data mattered most. This behavior manifested itself by artificially narrowing (Fe(III)GlbN-A, concentration increasing during indirect acquisition) and broadening (Fe(III)GlbN-R, concentration decreasing during indirect acquisition) HSQC cross peaks in the ^{15}N dimension. The evolving concentrations invalidated a simple quantitative relationship between cross peak volume and concentration, and the 2D approach was abandoned. Raising the pH further was not acceptable because of sample viability.

2D ^{15}N-^{1}H ZZ exchange NMR spectroscopy

The concentration of ferric and ferrous GlbNs must remain constant over the course of the ^{15}N ZZ mixing series in order to extract reliable exchange and relaxation parameters. Experiments were initiated after the solutions had reached a stable ratio of Fe(II) and Fe(III)GlbN. The concentrations were estimated by comparing the integrated intensities of corresponding ferrous and ferric NH peaks obtained from fully relaxed (relaxation delay = 5 s) ^{1}H-^{15}N HSQC experiments. Fe(III):Fe(II) ratios were estimated from these data and were in good agreement with the fitted autopeak intensities (Fe(II)GlbN: $I_{22}(0)$, Fe(III)GlbN: $I_{33}(0)$) at $\tau_{mix} = 0$.

4. Kinetics of the GlbN-R → GlbN-A conversion in partially reduced samples

The chemical shift of Thr80 HN (Asn80 in *Synechocystis* GlbN) in each form of *Synechococcus* GlbN (Fe(II)GlbN-R/-A and Fe(III)GlbN-R/-A), is shifted far downfield owing to strong hydrogen bonding with the side chain of His83. Since the ^{15}N chemical shift is nearly degenerate in each form, it was assumed that the band-selective pulse excited Thr80 (Asn80) ^{15}N resonances uniformly. Hence, peak areas obtained using the Topspin 2.1 deconvolution routine were used without correction as target input for kinetic modeling. Although the 80 HN signals from Fe(III)GlbN-R and Fe(III)GlbN-A are resolved because of the distinct magnetic susceptibility tensors of these two paramagnetic species, we suspect that the 80 HN signals of Fe(II)GlbN-R and Fe(II)GlbN-A (both diamagnetic) overlap. We cannot generate a sample of pure wild-type Fe(II)GlbN-R and therefore used *Synechococcus* H117A Fe(II)GlbN as a wild-type

GlbN-R mimic to confirm the expectation. Figure S3.4 shows that the two reduced state cross peaks are not resolved from each other.

The following reactions were used to model the substoichiometric conversion of GlbN at pH 9.2:

$$\text{Fe(II)GlbN-R} \rightarrow \text{Fe(II)GlbN-A} \quad (1)$$

k_{xl} fixed to 0.0023 s^{-1} as obtained from the optical data (*Synechococcus* GlbN)

or k_{xl} fixed to 0.005 s^{-1}, interpolated from the optical data (*Synechocystis* GlbN)

$$\text{Fe(II)GlbN-A} + \text{Fe(III)GlbN-R} \leftrightarrows \text{Fe(III)GlbN-A} + \text{Fe(II)GlbN-R} \quad (2)$$

ET step, k_{12} and k_{21} allowed to float

These steps accounted for cross-link propagation. *Synechococcus* GlbN data modeling required an additional minor oxidation phase:

$$\text{Fe(II)GlbN-A} + \text{ox} \rightarrow \text{Fe(III)GlbN-A} + \text{red} \quad (3)$$

$k_{ox,app}$ allowed to float

This last reaction appeared unrelated to the cross-linking reaction because it was also observed with the unreactive H117A variant (Figure S3.5). It was assumed that it occurred as a result of contamination with an oxidant (dioxygen). An arbitrary ox concentration was used in Equation 3.3. The oxidation step was unnecessary to describe the self-amplified conversion of *Synechocystis* GlbN-R to GlbN-A.

Although only a lower limit on individual k_{12} and k_{21} values could be determined reliably, their ratio $K_{12} = k_{12}/k_{21}$ was well described by the data. Since electron self-exchange (ESE) was detected between free GlbN monomers and no additional species was observed in solution, it was likely that GlbN encounter complexes (EC) were weakly populated and in fast exchange on the chemical shift timescale. There was no evidence of Fe(III)GlbN dimerization in samples up to 5 mM protein. If one assumes that k_{12} and k_{21} report on ET directly or that association and dissociation steps are fast and symmetrical about encounter complex formation (identical work terms):

$$\text{Fe(II)GlbN-A} + \text{Fe(III)GlbN-R} \leftrightarrows \text{EC} \leftrightarrows \text{EC*} \leftrightarrows \text{Fe(III)GlbN-A} + \text{Fe(II)GlbN-R} \quad (4),$$

then the ratio k_{12}/k_{21} may be interpreted in terms of redox potential using:

$$\Delta E^\circ = -\Delta G^\circ/nF = (RT \ln K_{12})/nF \quad (5)$$

For *Synechococcus* GlbN, K_{12} was 0.1, returning $\Delta E° \sim -60$ mV (GlbN-R lower potential). For *Synechocystis* GlbN, K_{12} was 0.25, returning $\Delta E° \sim -35$ mV (GlbN-R lower potential). The latter $\Delta E°$ is about 50 mV less than the difference in measured redox potentials between WT *Synechocystis* GlbN-A and H117A *Synechocystis* GlbN (–85 mV, H117A lower) [6]. This suggests either that the observed K_{12} does not report directly on ET and/or that H117A *Synechocystis* GlbN is an imperfect analogue for WT *Synechocystis* GlbN-R redox behavior.

5. Determination of electron self-exchange rate

The exchange of longitudinal $^{15}N_Z$ magnetization as a function of τ_{mix} generates self (I_{22} and I_{33}) and cross (I_{23} and I_{32}) peaks whose integrated intensities are described by the Bloch-McConnell equations [7].

In the GlbN analyses, fitting was reserved for residues that exhibited completely resolved $^{15}N_Z$ exchange quartets and similar ^{1}H line widths in the Fe(II) and Fe(III) states (data for WT *Synechococcus* GlbN-A shown, Figure S3.10, with probe location in the 2KSC structure shown in Figure S3.12). The proximal coordinating histidine (His70), though resolved, was removed from the analysis because of its non-negligible paramagnetic relaxation enhancement in the ferric state. No correction was made for differential relaxation during INEPT transfers. Peak volumes were estimated by Gaussian simulation using the program Sparky. The relevant equations are [8]:

$$I_{22}(\tau) = I_{22}(0)[-(\lambda_2-a_{11})\exp(-\lambda_1\tau) + (\lambda_1-a_{11})\exp(-\lambda_2\tau)]/(\lambda_1-\lambda_2)$$

$$I_{33}(\tau) = I_{33}(0)[-(\lambda_2-a_{22})\exp(-\lambda_1\tau) + (\lambda_1-a_{22})\exp(-\lambda_2\tau)]/(\lambda_1-\lambda_2)$$

$$I_{23}(\tau) = I_{22}(0)[a_{21}\exp(-\lambda_1\tau)-a_{21}\exp(-\lambda_2\tau)]/(\lambda_1-\lambda_2)$$

$$I_{32}(\tau) = I_{33}(0)[a_{12}\exp(-\lambda_1\tau)-a_{12}\exp(-\lambda_2\tau)]/(\lambda_1-\lambda_2)$$

with $\lambda_{1,2} = (1/2)\{(a_{11} + a_{22}) \pm [(a_{11} - a_{22})^2 + 4k_{23}k_{32}]^{1/2}\}$ and $a_{11} = R_{1,2} + k_{23}$, $a_{22} = R_{1,3} + k_{32}$, $a_{12} = -k_{32}$, $a_{21} = -k_{23}$. I_{22} and I_{33} refer to the intensity of the Fe(II) auto peak and Fe(III) auto peak, respectively; I_{23} refers to the intensity of the exchange cross peak with Fe(II) ^{15}N and Fe(III) ^{1}H shifts; I_{32} refers to the intensity of the exchange cross peak with Fe(III) ^{15}N and Fe(II) ^{1}H shifts; k_{23} refers to the apparent first order rate constant for Fe(II) $\rightarrow$ Fe(III), and k_{32} refers to the apparent first order rate constant for Fe(III) $\rightarrow$ Fe(II).

The data (truncated to $\tau_{mix} < 500$ ms) were first inspected by calculating the $\Xi(t)$ ratio, which is unaffected by relaxation [9].

$$\Xi(t) = \frac{I_{23}(t)I_{32}(t)}{I_{22}(t)I_{33}(t) - I_{23}(t)I_{32}(t)} = k_{32}k_{23}t^2$$

Global fitting (Scilab 5.2, not shown) indicated that all chosen residues (and presumably the whole protein) registered a single exchange process. A local nonlinear least squares fitting routine (Scilab 5.2) was also employed to extract preliminary guesses for exchange (k_{23}, k_{32}), relaxation ($R_{1,2}$, $R_{1,3}$), and intensity ($I_{22}(0)$, $I_{33}(0)$) parameters by minimization of the sum of the squared error residuals.

Uncertainties in peak volumes were re-estimated by calculating the standard deviation of the residuals between best-fit curves and data. Uncertainties estimated in this fashion were of similar magnitude as those estimated from the analysis of duplicate τ_{mix} points. The program KinTek Explorer was used to extract the reported exchange and relaxation parameters by simulation and nonlinear least squares fitting of integrated peak intensity data directly to the kinetic scheme:

$$Fe(II)N_z(eq) \leftarrow Fe(II)N_z(\tau_{mix}) \leftrightarrows Fe(III)N_z(\tau_{mix}) \rightarrow Fe(III)N_z(eq)$$

Owing to a small uncertainty in the relative concentration of Fe(II) and Fe(III) GlbNs, $I_{22}(0)$ and $I_{33}(0)$ were allowed to float independently in this analysis. The relative intensity of initial Fe(II) and Fe(III) GlbN magnetization ($I_{22}(0)/I_{33}(0)$) converged to values that were in good agreement with those determined by analysis of fully-relaxed HSQC data. Extracted ^{15}N R_1 values were generally within error of those determined by conventional methods (these are shown in Figure S3.10 for *Synechococcus* WT GlbN-A and illustrate the influence of the paramagnetic center). Residue-specific exchange parameters were within error of one another and further confirmed that a single exchange event was sufficient to characterize the ET reaction between free Fe(II) and Fe(III) GlbN monomers in solution.

KinTek Explorer confidence contours were calculated on a per-residue basis (one example is shown in Figure S3.13). Individual best-fit parameters were allowed to vary at least 20-fold when computing confidence contours. A threshold $\chi^2_{limits}/\chi^2_{min}$ value of 1.5 was used to evaluate the confidence contour error bounds. The 1.5-fold χ^2_{min} acceptable error boundary was chosen based on the following information: number of data points = 44–48, number of fitted parameters = 6 (k_{23}, k_{32}, $R_{1,2}$, $R_{,3}$ and $I_{22}(0)$, $I_{33}(0)$ scaling factors), and degrees of freedom = 38–42.

This method of analysis was applied to H117A *Synechococcus* GlbN (4 resolved ZZ quartets) and WT *Synechocystis* GlbN-A (3 resolved ZZ quartets). Figure S3.14 shows the fitted data used to provide k_{22} in the text for H117A *Synechococcus* GlbN, and Figure S3.15 presents an example N_z evolution curve used to estimate the kinetics of ESE in WT *Synechocystis* GlbN-A.

6. Structural calculations

The magnetic susceptibility tensor obtained by refinement had moderate anisotropy: $\Delta\chi_{ax} = (3.94 \pm 0.07) \times 10^{-32}$ m^3 and $\Delta\chi_{rh} = (1.07 \pm 0.06) \times 10^{-32}$ m^3. The Euler angles were $\alpha = 257°$, $\beta = -16°$, $\gamma = 27°$. (Note that Numbat [10] uses the "ZYZ" convention for Euler rotation.) The molecular frame of reference had x axis along the Fe–NC direction and z axis pointing toward the proximal histidine; the coordinates of the Fe center were fixed at the origin. Tensor representations in Figures S19–S21 were prepared with Numbat [10]. Additional spectral assignments will be necessary to improve the tensor calculation.

SUPPORTING INFORMATION REFERENCES

1. Scott NL, Lecomte JTJ (2000) Protein Sci. 9:587-597
2. Scott NL, Falzone CJ, Vuletich DA, Zhao J, Bryant DA, Lecomte JTJ (2002) Biochemistry 41:6902-6910
3. Vuletich DA, Falzone CJ, Lecomte JTJ (2006) Biochemistry 45:14075-14084
4. Nothnagel HJ, Preimesberger MR, Pond MP, Winer BY, Adney EM, Lecomte JTJ (2011) J. Biol. Inorg. Chem. 16:539-552
5. de Duve C (1948) Acta Chem. Scan. 2:264-289
6. Hoy JA, Smagghe BJ, Halder P, Hargrove MS (2007) Protein Sci. 16:250-260
7. John M, Headlam MJ, Dixon NE, Otting G (2007) J. Biomol. NMR 37:43-51
8. Farrow NA, Zhang O, Forman-Kay JD, Kay LE (1995) Biochemistry 34:868-878
9. Miloushev VZ, Bahna F, Ciatto C, Ahlsen G, Honig B, Shapiro L, Palmer AG, 3rd (2008) Structure 16:1195-1205
10. Schmitz C, Stanton-Cook MJ, Su XC, Otting G, Huber T (2008) J. Biomol. NMR 41:179-189

Table S3.1. Electron self-exchange rate in wild-type *Synechococcus* GlbN-A. Values were determined at 298 K and pH 9.2. The second order rate constant for ESE was obtained by summing k_{23} and k_{32} and dividing by the total concentration of protein (~2 mM). Subscripts 2 and 3 stand for Fe(II) and Fe(III), respectively.

Residue (local)		$R_{1,2}$	$R_{1,3}$	k_{23} (s^{-1})	k_{32} (s^{-1})
Leu91, best fit		1.66	1.67	0.52	0.42
	lower bound	1.56	1.58	0.41	0.33
	upper bound	1.80	1.75	0.62	0.50
Leu122, best fit		1.56	1.64	0.61	0.36
	lower bound	1.39	1.58	0.48	0.34
	upper bound	1.71	1.70	0.71	0.39
Gln71, best fit		1.72	1.65	0.42	0.39
	lower bound	1.64	1.60	0.38	0.36
	upper bound	1.79	1.69	0.47	0.42
Phe35, best fit		1.61	1.72	0.50	0.34
	lower bound	1.50	1.64	0.44	0.27
	upper bound	1.72	1.81	0.57	0.40
Arg124, best fit		1.57	1.62	0.56	0.44
	lower bound	1.41	1.51	0.47	0.39
	upper bound	1.73	1.73	0.65	0.48

Table S3.2. Electron self-exchange rate in H117A *Synechococcus* GlbN. Values determined at 298 K and pH 9.2. The total concentration of protein was ~1.7 mM, which was used to calculate the second order rate constant for ESE as described above. Subscripts 2 and 3 stand for Fe(II) and Fe(III), respectively.

Residue (local)		$R_{1,2}$	$R_{1,3}$	k_{23} (s^{-1})	k_{32} (s^{-1})
Gln47, best fit		1.72	1.67	0.45	0.17
	lower bound	1.61	1.58	0.41	0.15
	upper bound	1.84	1.75	0.48	0.19
Peak X, best fit		1.64	1.63	0.44	0.14
	lower bound	1.53	1.55	0.38	0.13
	upper bound	1.77	1.70	0.50	0.16
Gln71, best fit		1.68	1.66	0.40	0.16
	lower bound	1.57	1.55	0.34	0.15
	upper bound	1.81	1.78	0.45	0.17
Arg124, best fit		1.52	1.61	0.58	0.17
	lower bound	1.41	1.53	0.49	0.15
	upper bound	1.63	1.68	0.67	0.20

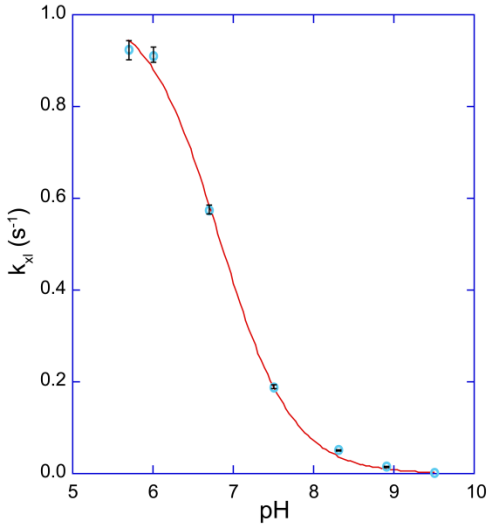


Figure S3.1. pH dependence of the cross-linking rate constant for *Synechocystis* WT GlbN. The k_{xl} value was determined in triplicate at each pH. The solid line represents a fit with pK_a = 6.83 ± 0.06 and Hill coefficient n = 0.96 ± 0.09 (Equation 3.1 in the text). The maximum rate constant was estimated as $k_{xl,max}$ = 1.02 ± 0.04 s^{-1}. Because of the lack of an acid baseline, the pK_a value is tentative. Error bars were generated by taking the maximal upper and lower bounds of k_{xl} derived from KinTek Explorer confidence contour analysis using a threshold $\chi^2_{limits}/\chi^2_{min}$ value of 1.2. The pH 9.5 k_{xl} value has been previously determined by manual mixing methods [4], and was included in this analysis.

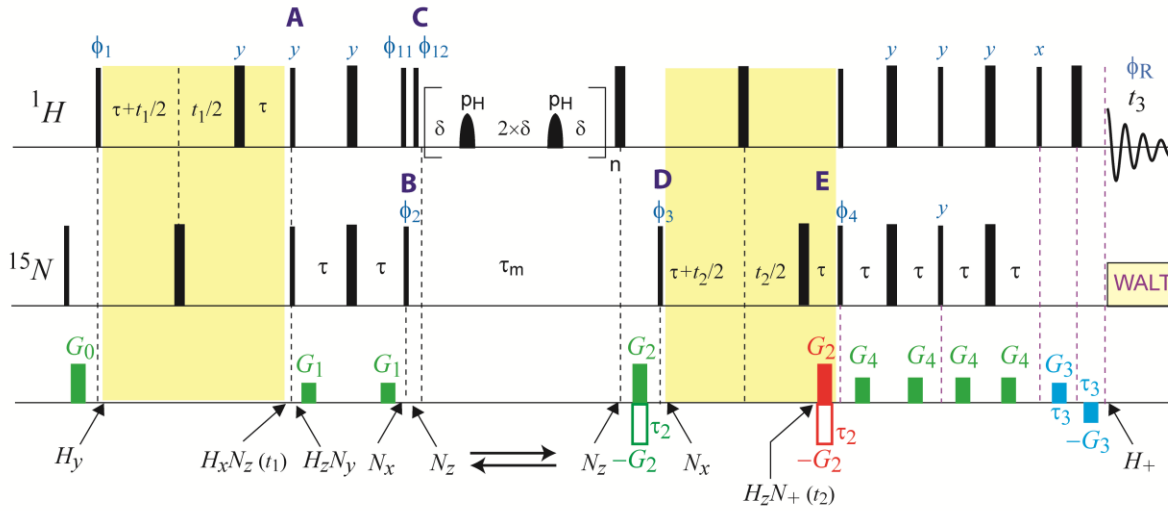


Figure S3.2. Pulse sequence for the SE H–Nz-N–H ZZ exchange experiment. Narrow and wide black bars denote 90° and 180° pulses, respectively. $\phi_1 = x, -x$ + States/TPPI; $\phi_{11} = y$ + States/TPPI; $\phi_{12} = -x, x$; $\phi_2 = 2(y),2(-y)$; $\phi_3 = 4(y),4(-y)$ + States/TPPI; $\phi_4 = x$ + echo/antiecho; $\phi_R = x, -x, -x, x, -x, x, x, -x$. The minimum number of scans is 4. At point A, ^{1}H magnetization has been frequency labeled (t_1 max = 25 ms) with INEPT transfer achieved by setting τ to 1/(4J_{NH}) = 2.3 ms. Anti-phase H_zN_y magnetization evolves to in-phase $^{15}N_x$ during the second INEPT period and is placed along the z-axis after the ^{15}N pulse at point B. Quadrature detection of ^{1}H(ω_1) is achieved through States-TPPI cycling. Additional phase cycling of the initial ^{1}H 90° and the ^{1}H 90° at point C is required to place H_2O magnetization at equilibrium during the mixing period. The ZZ mixing time in Figure 3.5 was 484 ms (approximate optimal τ_{mix} for cross-peak buildup during the *Synechococcus* WT GlbN-A ESE reaction). Similar to the 2D ^{15}N ZZ experiments, an even number of equally spaced 180° G^3 ^{1}H selective 1.0-ms pulses is applied 3.3 kHz downfield of the carrier (4.76 ppm) to eliminate heteronuclear cross-correlated relaxation and minimize water saturation during τ_{mix} (τ_{mix} = n × (4δ + 2 p_H), δ = 5 ms, p_H ~ H^N selective 180° pulse). After the mixing period, $^{15}N_z$ magnetization is rotated into the transverse plane (point D) for chemical shift evolution (maximum t_2 = 49.4 ms). Gradient-encoded H_zN_+ magnetization is generated at point E. ^{15}N (ω_2) quadrature detection is achieved using echo-antiecho coherence selection. A sensitivity-enhanced reverse INEPT element was utilized to prepare ^{1}H magnetization for direct detection (τ_2 = 1 ms; τ_3 = 0.1 ms; G_3 ~ 0.5 × G_2). WALTZ16 ^{15}N decoupling was performed during FID acquisition (t_3 = 60 ms). The relaxation delay was set to 0.9 s and the total acquisition time was approximately 60 h.

One advantage of this exchange experiment compared to collecting spectra of the pure reduced (Figure S3.16) and oxidized forms individually is that the chemical shifts of each species are obtained under identical conditions of ionic strength, pH and temperature.

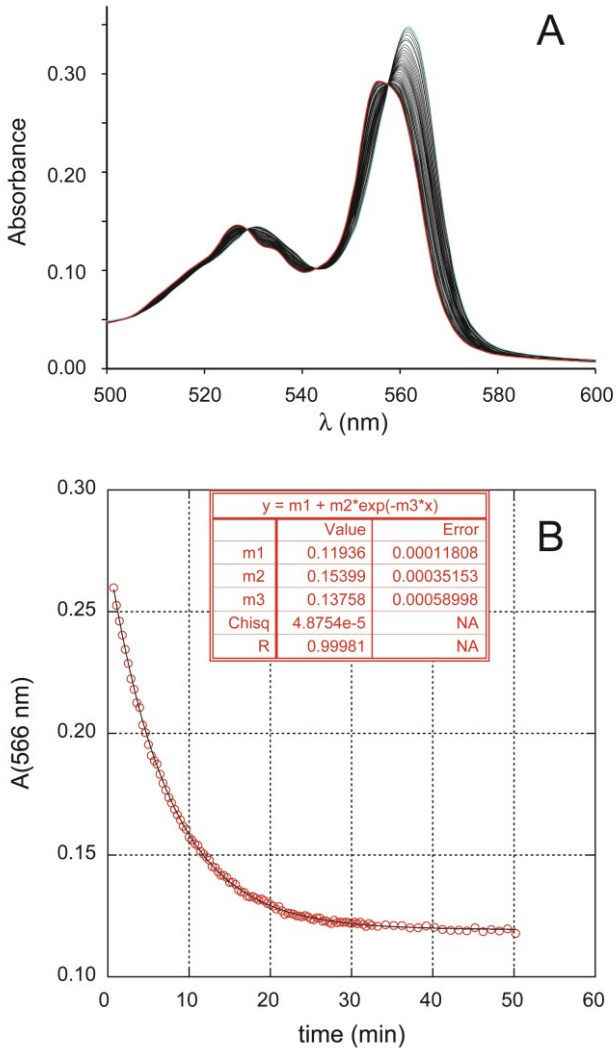


Figure S3.3. Conversion of *Synechococcus* WT GlbN-R to GlbN-A following reduction with excess DT at pH 9.2 monitored by optical absorbance. (A) The first spectrum is shown in cyan; the spectrum collected after 1 h is in red. Isosbestic points were detected for the first 30 to 45 min. After this period of time, spectral deterioration resembling that when treating the protein with H_2O_2 was observed. This was attributed to the release of peroxide following deactivation of catalase in the dioxygen scavenging system in the presence of excess DT. The data were collected over 20 h and processed by SVD in order to extract the rate constant for the conversion step. (B) Time course of the absorbance at 566 nm. The value of k_{xl} was $(2.3 \pm 0.1) \times 10^{-3}$ s^{-1}.

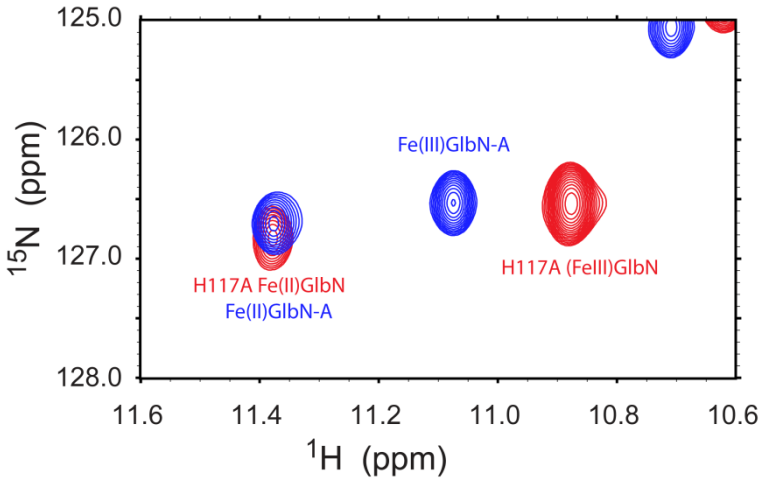


Figure S3.4. Overlay of two ^{1}H-^{15}N HSQC spectra of *Synechococcus* GlbN. The region contains signals from Thr80 HN. The two samples were mixtures of WT Fe(II) and Fe(III) GlbN-A (blue) and H117A Fe(II) and Fe(III) GlbN (red). The figure illustrates the overlap of the H117A Fe(II)GlbN and Fe(II)GlbN-A signals, which leads us to assume that the signals of WT Fe(II)GlbN-R and Fe(II)GlbN-A also overlap. Trace "a" in Figure 3.4a in the main text therefore contains intensity from both species. The *Synechocystis* GlbN conversion data were modeled assuming a similar overlap between the Asn80 HN signals in Fe(II)GlbN-R and Fe(II)GlbN-A (trace "d" in Figure 3.4b).

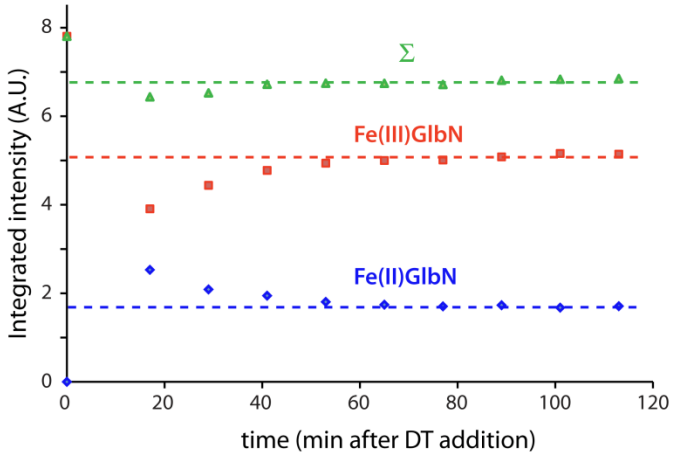


Figure S3.5. Time course of Thr80 HN signals for *Synechococcus* H117A Fe(II)GlbN (blue) and Fe(III)GlbN (red) after substoichiometric reduction in the presence of GODCAT. Integrated intensities were estimated by Gaussian simulation of cross peaks in 2-D ^{1}H-^{15}N HSQC data. The decrease in the intensity of the Fe(II) form is compensated by an increase in the intensity of the Fe(III) form. This is attributed to partial reoxidation of the sample, likely by dioxygen as a contaminant.

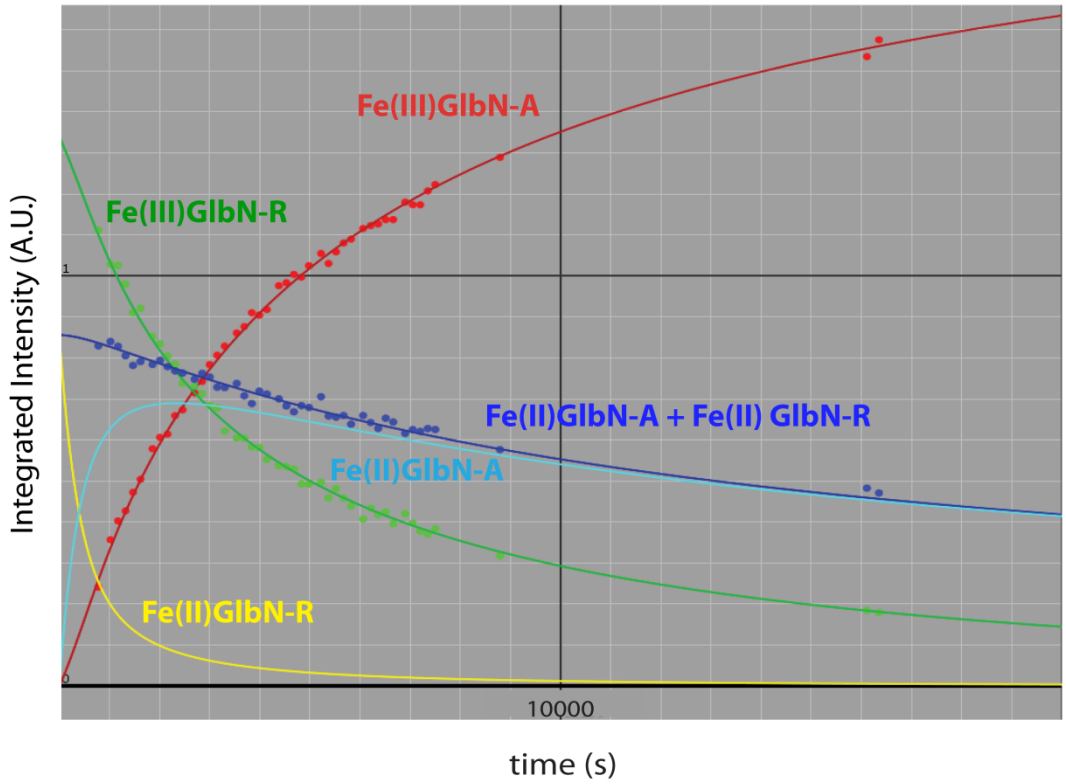


Figure S3.6. Time course of *Synechococcus* Fe(II)GlbN-R, Fe(III)GlbN-R, Fe(II)GlbN-A, and Fe(III)GlbN-A after substoichiometric reduction in the presence of GODCAT. The decrease in the intensity of the total Fe(II) protein is compensated by an increase in the intensity of the total Fe(III) protein. The data were simulated using the program KinTek Explorer and the lines represent a nonlinear least-squares fit of the data to the ET model (Equations 1, 2, and 3 in text above). The uncertainty of fitted parameters is discussed in the text and in Figure S3.7.

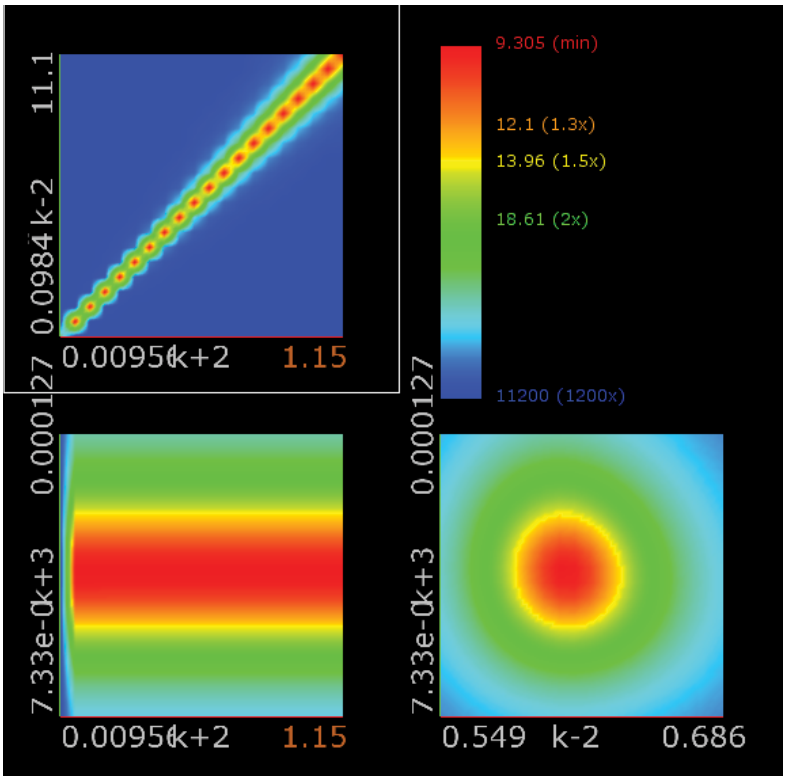


Figure S3.7. Confidence contour for the fit in Figure S3.6. Here, k_{+2} is the same as k_{12} in the text and k_{-2} corresponds to k_{21}; k_{+3} is the bimolecular rate constant for reoxidation. χ^2 Contours in yellow correspond to a $\chi^2_{limits}/\chi^2_{min} = 1.5$. This plot shows that the ratio of k_{12} and k_{21} (and not individual values) is well defined by the data.

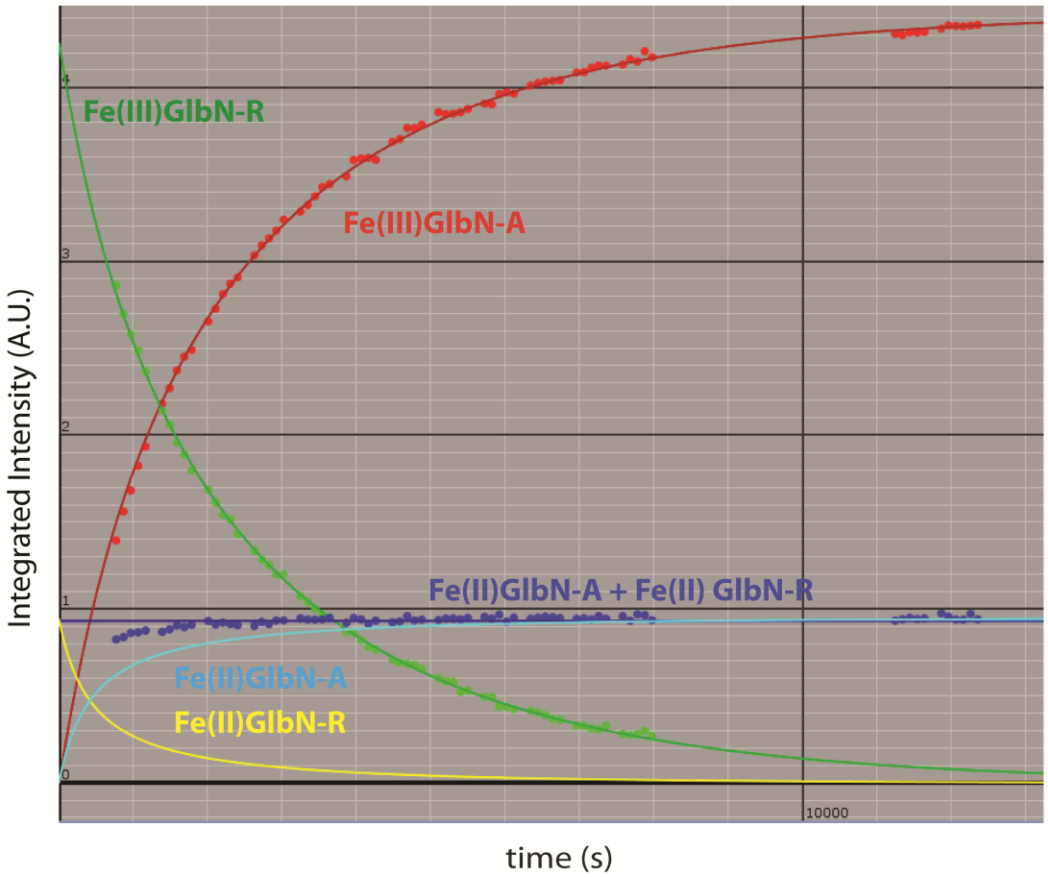


Figure S3.8. Time course of *Synechocystis* Fe(II)GlbN-R, Fe(III)GlbN-R, Fe(II)GlbN-A, and Fe(III)GlbN-A after substoichiometric reduction in the presence of GODCAT. The decrease in the intensity of the total ferrous protein is compensated by an increase in the intensity of the total ferric protein and the kinetics are consistent with a single ET step. *Synechocystis* Fe(II)GlbN remained at a constant concentration throughout the self-amplified conversion reaction. The data were simulated using the program KinTek Explorer and the lines represent a nonlinear least squares fit of the data directly to the ET model (Supplemental equations 1 and 2).

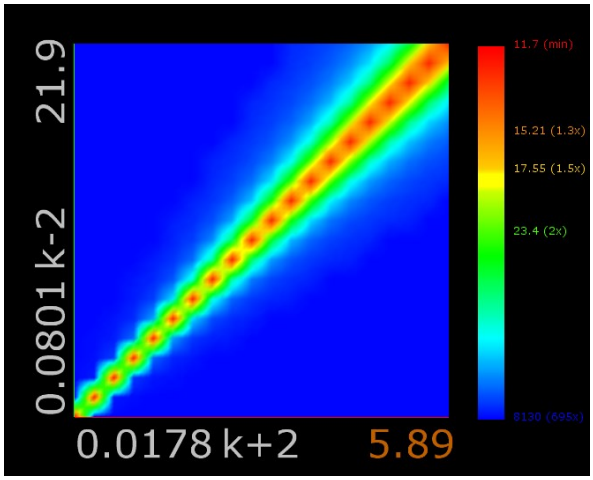


Figure S3.9. Confidence contour for the fit in Figure S3.8. Here, k_{+2} (in text, k_{12})corresponds to the *Synechocystis* WT Fe(II)GlbN-A to Fe(III)GlbN-R cross-exchange rate constant, and k_{-2} (in text, k_{21}) corresponds to the reverse reaction, in which Fe(II)GlbN-R transfers an electron back to Fe(III)GlbN-A. χ^2 Contours in yellow correspond to a $\chi^2_{limits}/\chi^2_{min} = 1.5$. The NMR data define the ratio of k_{12} and k_{21} but only provide a lower bound on $k_{12} \sim 10\ M^{-1}\ s^{-1}$. The apparent equilibrium constant observed for *Synechocystis* WT GlbN was $K_{12} = k_{12}/k_{21} = 0.25$.

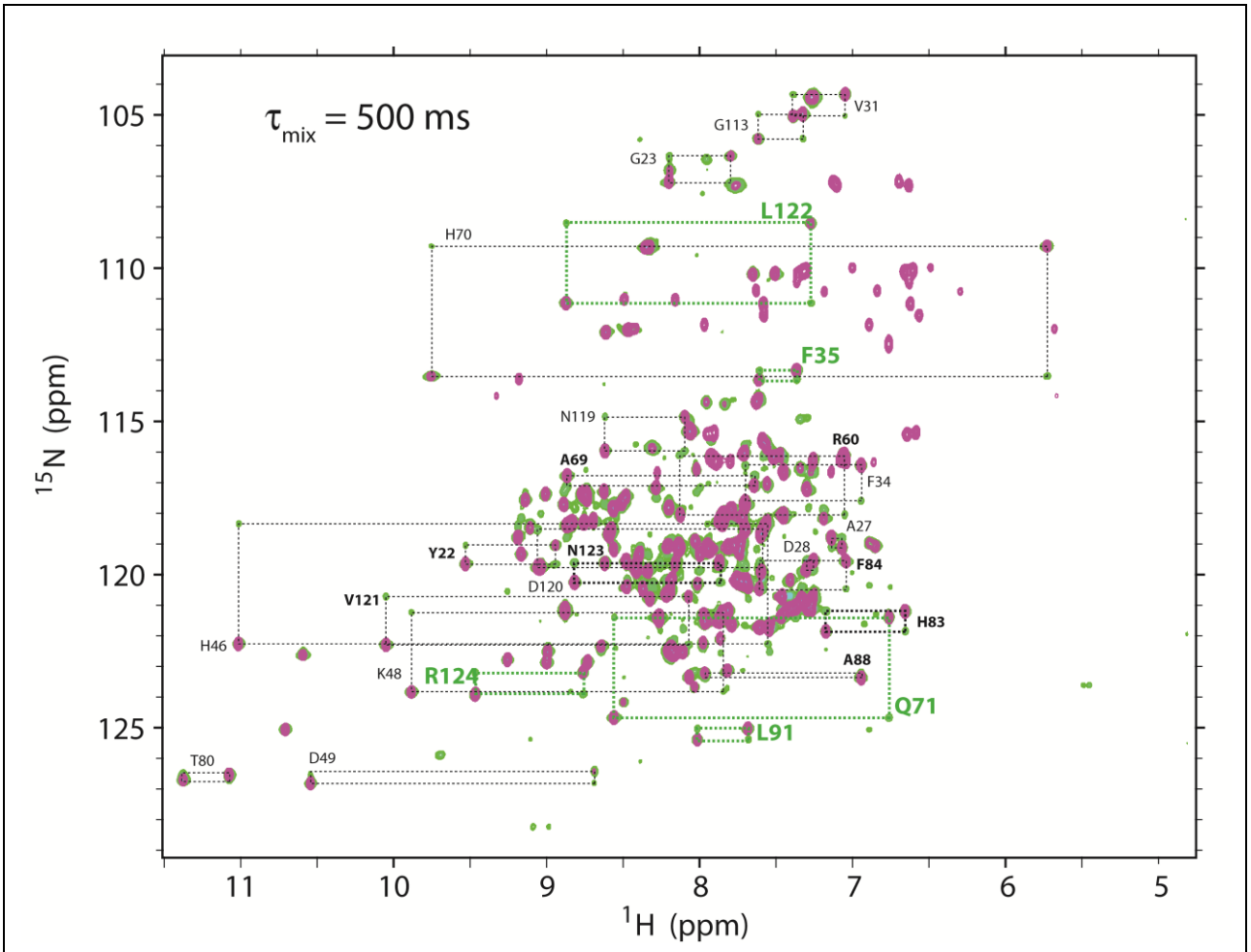


Figure S3.10. An ^{15}N ZZ exchange spectrum (green) overlaid onto a standard ^{1}H-^{15}N HSQC spectrum (magenta) collected on a 40:60 Fe(II):Fe(III) redox mixture of *Synechococcus* WT GlbN-A (inset of Figure 3.5 in the main text) showing the resolved exchange quartets chosen for ESE rate determination (annotated in green).

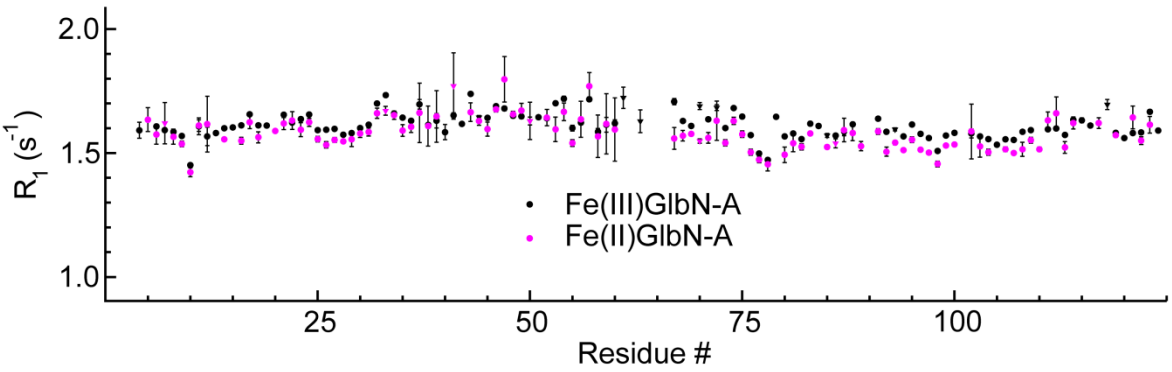


Figure S3.11. ^{15}N R_1 versus residue number for *Synechococcus* WT Fe(III)GlbN-A and Fe(II)GlbN-A. Note that the values for the reduced state are systematically lower than those of the oxidized state, but by a small percentage that in most cases is within experimental error. The values were consistent with those obtained in the fitting of the ^{15}N ZZ exchange experiments.

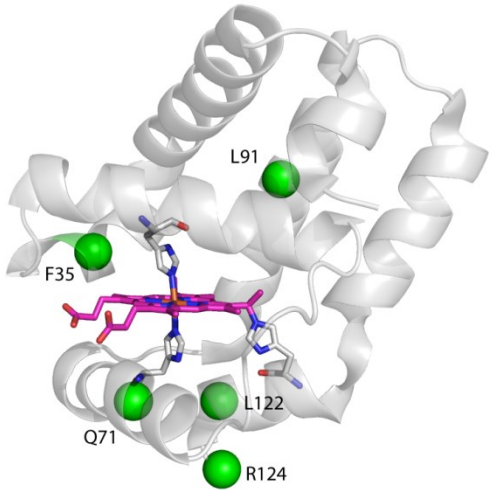


Figure S3.12. The five probes used in the analysis of the *Synechococcus* WT GlbN-A ^{15}N ZZ exchange data are shown as green spheres on the solution structure of *Synechococcus* WT GlbN-A (2KSC). As expected, the probes all report on the same exchange event corresponding to ESE.

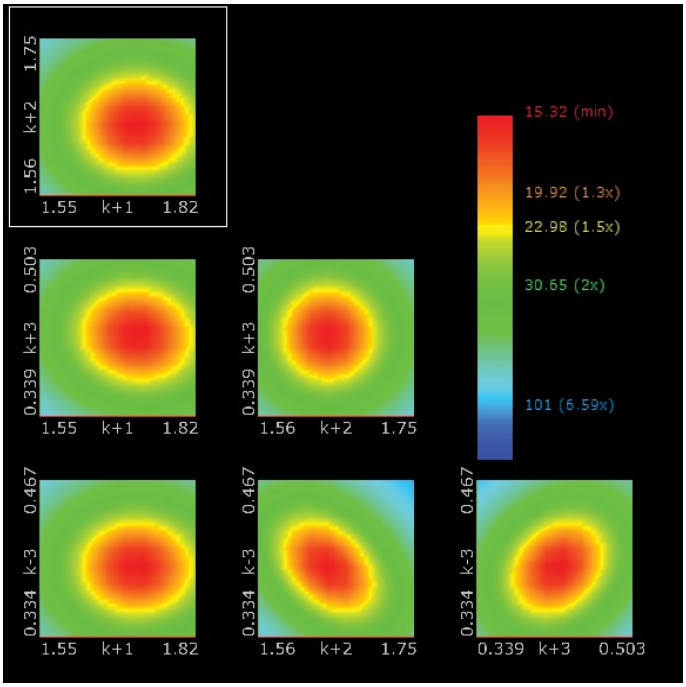


Figure S3.13. Confidence contour for the fit shown in Figure 3.5 of the main text (Gln71). Here, k_{+1} and k_{+2} are the ^{15}N R_1 values in *Synechococcus* WT Fe(II) and Fe(III) GlbN-A respectively. k_{+3} and k_{-3} are the apparent 1st order rate constants that characterize N_z magnetization exchange during τ_{mix} due to ESE. k_{+3} (k_{-3}) reports on indirectly detected ferrous (ferric) magnetization that was subsequently transferred to ferric (ferrous) GlbN-A for direct detection. The yellow regions of the χ^2 contour plots correspond to a $\chi^2_{limits}/\chi^2_{min} = 1.5$.

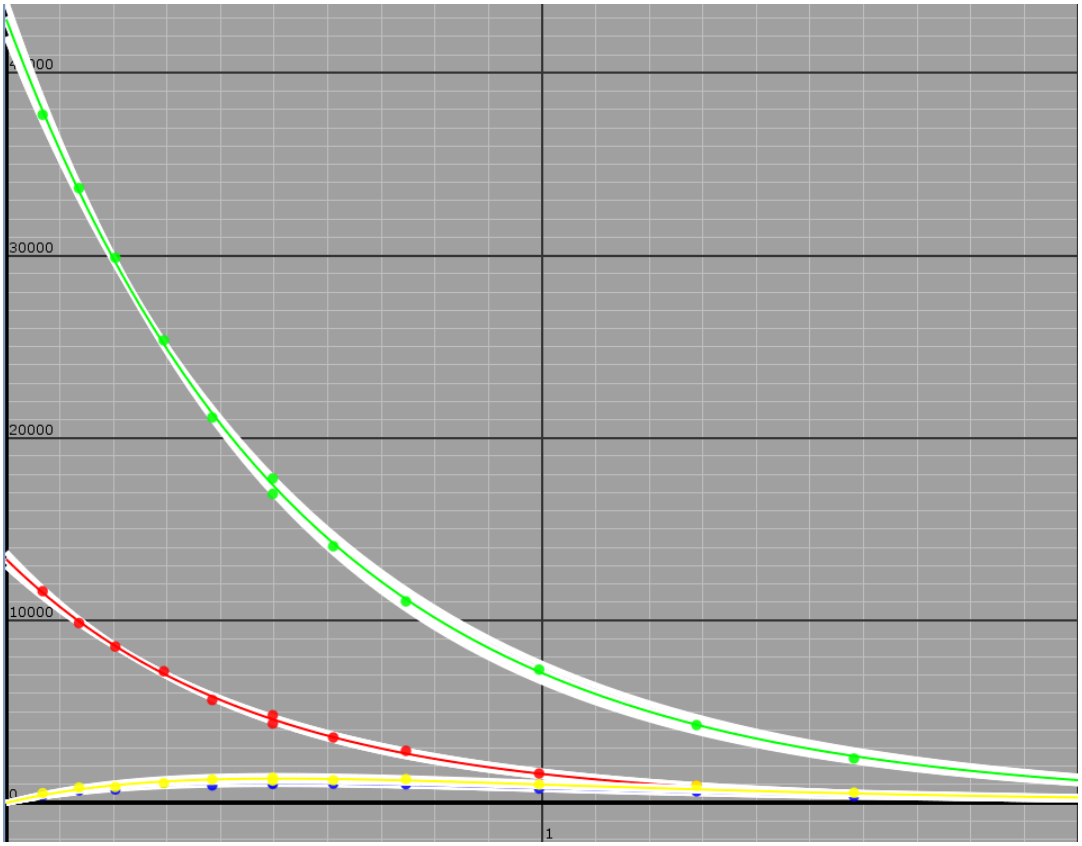


Figure S3.14. Sample fit for Nz ZZ exchange detected ESE in *Synechococcus* H117A GlbN (Q47 HN). The white regions define the lower and upper bounds determined from KinTek Explorer confidence contour analysis ($\chi^2_{limits}/\chi^2_{min}$ = 1.5).

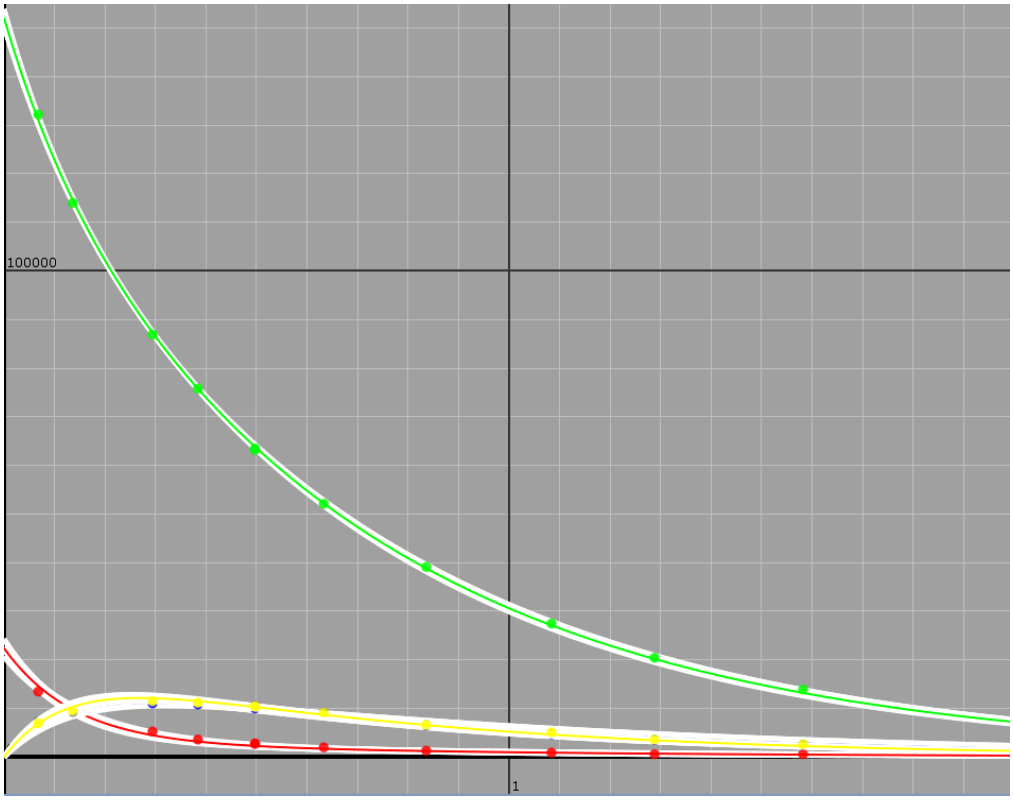


Figure S3.15. Sample fit to determine the ESE kinetics of *Synechocystis* WT GlbN-A using Nz ZZ exchange NMR. The Fe(II) GlbN-A ^{15}N R_1 uncertainty is relatively large owing to low auto-peak intensities (red curve). The ESE rate constants were determined with reasonable accuracy; the white traces describe the upper and lower error bounds for $\chi^2_{limits}/\chi^2_{min} = 1.5$.

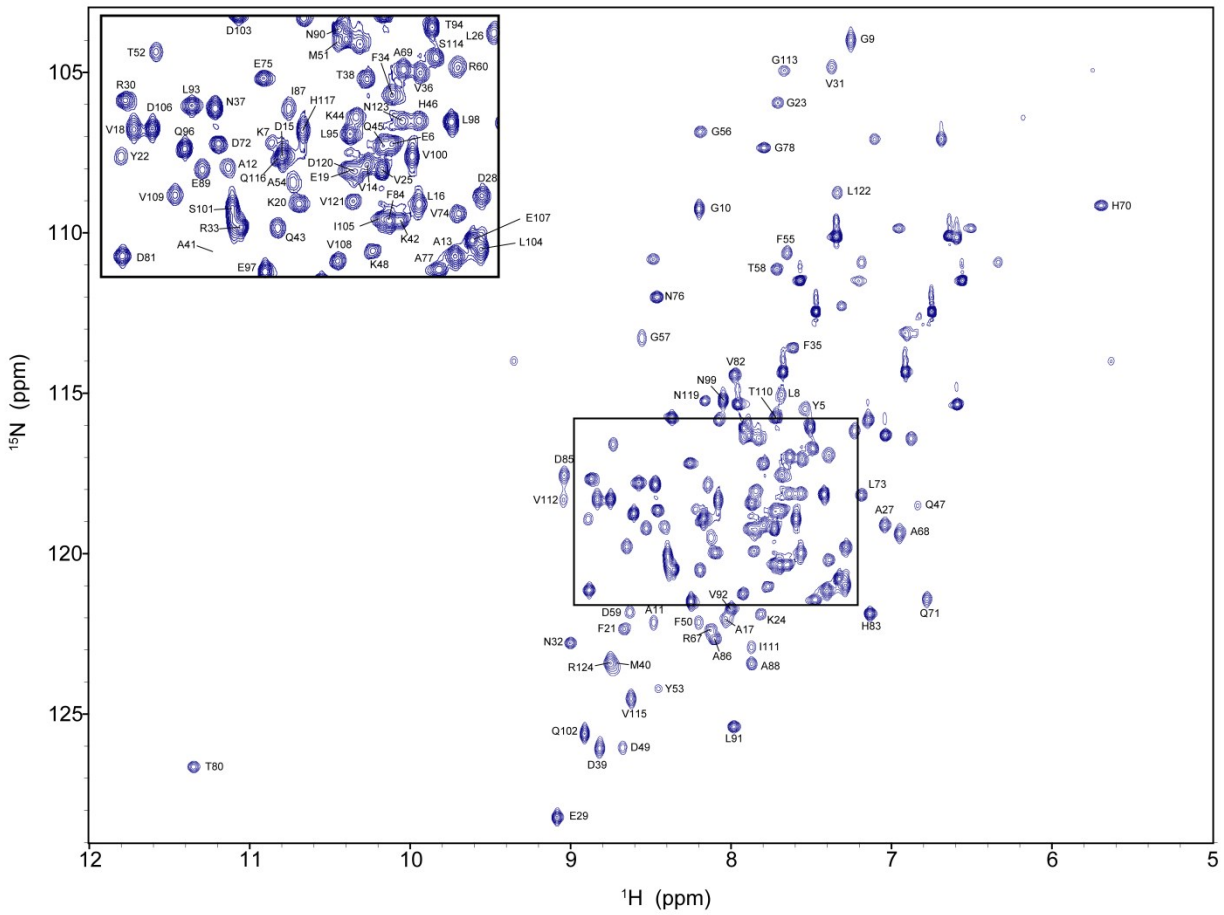


Figure S3.16. An assigned ^{1}H-^{15}N HSQC spectrum of *Synechococcus* WT Fe(II)GlbN-A (600 MHz, 298 K, 100 mM phosphate buffer, pH 7.2, 10% D_2O). The sample contained ~1.6 mM GlbN and GODCAT. Note that the line widths are variable. For example, broader cross peaks are observed at the end of the E helix (residues 47–54). Backbone ^{1}H chemical shifts of Fe(II)GlbN-A used in the TALOS+ and pseudocontact shift calculations have been deposited at the BMRB (accession number 17947).

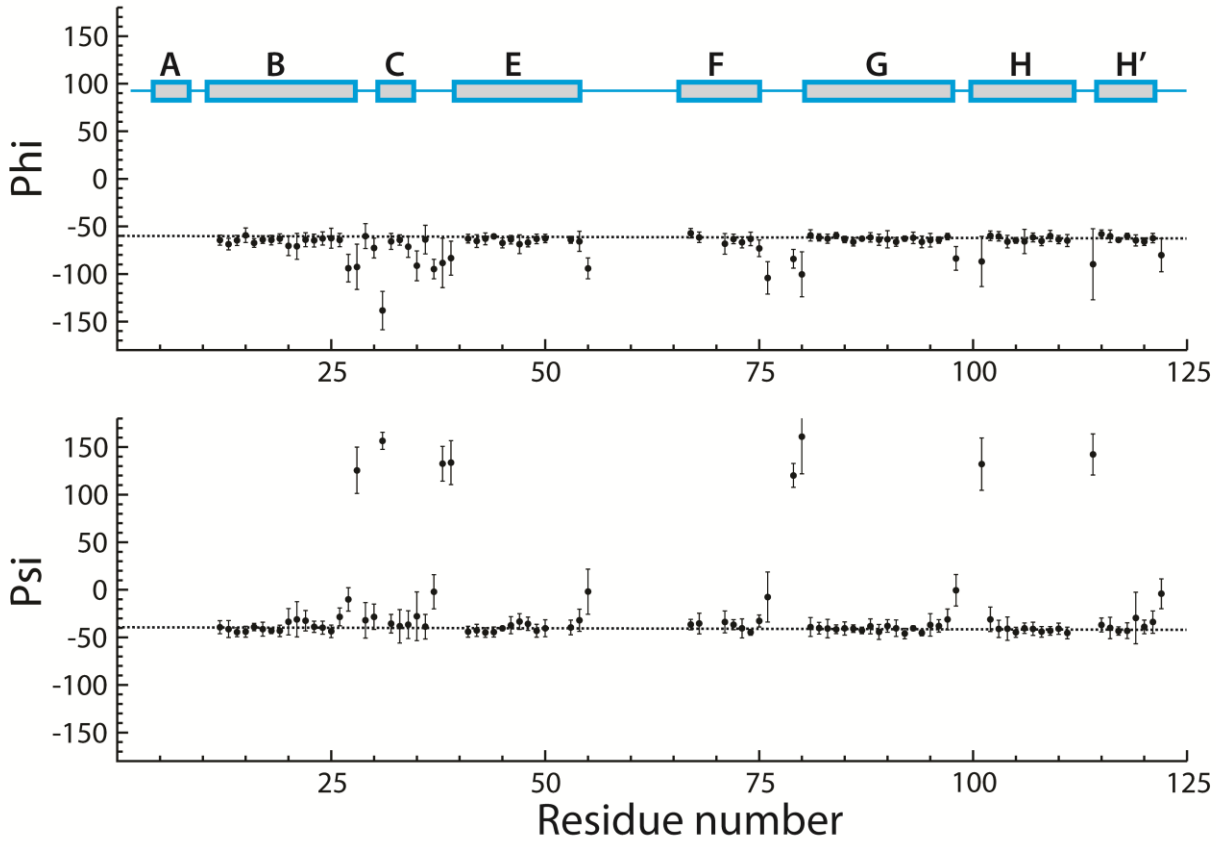


Figure S3.17. Backbone dihedral angles and their uncertainties for *Synechococcus* WT Fe(II)GlbN-A calculated with TALOS+. The secondary structure observed in Fe(III)GlbN-A, indicated by rectangles, is maintained. Dotted lines are drawn at φ = –60° and ψ = –40°.

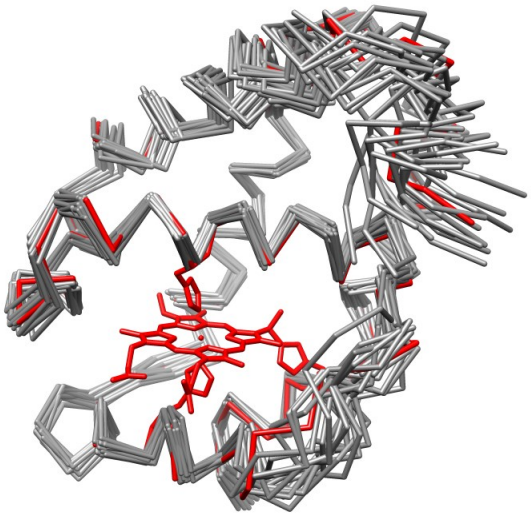


Figure S3.18. The structure refined with the pseudocontact shifts is shown in red among the 2KSC family. It is comprised within it, which indicates similar sampling of the conformational space with XPLOR. The regions of largest deviations between the refined model and the 2KSC average are those that display the largest ensemble rmsds without refinement (Figure 3.1a). We conclude that the measured pseudocontact shifts can be accommodated by the current family and that the backbone structures of the reduced and oxidized *Synechococcus* WT GlbN-A are, on average, similar.

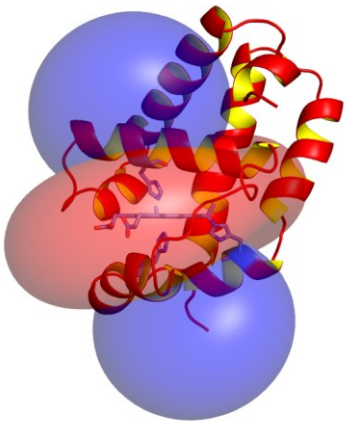


Figure S3.19. Pseudocontact shift map for *Synechococcus* WT GlbN-A in the ferric bishistidine state. Representation with Pymol (Delano Scientific) of the Numbat output [10]. The orientation of the structure is similar to that in Figure 3.1a of the text, with proximal His below the heme. The isoshift surfaces are drawn at +0.13 ppm (blue) and –0.13 ppm (red).

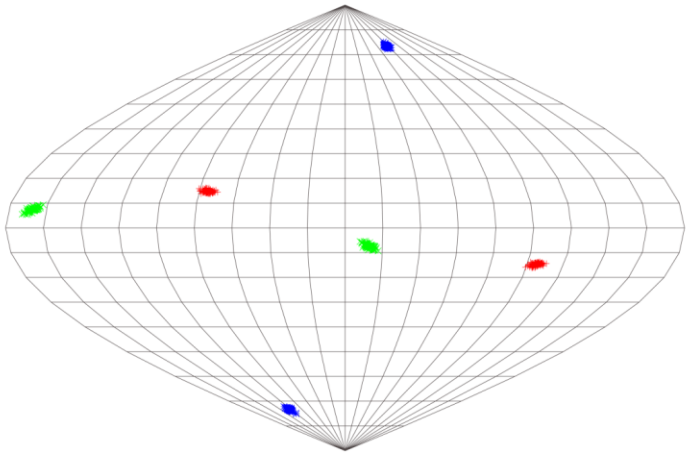


Figure S3.20. Sanson-Flamsteed equivalent projection of the magnetic susceptibility tensor for *Synechococcus* WT GlbN-A in the ferric bishistidine state. The z axis (vertical) is normal to the heme plane. The figure illustrates the tilt and the well-defined orientation of the tensor. Generated with Numbat [10].

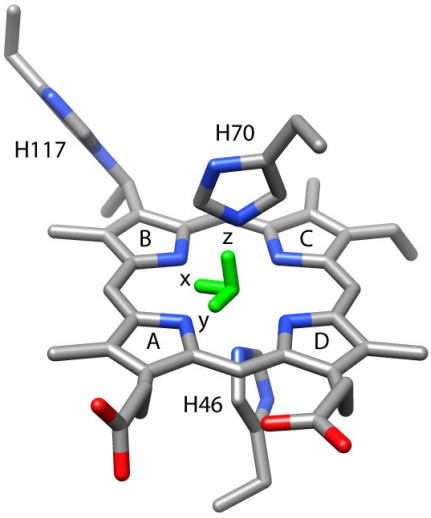


Figure S3.21. Magnetic susceptibility tensor of *Synechococcus* WT Fe(III)GlbN-A after refinement of 2KSC model 2. The tensor axes are represented as green sticks. The z direction lies practically in the plane normal to the heme passing through the nitrogens of pyrrole B and D.

Chapter 4.

The histidine-heme posttranslational modification in *Synechococcus* hemoglobin conditions its reactivity toward nitric oxide

Matthew R. Preimesberger, Lukas Gilevicius, and Juliette T.J. Lecomte

Author contributions:

MRP: designed research, prepared protein samples, collected and analyzed $NO^{\bullet}$, O_2, NOD (single addition) UV-visible absorbance data, acquired and analyzed NMR data, assisted with NOD and Griess assay development, wrote the chapter

LG: assisted with protein preparation, conducted NOD (multiple addition) and Griess assays, analyzed UV-visible data

JTJL: designed research, analyzed data

Abbreviations: O_2, oxygen; $O_2^{\bullet-}$, superoxide; $NO^{\bullet}$, nitric oxide; CO, carbon monoxide; H_2O_2, hydrogen peroxide; NO_2^-, nitrite; NO_3^-, nitrate; NOD, nitric oxide dioxygenase; NOESY, nuclear Overhauser effect spectroscopy; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy

Abstract

Synechococcus sp. PCC 7002 is a robust model cyanobacterium capable of growing under conditions that promote reactive oxygen species (ROS) and reactive nitrogen species (RNS) stress.[1,2] The organism contains an 2/2 hemoglobin (GlbN) implicated in ROS/RNS detoxification caused when cells are grown with nitrate as their only source of nitrogen.[12] In vitro structural studies on the recombinant protein have indicated that GlbN is a member of a growing class of hemoglobins that utilize two histidines (H70, "proximal"; and H46 "distal") to coordinate the heme iron. Relative to the classic scheme exhibited by hemoglobin and myoglobin, *bis*-histidine iron coordination is thought to enhance electron transfer kinetics, lower reduction potential, stabilize the ferric form the protein, and condition ligand binding properties.

GlbN is distinguishable among virtually all other hemoglobins. Under reducing conditions in the absence of oxygen, H117 undergoes an irreversible spontaneous posttranslational modification (PTM) with the heme 2-vinyl group to form a His117 Nε2-2-Cα covalent bond. The PTM causes substantial stabilization of the holoprotein, with only minor structural and dynamic perturbation. Interestingly, under oxic conditions, unmodified GlbN is primarily obtained when overexpressed in synechococcal cells; however under microoxic conditions, GlbN with PTM (GlbN-A) is predominantly detected. [12] Thus, O_2 may inhibit the spontaneous PTM; additionally the data support that both GlbN and GlbN-A are physiologically active in the cell (depending on the external growth conditions). Despite much effort, no clear chemical role for the His-heme PTM has been determined. In this work, we examined the influence

of His-heme crosslinking by comparing GlbN and GlbN-A and their reactions with nitric oxide ($NO^{\bullet}$).

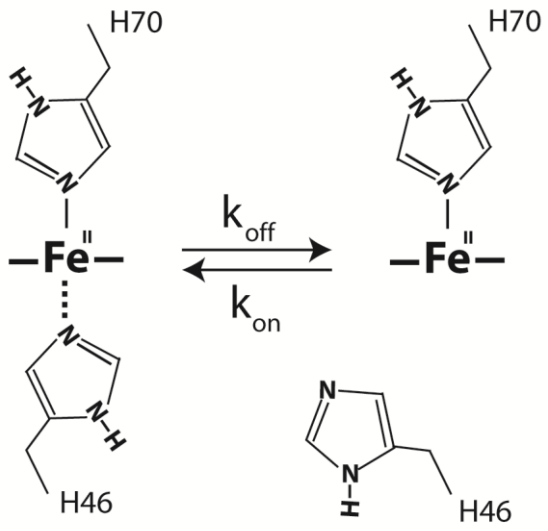
The ability for hemoglobins to both destroy (via $NO^{\bullet}$dioxygenase, NOD) and produce $NO^{\bullet}$ (via nitrite reductase) depends on the presence (NOD) or absence (nitrite reductase) of dioxygen. Electron absorption spectrophotometry was utilized to characterize the ability for GlbN and GlbN-A to undergo NOD chemistry. Both GlbN and GlbN-A are capable of catalyzing multiple NOD cycles under oxic conditions using a Ferredoxin-NADP+ reductase system for heme re-reduction. Upon multiple $NO^{\bullet}$ challenges, heme bleaching was observed for both proteins which casts some doubt on an in vivo NOD function for GlbNs.

Finally, NMR spectroscopy complemented UV-visible absorbance data to study the binding of $NO^{\bullet}$ to both Fe(III) and Fe(II) GlbNs. Both GlbN and GlbN-A were found to bind $NO^{\bullet}$ in the ferric state (Fe(III)-$NO^{\bullet}$) forming diamagnetic complexes which resisted autoreduction to the paramagnetic Fe(II)-$NO^{\bullet}$ forms. Dithionite reduction of Fe(III)-$NO^{\bullet}$GlbN and GlbN-A resulted in marked spectroscopic differences. Whereas GlbN-A formed a stable Fe(II)-$NO^{\bullet}$ complex, $NO^{\bullet}$ binding to Fe(II) GlbN caused immediate heme loss and remarkably, HNO formation. A nitrite reduction chase experiment using dithionite, Fe(III) GlbN and $^{15}NO_2^-$ was used to generate $^{15}NO^{\bullet}$ in situ and resulted in formation of Fe(II)-H^{15}NO GlbN. Although GlbN-A also appears capable of trapping HNO, formation of the His-heme PTM extinguishes the unusual $NO^{\bullet}$ reductase chemistry associated with the non crosslinked state (GlbN). Overall, the results suggest a potential role for the reductive covalent modification in Fe(II) GlbNs:

protection from $NO^{\bullet}$-mediated heme loss and prevention of undesired HNO formation.

Introduction

The model cyanobacterium *Synechococcus* sp. PCC 7002 contains a gene (*glbN*) encoding a monomeric hemoglobin (GlbN) of the group 1 truncated lineage (trHb1).[3] Structural and chemical reactivity studies indicate that GlbN binds the heme using two histidines.[3,4] H70 (F8 helical position in the Perutz nomenclature) is the strictly conserved "proximal" heme ligand and is tightly coordinated to the iron via its Nε2. The Nε2 of H46 (E10 helical position) reversibly coordinates the "distal" iron site (Scheme 4.1).



Scheme 4.1. The reversible *bis*-histidine iron coordination scheme of synechococcal GlbNs. Upon binding of exogenous ligands such as O_2 or CO, the proximal histidine (H70) remains coordinated to the heme iron while the second histidine (H46) is displaced. In the *bis*-histidine state, the octahedral iron coordination geometry results in a low spin diamagnetic (S=0) ferrous complex (left). In the decoordinated state (right), the ferrous iron adopts a square pyramidal high spin (S=2) complex. The rate constants k_{off} and k_{on} and equilibrium $K=k_{on}/k_{off}$ tune heme redox and ligand binding properties.

Along with *Synechocystis* sp. PCC 6803 GlbN,[5] ferrous *Synechococcus* GlbN is extraordinary among hemoglobins for its ability to undergo a spontaneous and irreversible posttranslational modification (PTM) involving a non-coordinating histidine (H117) to form a heme 2-Cα-Nε2 H117 covalent bond. [3,6,7] NMR and optical absorbance studies support that the GlbN His-heme PTM is mechanistically and structurally analogous to Cys-heme thioether bond formation observed in certain *c*-type cytochromes in vitro.[7–11] Much work has sought to understand the physicochemical consequences of PTM in GlbN. Although the structures of GlbN and the modified protein (hereafter GlbN-A) are

highly similar,[4,12] the thermal stability of GlbN-A is substantially enhanced relative to the non-crosslinked protein.[13] Only subtle dynamic modifications are observed in GlbN-A relative to GlbN.[13,14]

Low levels of *glbN* transcripts are detected constitutively in synechococcal cells, under a wide variety of conditions.[12] GlbN overexpressed from synechococcal cells under microoxic conditions contains covalently attached heme. However, when expressed under oxic conditions, GlbN contains primarily unmodified *b* heme.[12] Thus, O_2 appears to inhibit the spontaneous PTM by binding to Fe(II) GlbN and presumably withdrawing electron density from the porphyrin π-system of ferrous heme.[15] This effect, which assumes the formation of a complex with Fe(III)-$O_2^{\bullet-}$ character, is similar to the π-acceptor back-bonding effect of CO.[7]

The presence or absence of a functional *glbN* gene alters the ability of *Synechococcus* cells to grow on nitrate (NO_3^-) as their sole nitrogen source. Interestingly, elevated levels of ROS and reactive nitrogen species (RNS) are detected in a *Synechococcus glbN* knock-out strain under both aerobic and microoxic NO_3^- growth.[12] Additionally, the same *Synechococcus glbN* knock-out strain complemented with a plasmid overexpressing GlbN alleviates ROS/RNS levels and restores cell growth partially. In light of these observations and low *glbN* expression levels, simple O_2 transport or storage are unlikely roles for GlbN. Instead, GlbNs is thought to perform an enzymatic or sensor function.

The effect of NO_3^- on the knock-out strain suggests that GlbN participates in nitrogen-oxygen chemistry. There are numerous possibilities for such activity. Assimilatory NO_3^- reduction to ammonia (NH_3) requires a total of 8 e^- and

proceeds through the production of nitrite (NO_2^-). The enzymes involved in the reduction are known to leak $NO^\bullet$,[12,16] a short-lived, reactive radical in the cellular millieu. $NO^\bullet$ readily combines with another common cyanobacterial ROS, $O_2^{\bullet-}$, to form the powerful oxidant peroxynitrite ($O{=}NOO^-$).[17–20] This lead to the hypothesis that GlbN tempers NO_3^-–mediated ROS/RNS cellular stress in *Synechococcus* cells via elimination of hydrogen peroxide (H_2O_2), $O_2^{\bullet-}$, $NO^\bullet$[12,21] or peroxynitrite[22,23] (Ferric reaction 4.1).

Ferric reaction 4.1. Peroxynitrite isomerase

$$\text{Hb-Fe(III)} + \text{ONOOH} \leftrightarrows [\text{Hb-Fe(III)–ONOO}^-] + \text{H}^+ \quad (1)$$

$$[\text{Hb-Fe(III)–ONOO}^-] + \text{H}^+ \leftrightarrows [\text{Hb-Fe(III)}] + {}^\bullet\text{ONO} + \text{OH}^\bullet \quad (2)$$

$$[\text{Hb-Fe(III)}] + \text{ONO}^\bullet + \text{OH}^\bullet \rightarrow \text{Hb-Fe(III)} + \text{NO}_3^- \quad (3)$$

Bis-histidine hexacoordination occludes the distal iron site in ferrous GlbNs. Among other consequences, the result is a likely attenuation of highly undesirable Fenton chemistry (Ferrous reaction 4.1).[24]

Ferrous reaction 4.1. Fenton chemistry

$$\text{Hb-Fe(II)} + \text{H}_2\text{O}_2 \leftrightarrows \text{Hb-Fe(III)} + \text{OH}^- + \text{OH}^\bullet$$

Similarly, both GlbN and GlbN-A exhibit stable *bis*-histidine coordination to the ferric iron; this feature imparts protection from H_2O_2-mediated heme bleaching[7] by blocking peroxidase chemistry (Ferric reaction 4.2.)

Ferric reaction 4.2. Hydrogen peroxidase

$$\text{Hb-Fe(III)} + \text{H}_2\text{O}_2 \leftrightarrows [\text{Hb-Fe(III)–OOH}^-] + \text{H}^+ \quad (1)$$

$$[\text{Hb-Fe(III)–OOH}^{-}] + \text{H}^{+} \rightarrow [\text{Hb-Fe(IV)=O}^{+\bullet}] + \text{H}_2\text{O} \qquad (2)$$

$$[\text{Hb-Fe(IV)=O}^{+\bullet}] + \text{guiacol} + 2\text{H}^{+} \rightarrow \text{Hb-Fe(III)} + \text{H}_2\text{O} + \text{guiacol}^{2+} \qquad (3)$$

Unlike true heme peroxidases[25–27] or the proposed pseudo-peroxidase trHbs from *Arabidopsis thaliana* (k_{cat}/K_M ~ 10^3 M^{-1} s^{-1}),[28] *Synechococcus* GlbN and GlbN-A do not form a stable Fe(IV)=$O^{\bullet+}$ (Compound I) oxidant and accordingly display poor in vitro peroxidase activity (k_{cat}/K_M ~ 1 M^{-1} s^{-1}) when guaiacol is an external substrate.[12] Thus, it is doubtful that *Synechococcus* GlbNs play a direct role in H_2O_2 detoxification.

$NO^{\bullet}$ is a relatively stable radical in aerated aqueous solution ($t_{1/2}$ ~ 4-7 min)[29,30]; under in vivo conditions direct interaction with metal centers can lead to its rapid consumption ($t_{1/2}$ ~ 2 ms).[30] $NO^{\bullet}$ is a peculiar strong-field ligand which can bind both ferrous and ferric heme iron tightly.[16,17] Additionally, $NO^{\bullet}$ exerts a strong negative σ-*trans* effect, which weakens axial iron ligation on the opposite side of the heme.[31,32] Hemoglobins can both destroy and produce $NO^{\bullet}$ depending on the iron coordination and redox state. The most commonly accepted route for $NO^{\bullet}$ destruction is through its reaction with oxy-hemoglobin via the so-called $NO^{\bullet}$ dioxygenase (NOD, Oxy reaction 4.1) reaction exemplified by the dedicated NOD *E. coli* flavohemoglobin.[33–36]

Oxy reaction 4.1. Nitric oxide dioxygenase (NOD)

$$\text{Hb-Fe(II)} + O_2 \leftrightarrows \text{Hb-Fe(III)}{-}O_2^{\bullet -} \quad (1)$$

$$\text{Hb-Fe(III)}{-}O_2^{-} + NO^{\bullet} \rightarrow [\text{Hb-Fe(III)}{-}OONO^{-}] \quad (2)$$

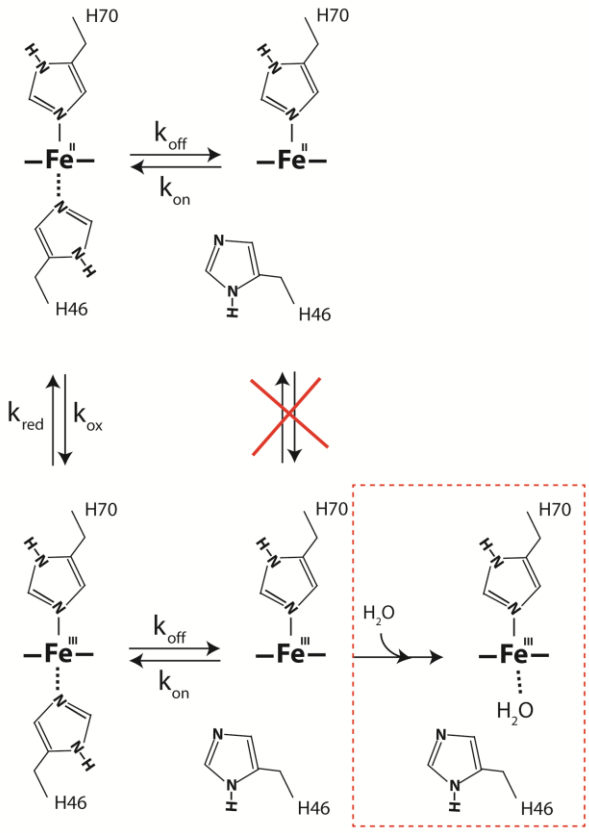
$$[\text{Hb-Fe(III)}{-}OONO^{-}] \rightarrow \text{Hb-Fe(III)} + NO_3^{-} \quad (3)$$

$$\text{Hb-Fe(III)} + e^{-} \rightarrow \text{Hb-Fe(II)} \quad (4)$$

Other trHbs from strains of *Mycobacterium*, *Tetrahymena pyriformis* (a ciliate), and microalgal raphidophytes (*Heterosigma akashiwo and Chattonella subsalsa*) are proposed to undergo similar chemistry.[37,21,38–40]

Hargrove and coworkers emphasize that many hemoglobins can undergo NOD chemistry in vitro, but are often limited by external factors in vivo.[46] To wit, the NOD reaction (Oxy reaction 4.1) consists of four steps: 1) Dioxygen (O_2) binding by Fe(II) GlbN, 2) radical combination of $NO^{\bullet}$ with activated Fe(III)-$O_2^{\bullet -}$, 3) peroxynitrite isomerization and rapid dissociation of product NO_3^{-} from Fe(III) GlbN, and 4) re-reduction of Fe(III) GlbN to the initial Fe(II) state. Although most hemoglobins readily undergo the radical combination reaction, re-reduction is often slow and rate-limiting.[46] The classical NOD flavohemoglobin contains an additional reductase domain for rapid *intra*molecular electron transfer (ET).[47,48] In contrast, GlbNs require *inter*molecular ET for NOD turnover. Notably, the *bis*-histidine hexacoordination observed in GlbN and GlbN-A is known to facilitate ET and electron self-exchange (Scheme 4.2).[11,49] Thus, an attractive hypothesis is that under oxic conditions and with the

participation of a reductase, GlbNs protect the NO_3^-–grown *Synechococcus* cells from $NO^{\bullet}$-mediated stress via efficient NOD chemistry.



Scheme 4.2. Model for electron transfer in *bis*-histidine GlbNs. With the distal histidine (H46) in the decoordinated state (right), a distinct ligand field geometry is preferred for the ferrous and ferric iron (if sterically possible, water will coordinate ferric heme iron in part owing to its formal +1 charge, the red dashed box is a hypothetical state in GlbNs, but is supported by water coordination in the H46A variant of *Synechocystis* GlbN). According to Marcus theory, electron transfer is retarded along the right-hand pathway because of the substantial nuclear reorganization required to fulfill the Franck-Condon principle. On the contrary, in the *bis*-histidine scheme (left), the ferric and

ferrous iron *both* adopt a low spin octahedral geometry with minimal structural alteration. Thus, the rates of reduction (k_{red}) and oxidation (k_{ox}) will be enhanced and rapid redox cycling made possible. Natural electron transfer heme proteins such as cytochrome b_5 (His-Fe-His) and cytochromes *c* (His-Fe-Met) utilize hexacoordinate heme ligation to achieve electron transfer rate constants as high as 10^5 - 10^7 M^{-1} s^{-1}.[50,51]

It is not obvious how GlbNs confer protection to *Synechococcus* cells grown on high levels of NO_3^- under microoxic conditions.[12] It is possible that ferric GlbNs perform the peroxynitrite isomerase reaction (Ferric reaction 4.1), which does not require redox cycling.[22,23] Also possible is the nitrosyl autoreduction reaction (Ferric reaction 4.3).

Ferric reaction 4.3. Nitrosyl autoreduction

$$\text{Hb-Fe(III)} + NO^{\bullet} \leftrightarrows \text{Hb-Fe(III)–}NO^{\bullet} \quad (1)$$

$$\text{Hb-Fe(III)–}NO^{\bullet} \leftrightarrows [\text{Hb-Fe(II)–}NO^{+}] \quad (2)$$

$$[\text{Hb-Fe(II)–}NO^{+}] + OH^- \leftrightarrows \text{Hb-Fe(II)} + NO_2^- + H^+ \quad (3)$$

$$\text{Hb-Fe(II)} + NO^{\bullet} \rightarrow \text{Hb-Fe(II)–}NO^{\bullet} \quad (4)$$

Following $NO^{\bullet}$ binding and activation to form the Fe(II)-NO^+ state, reduction is catalyzed by NO_2^- or hydroxide anion (Ferric reaction 4.3, step 3)[52,53] and eliminates $NO^{\bullet}$.[54] To undergo an additional cycle, Fe(II) GlbN must be reoxidized to the initial Fe(III) state.

Finally, under microoxic conditions hemoglobins such as the Fe(II) GlbN can produce $NO^{\bullet}$ via well-known NO_2^- reductase chemistry (Ferrous reaction 4.2).[41-45,55,56]

Ferrous reaction 4.2. Nitrite reductase

$$\text{Hb-Fe(II)} + NO_2^- \leftrightarrows [\text{Hb-Fe(II)}–NO_2^-] \quad (1)$$

$$[\text{Hb-Fe(II)}–NO_2^-] + H^+ \leftrightarrows \text{Hb-Fe(III)} + NO^\bullet + OH^- \quad (2)$$

$$\text{Hb-Fe(III)} + e^- \rightarrow \text{Hb-Fe(II)} \quad (3)$$

$$\text{Hb-Fe(II)} + NO^\bullet \rightarrow \text{Hb-Fe(II)}–NO^\bullet \quad (4)$$

Plant and cyanobacterial Hbs are thought to be especially adept at the nitrite reductase chemistry.[43,45] The NO_2^- reductase reaction is essentially the reverse of the previously described $NO^\bullet$ autoreduction reaction (Ferric reaction 4.3). Specifically, NO_2^- reductase chemistry (Ferrous reaction 4.2) is initiated by NO_2^- binding to the Fe(II) center (step 1). Acid-catalyzed reduction to $NO^\bullet$ is followed by $NO^\bullet$ release from the ferric iron (step 2). Re-reduction of Fe(III) GlbN to Fe(II) GlbN is required (step 3) to regenerate the starting enzyme. Interestingly, the $NO^\bullet$ generated in step 2 can inhibit subsequent cycles by competing with NO_2^- for the ferrous iron (step 4).[43] An unusual distinct mechanism of $NO^\bullet$ production has been proposed in the heme protein nitrophorin 7, and involves only the ferric heme and a NO_2^- disproportionation reaction ($3NO_2^- + 2H^+ \rightarrow 2NO^\bullet + NO_3^- + H_2O$).[57]

Clearly, GlbN and GlbN-A could catalyze a large number of reactions making difficult an unambiguous assignment of these protein's role in *Synechococcus* ROS/RNS protection. The aim of the current chapter is to address potential differences in the in vitro redox reactivity of these two forms of GlbN towards the common cyanobacterial RNS, $NO^\bullet$. The ability of GlbN and GlbN-A to bind oxygen and destroy (NOD) or produce (NO_2^- reductase) $NO^\bullet$ is compared.

Together, the previous physiological[12] and current chemical study allows for a tentative suggestion for the physiological role of His-heme PTM in *Synechococcus* GlbN.

Materials and Methods

Protein preparation

Wild-type and H117A *Synechococcus* sp. PCC 7002 GlbNs were produced recombinantly using *E. coli* BL21(DE3) as described previously.[3] GlbNs primarily partition into inclusion bodies,which were re-solubilized in freshly prepared 8 M urea and applied to a G-50 size exclusion column for apoprotein re-folding and purification. Fractions deemed pure by denaturing SDS-PAGE were pooled and the concentration of the apoprotein evaluated with a predicted ε_{280} of 4.47 mM^{-1} cm^{-1} prior to coarse hemin titration to generate the ferric holoprotein. Passage of the resulting solution through a DEAE anion exchange column produced purified Fe(III) GlbN (> 95 %). The purified protein samples were concentrated and buffer exchanged into pH-adjusted H_2O or 1 mM Na/K-phosphate, pH 7.0-7.2 until further use.

Glucose oxidase/D-(+)-glucose/catalase (GODCAT) O_2 scavenging system

To generate microoxic in vitro conditions, the GODCAT[58] O_2 elimination system was used. For optical absorbance experiments, individual components

included (final concentrations): 40 μg/mL bovine catalase (Sigma), 100-200 μg/mL *Aspergillus niger* glucose oxidase (Sigma), and 0.04% m/v glucose. The GODCAT components were added to GlbN samples and allowed to incubate for at least 5 minutes prior to use.

Electronic absorption spectrophotometry

Optical absorbance data were acquired at room temperature on a Cary50 UV-vis spectrophotometer. Single scan mode was used to evaluate protein (H117A and WT GlbN[3]: ε_{411} =96 mM^{-1} cm^{-1}, WT GlbN-A: ε_{409} = 87 mM^{-1} cm^{-1}) and $NO^{\bullet}$ donor (see below) concentrations. Single scans typically utilized a spectral window of 700-250 nm, with 0.5 s averaging time, and 1 nm step size. Scanning kinetics experiments were used to monitor ligand binding and enzymatic reaction of GlbNs (see below). The acquisition parameters were: spectral window 640–350 nm, 0.1-s averaging time, 1 nm step size. For kinetic experiments, manual mixing dead times were ~15 s. Unless otherwise noted, optical spectra were collected every 0.5 min for 20 min, and then every 5 min for 2 h.

Ferredoxin/$NADP^+$ reduction system

When dithionite was not desirable for GlbN reduction (when oxic conditions were required), a ferredoxin-based enzymatic reduction system[59] was used. This system contains glucose-6-phosphate (G6P) as an electron source, $NADP^+$, G6P dehydrogenase to generate NADPH from G6P and $NADP^+$, and ferredoxin-$NADP^+$ reductase to generated reduced ferredoxin (Fd). Typical optical absorbance experiments included the following components (all from Sigma,

final concentrations): 0.1 % m/v G6P, 0.01 - 0.2 mM $NADP^+$, ~20 U *Leuconostoc mesenteroides* G6P-dehydrogenase, ~0.1 U spinach Fd/$NADP^+$ reductase, 100 μg/mL spinach ferredoxin, and 40 μg/mL bovine catalase (to eliminate any H_2O_2). "Order-of-addition" control experiments indicate that Fd is the active reductant for GlbNs, and that NADPH-mediated reduction of GlbN is negligible.

$NO^\bullet$ binding to ferric GlbNs monitored by optical absorbance

Stock solutions of the $NO^\bullet$ donors dipropylenetriamine-NONOate (DPTA-NONOate, Cayman Chemical) or methylamine hexamethylene methylamine-NONOate (MAHMA-NONOate, Cayman Chemical) were prepared in 0.1 M NaOH. The concentration was evaluated after dilution in H_2O using published extinction coefficients (DPTA-NONOate, ε_{252} = 7.9 mM^{-1} cm^{-1}; MAHMA-NONOate, ε_{250} = 7 mM^{-1} cm^{-1}).[60,61] At pH 13, the NONOates are stable for days to weeks and degradation can be assessed by a decrease in the ~250 nm absorbance band. At pH 7, the NONOates degrade by an apparent single exponential process (DPTA-NONOate, $t_{1/2}$ ~ 3-4 h, MAHMA-NONOate $t_{1/2}$ ~ 3-4 min) to produce two molecules of $NO^\bullet$ per one molecule donor. Fe(III)-GlbN (4-10 μM) samples used for optical absorbance were treated with 0.8–1 mM NONOate in order to produce their Fe(III)-$NO^\bullet$ forms.

$NO^\bullet$ dioxygenase (NOD) assay monitored by optical absorbance

Fe(III) GlbNs (4-10 μM, pH 7.0–7.2) were incubated in O_2-saturated buffer (O_2, Airgas) and in the presence of the Fd-$NADP^+$ reductase system to produce Fe(II)-O_2 GlbNs. Complete reduction and oxygen binding was generally achieved within ~15 min. At this point, 170–900 μM MAHMA-NONOate was added

(depending on the particular experiment) to initiate $NO^{\bullet}$ release and NOD activity. After oxidation, GlbNs were re-reduced and returned to the Fe(II)-O_2 form. For multiple (usually five) $NO^{\bullet}$ addition experiments, 2, 10, or 20-fold MAHMA-NONOate was introduced after each NOD/O_2 re-binding cycle.

Griess assay for detection of NO_3^- and NO_2^-

The acidic Griess assay (Life Technologies) was used for measuring the amount of NO_2^- and NO_3^- produced in GlbN NOD reactions. The assay depends upon the reaction of NO_2^- with sulfanilamide to form a diazonium salt, the latter ultimately reacting to produce a colored azo-dye. A 1 mM NO_2^- standard (Life Technologies) was used to generate a linear calibration curve and quantify product formation (ε_{520} = 26.9 mM^{-1} cm^{-1}). Following an NOD experiment, GlbN samples were split into two; to one half ~4 μM FAD and catalytic amounts of *Aspergillus niger* NO_3^- reductase (+NR) were added in order to convert NO_3^- into NO_2^- stoichiometrically (20-30 minute incubation), to the other half no NO_3^- reductase was added (–NR). The Griess reagents were then added to both +NR and –NR samples to determine total NO_2^- concentrations. Following an additional ~20 min incubation, relative product formation was assessed by recording the absorbance in the 400–600 nm region. The difference in absorbance between +NR and –NR treated samples is proportional to the amount of NO_3^- produced during the GlbN NOD assay.

Dithionite reduction of Fe(III)-$NO^{\bullet}$ GlbNs monitored by optical absorbance

Fe(III) GlbNs were treated with DPTA- or MAHMA-NONOate to produce the Fe(III)-$NO^{\bullet}$ (4-10 μM, pH 7.0-7.2) form. At this point, a stock of sodium

dithionite (DT, Alfa-Aesar) was freshly prepared, and 2 mM DT (final concentration) was added to the Fe(III)-NO• GlbNs to initiate reduction. In a similar experiment, a 10-fold molar excess (40–100 μM) horse skeletal apomyoglobin was added immediately following the DT reduction step (~15 s). Under these conditions, heme dissociation from GlbN and capture by apomyoglobin (forming apo GlbN and myoglobin) can be monitored kinetically owing to the different visible spectra of GlbNs and myoglobin.

NMR spectroscopy of nitrosyl-GlbNs

NMR data were acquired on Bruker Avance or Avance-II NMR spectrometers operating at 600 MHz, each equipped with a cryoprobe. An unlabeled ~1.1 mM Fe(III) GlbN sample (100 mM P, pH 7.1, GODCAT, 298 K) was treated with four-fold excess MAHMA-NONOate to generate the Fe(III)-NO• state. A mixture of Fe(III) and Fe(III)-NO• forms was produced and remained relatively stable for ~6 h. Water presaturation ^{1}H 1-D, ^{1}H-^{1}H NOESY (τ_{mix} = 80 ms), and TOCSY (τ_{mix} = 45 ms) spectra were recorded on the resulting diamagnetic adduct in order to assign heme resonances. Since the Fe(III)-NO• form decayed in less than 12 h, only a preliminary analysis could be achieved. Similarly, an ^{15}N-labeled ~1.4 mM Fe(III) GlbN-A (250 mM P, pH 7.1, GODCAT, 298 K) sample was treated with five-fold molar excess MAHMA-NONOate to produce a mixture of Fe(III) and Fe(III)-NO• GlbN-A. To facilitate heme resonance assignments, presaturation ^{1}H 1-D, flipback-WATERGATE ^{1}H-^{1}H NOESY, and presaturation ^{1}H-^{1}H TOWNY-TOCSY were acquired with ^{15}N-decoupling. ^{1}H-^{15}N HSQC spectra were recorded on a separate ^{15}N-labeled Fe(III)-NO• GlbN-A sample (800 μM protein, 5 mM DPTA-NONOate, 250 mM P, pH 7.1, 298 K). After data acquisition, the

diamagnetic Fe(III)-NO• GlbN-A sample was treated with 8 mM DT, under argon (Airgas) atmosphere. Formation of paramagnetic Fe(II)-NO• GlbN-A was manifest in the absence of detectable heme signals; ^{1}H-^{15}N HSQC spectra were recorded on Fe(II)-NO• GlbN-A to compare with its Fe(III)-NO• state.

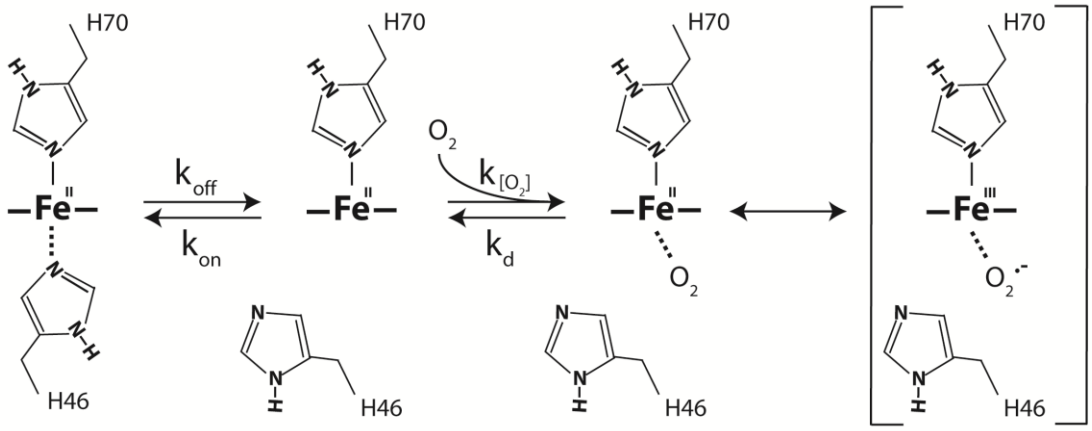
To generate Fe(II)-HNO GlbN, two separate procedures were used. 1) Fe(III) GlbN was treated with MAHMA-NONOate to produce Fe(III)-NO• GlbN (as above) and subsequently reduced with excess DT to form the Fe(II)-HNO state. 2) Fe(III) wild-type (or H117A GlbN) was incubated in the presence of $^{15}NO_2^-$ (or NO_2^-) and reduced with excess DT, resulting in formation of Fe(II)-HNO GlbN. In both cases, the same Fe(II)-HNO form persisted for several hours, and presaturation ^{1}H 1-D, ^{1}H-^{1}H DQF-COSY, and NOESY spectra were used for preliminary heme and HNO assignments.

NMR data processing and analysis

NMR data were processed using Topspin 1.3, Topspin 2.1, or NMRpipe.[62] Spectra were analyzed using Sparky 3.[63] ^{1}H chemical shifts were referenced with respect to water (^{1}H = 4.76 ppm, 298 K).

Results

Because only the non-crosslinked GlbN was obtained from *Synechococcus* cells grown under oxic conditions,[12] the ability of O_2 to bind ferrous GlbN and inhibit spontaneous formation of GlbN-A was first investigated. Scheme 4.3 presents the basic O_2 binding reaction for a *bis*-histidine hemoglobin.[64]



Scheme 4.3. Dioxygen binding and autooxidation in GlbNs. Exogenous ligands such as dioxygen require displacement of the distal histidine prior to heme binding. If k_{off} is slow, the kinetics of oxygen binding ($k_{[O2]}$) can become rate-limited by H46 dissociation and independent of oxygen concentration. GlbNs have a tendency to autooxidize spontaneously to the ferric state and, in the process, reduce oxygen to superoxide ($O_2^{\bullet-}$, right). The superoxide radical can be used for enzymatic purposes (NOD, see below) or damage the porphyrin via heme oxygenase chemistry. Alternatively, escape of $O_2^{\bullet-}$ into solution where H+ ions are available will lead to its spontaneous dismutation into H_2O_2 and O_2. Note that superoxide can also reduce ferric iron to yield ferrous iron and dioxygen.

Exogenous ligand binding requires displacement of the distal histidine (H46). Figure 4.1 presents the X-ray crystal structure of *bis*-histidine GlbN-A (left, PDB: 4MAX) and its cyanide bound form (right, PDB: 4L2M), the latter

used as a proxy for the O_2 bound structure.[4] *Synechococcus* GlbN-A, like several closely related trHbs, undergoes a large conformational change upon switching between *bis*-histidine and "ligand-bound" states.[4,39,65–70] In the ligand-bound structure, the B and E-helices have rotated to present a distinct array of amino acids for interaction with the distal cyanide. The conserved Tyr22 OηH donates an H-bond to the cyanide nitrogen; additionally Gln43 and Gln47 NεH2 groups are positioned to form a network of H-bonds with the cyanide carbon (or π electrons) and the Tyr22 Oη. These interactions are thought to be critical in orienting the distal oxygen for superoxide activation (Scheme 4.3) and promotion of NOD activity.[34,39]

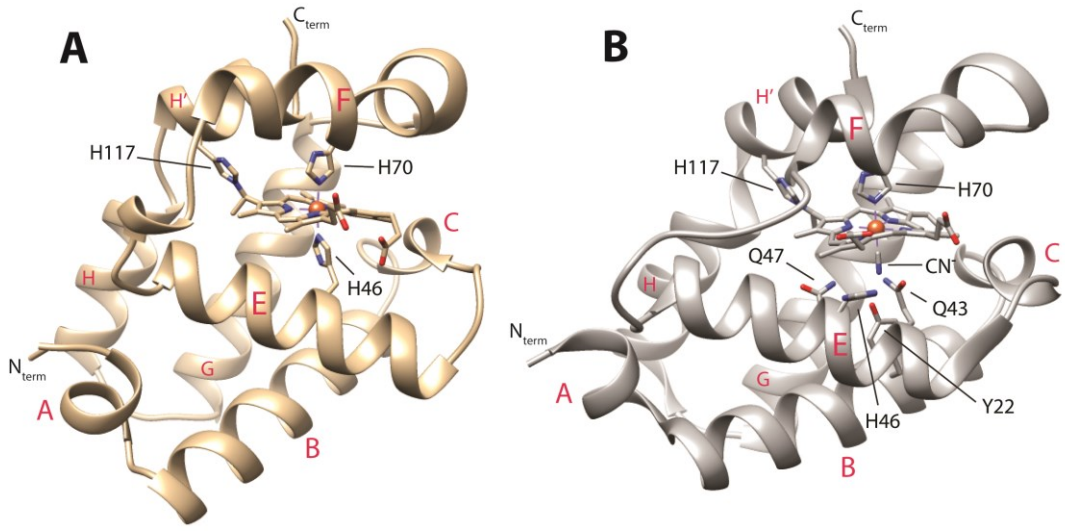


Figure 4.1. A) Crystallographic model (PDB: 4MAX) of *Synechococcus* GlbN with covalently attached heme (hereafter: GlbN-A). The posttranslational modification saturates the 2-vinyl group to form a H117 Nε2-2-Cα heme linkage. The proximal (H70) and distal (H46) histidines are coordinated to the ferric iron. α-helices are labeled in red according to Perutz notation. B) Crystallographic model (PDB: 4L2M) of GlbN-A determined in the ferric cyanide-bound (cyanomet) state. The distal pocket underwent a large structural rearrangement: H46 has flipped away from the heme iron to interact with propionate-7. The rotation of the E- and B-helices introduces Q43 and Q47, along with Y20 Oη-H near the cyanide to stabilize the exogenous ligand via a hydrogen bond network.

Figure 4.2 shows O_2 binding by GlbN monitored by optical absorbance spectrophotometry. The experiment makes use of spinach ferredoxin (Fd) as an electron shuttle along with an enzymatic re-reduction system to maintain GlbN in the reduced state.

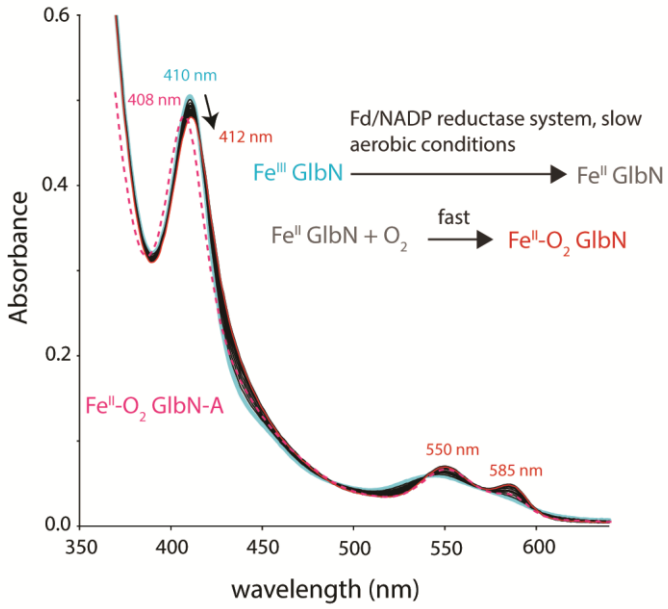


Figure 4.2. Dioxygen binding monitored by optical absorbance spectrophotometry. A ~5 μM ferric GlbN pH 7 sample (cyan spectrum) was incubated in the presence of a catalytic ferredoxin (Fd) reducing system and visible spectra were recorded as a function of time (black traces). The combination of glucose-6-phosphate (G6P), G6P-dehydrogenase (generates NADPH), Fd/NADPH reductase (generates reduced Fd), Fd (to reduce GlbN), and GlbN under oxic conditions results in slow but observable Fe(III) → Fe(II)-O_2 GlbN (red spectrum) conversion. A Fe(II) *bis*-histidine intermediate was not observed indicating that Fd-mediated reduction of GlbN is slow relative to dioxygen binding. The Fe(II)-O_2 GlbN spectrum persists for hours and is distinct from that observed for Fe(II)-O_2 GlbN-A (magenta, dotted spectrum).

Under oxic conditions, slow Fd-mediated reduction of GlbN is followed by rapid O_2 binding resulting in an apparent Fe(III) → Fe(II)-O_2 transition (Figure 2, cyan spectrum to red spectrum). These data indicate that Fd readily reduces GlbN, as expected from its propensity for facile ET.[11]

In the absence of oxygen at neutral pH, reduction of GlbN to the Fe(II) state is followed by rapid (< 1 s) PTM. [7,11] Under oxic conditions, the Fe(II)-O_2 GlbN

species (red spectrum) persists and does not convert to Fe(II)-O_2 GlbN-A (magenta dashed spectrum) for a period of hours. Thus, these spectra support an inhibitory effect of O_2 toward the PTM and are therefore consistent with the presence of GlbN in synechococcal cells grown under oxic conditions.

At substoichiometric oxygen:hemoglobin levels, oxygen is less effective at PTM inhibition. O_2 was bubbled in a 1.4 mM sample of ^{15}N-labeled Fe(III) GlbN NMR to reach an estimated [O2] ~ 1 mM. The sample was subsequently treated with the Fd-reductase system. Under these conditions, a majority of GlbN converted to GlbN-A within the ~10 min dead time. I hypothesize that Fe(II) GlbN crosslinking to Fe(II) GlbN-A followed by subsequent oxygen exchange with unreacted Fe(II)-O_2 GlbN (to produce Fe(II)-O_2 GlbN-A and Fe(II) GlbN) makes O_2 less effective at inhibiting the PTM under these NMR conditions. More interestingly, the ^{1}H 1-D data shown in Figure 4.3 support that Fe(II)-O_2 GlbN-A slowly converted to Fe(II) *bis*-histidine GlbN-A. Prolonged incubation leads to complete elimination of the O_2 bound form. Thus, in the presence of the Fd reducing system, it appears that GlbN-A can consume dioxygen. Slow autooxidation (Fe(II)-O_2 $\rightarrow$ Fe(III) + $O_2^{\bullet-}$), followed by Fd reduction of Fe(III) to Fe(II) GlbN-A may account for the observed spectral changes (Figure 4.3). Overall, the binding and activation of O_2 by GlbNs suggests a capacity for enzymatic oxidative chemistry.

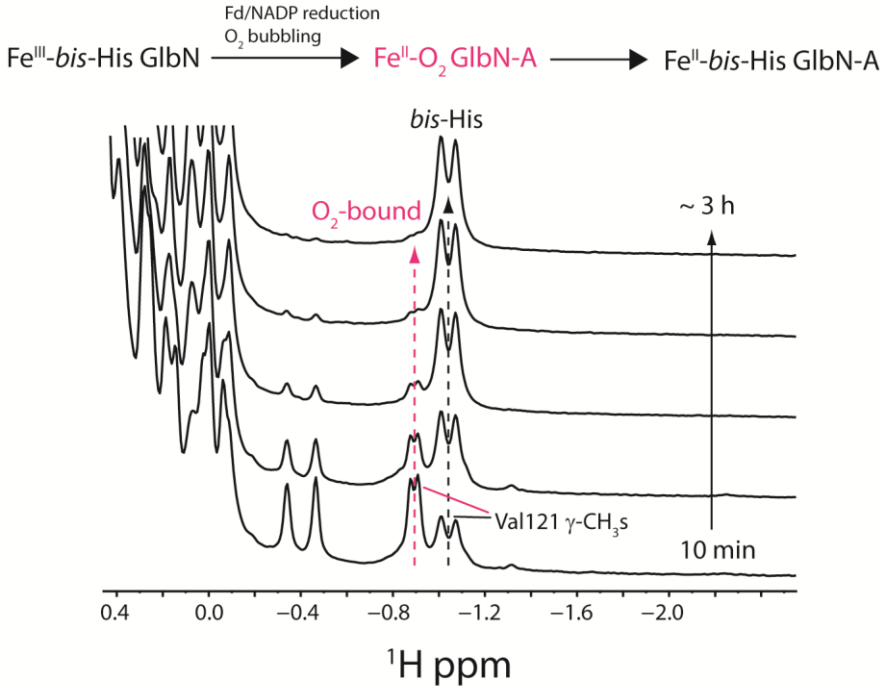
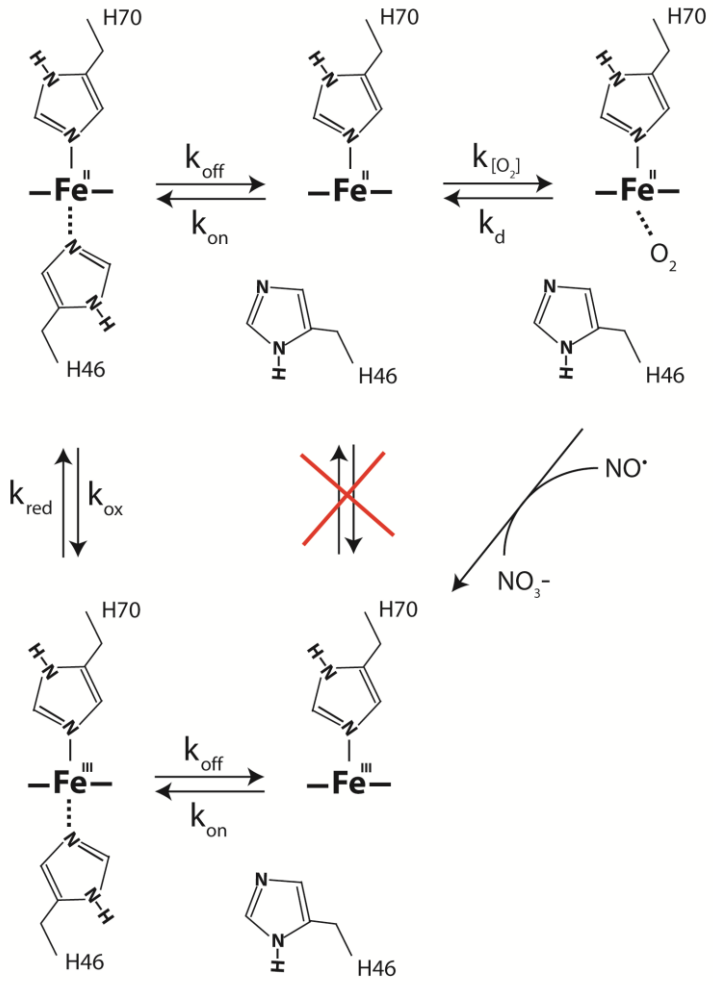


Figure 4.3. ^{1}H NMR evidence for GlbN-A activation of O_2 to $O_2^{\bullet-}$. An ~1.5 mM GlbN sample was saturated with dioxygen gas (~ 1 mM), incubated with the Fd reduction system, and transferred to a Shigemi NMR tube under argon. Under these conditions, conversion to GlbN-A occurred. Putative signals from Val121 were used to track initial generation of Fe(II)-O_2 GlbN-A. Prolonged incubation lead to the gradual depletion of Fe(II)-O_2 GlbN-A and stable production of *bis*-histidine Fe(II) GlbN-A.

To examine the ability for GlbNs to undergo the NO$^\bullet$ dioxygenase reaction (NOD, Scheme 4.4), an optically-monitored method was utilized.



Scheme 4.4. Nitric oxide dioxygenase (NOD) reaction cycle in GlbNs. Starting in the ferrous *bis*-histidine state (top left), the distal histidine (H46) is displaced and dioxygen binds (top right). The activated Fe(III)-$O_2^{\bullet-}$ complex reacts with exogenous $NO^{\bullet}$ to form less reactive nitrate anion (NO_3^-) and oxidized Fe(III) GlbN (bottom right). To complete a single NOD cycle, the Fe(III) protein must undergo electron transfer to return to its initial Fe(II) state.

GlbN (or GlbN-A) was treated with the Fd-$NADP^+$ reduction system under oxic conditions to generate Fe(II)-O_2 GlbN. After oxygen saturation, the $NO^{\bullet}$ releasing agent MAHMA-NONOate was added. Figure 4.4 shows representative results for such an experiment. Within the dead-time (15 s) of $NO^{\bullet}$ addition,

Fe(II)-O_2 GlbN (red spectrum) had oxidized to the Fe(III) *bis*-histidine state (blue spectrum). This latter spectrum persisted for ~10 minutes prior to returning to that of Fe(II)-O_2 GlbN (see 550 nm and 585 nm time traces, also pink spectrum). The buildup of the Fe(III) *bis*-histidine GlbN intermediate during the "turnover" period indicates that under these conditions, Fd-mediated reduction of Fe(III) to Fe(II) GlbN is rate-limiting for an NOD cycle.

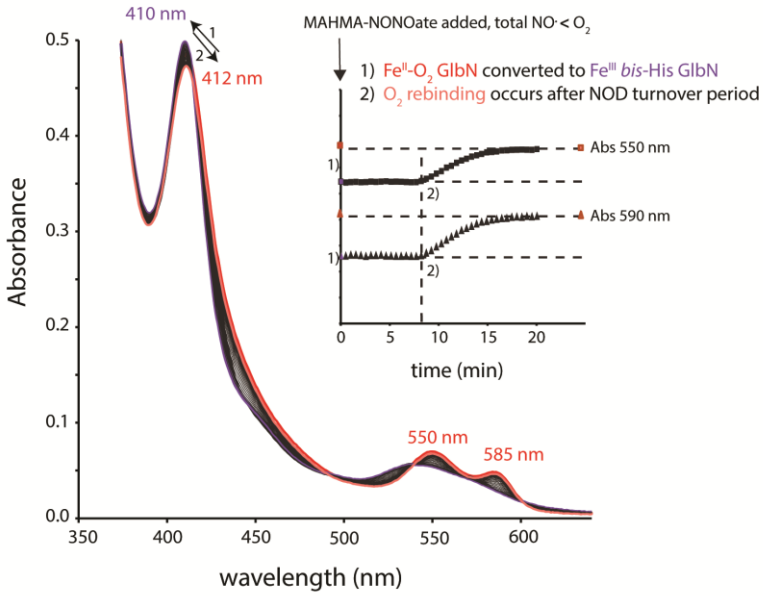


Figure 4.4. NOD assay performed with ~5 μM Fe(II)-O_2 GlbN in oxygen-saturated buffer. Following Fd-mediated reduction and dioxygen binding, Fe(II)-O_2 GlbN (initial red spectrum) was treated with ~175 μM of the nitric oxide releasing agent MAHMA-NONOate (addition marked with arrow, ~350 μM total $NO^{\bullet}$ released, $t_{1/2}$ ~ 2-3 min at pH 7, room temperature). Within the dead time, Fe(II)-O_2 GlbN is oxidized to the Fe(III) *bis*-histidine form (cyan spectrum) and this state persists for ~8 minutes during the "NOD turnover" period (see 550, 590 nm traces, inset). Once $NO^{\bullet}$ levels drop below the protein concentration (approximated with vertical dashed line), apparent conversion of Fe(III) *bis*-histidine to Fe(II)-O_2 is observed (final pink spectrum).

The oxidation of Fe(II)-O_2 GlbN by $NO^{\bullet}$ to form Fe(III) GlbN is highly suggestive of NOD chemistry. At this point, however, the nature of the product remains undetermined. In the absence of any enzyme, $NO^{\bullet}$ may react with O_2 in solution to form exclusively NO_2^-.[71] Therefore we next examined whether NO_3^- was the primary product generated during putative GlbN NOD activity. Under the condition of excess oxygen, GlbN NOD chemistry would result in a stoichiometric conversion of $NO^{\bullet} \rightarrow NO_3^-$. NOD competes with the spontaneous $2NO^{\bullet} + O_2 \rightarrow 2NO_2^-$ conversion process (NO_2^- product) for $NO^{\bullet}$ consumption. Following an NOD assay, the GlbN sample is split, and to one half the enzyme NO_3^- reductase is added (+NR) for catalytic reduction of nitrate to nitrite ($NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$). The second half of the GlbN sample is left without added NO_3^- reductase (–NR). The Griess test is an established colorimetric assay used for quantitative evaluation of NO_2^- concentration (Figure 4.5).

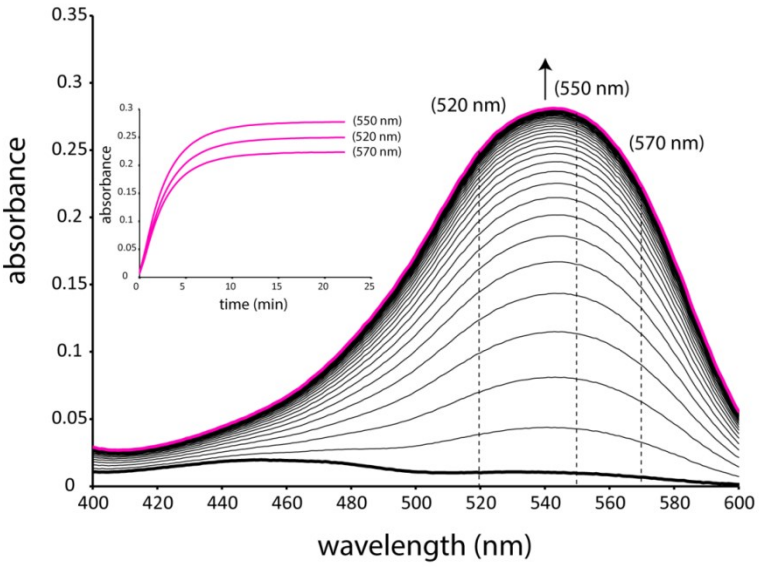


Figure 4.5. Absorbance data from Griess assay used for the detection of nitrate and nitrite produced in the GlbN NOD assay. The GlbN NOD reaction produces nitrate which can be transformed by nitrate reductase (+NR) to nitrite. Subsequent addition of the Griess reagents to a GlbN NOD (+NR) sample leads to colorimetric product formation. In the absence of NR, NOD chemistry produces nitrate and addition of the Griess reagents leads to minimal buildup of colored product (not shown). Inset: kinetics of azo-dye formation during the Griess reaction of a GlbN NOD sample. Product concentration was assayed after 20 min.

For the GlbN NOD sample lacking NR (–NR), all colored product will correspond to NO_2^- yield. The GlbN NOD sample treated with NR will show color formation related to the NO_3^- + NO_2^- yield; therefore the difference in color between +NR and –NR samples corresponds solely to the NO_3^- yield. When performed on the GlbN and GlbN-A NOD reactions, the Griess assay indicated that ~50-90 % of the added $NO^\bullet$ was converted into NO_3^- (data not shown). In the absence of GlbNs, $NO^\bullet$ addition leads to predominant formation of NO_2^-, as expected.[71] Therefore these results support that GlbNs can catalyze the dioxygenation of $NO^\bullet$ to form NO_3^-.

To test whether GlbNs can tolerate repeated $NO^{\bullet}$ challenges, the NOD assay was extended to include consecutive MAHMA-NONOate additions. Figure 4.6 compares single (Figure 4.6A,B) and multiple $NO^{\bullet}$ addition (Figure 4.6C) experiments performed on GlbN under aerobic conditions. Horse skeletal myoglobin was used as a reference and its response is shown in Figure 4.6D. As seen from the plots of the α and β bands (Figure 4.6C), repeated MAHMA-NONOate additions result in multiple Fe(II)-O_2 (oxidation) $\rightarrow$ Fe(III) *bis*-histidine ("turnover") $\rightarrow$ Fe(II)-O_2 (re-binding) cycles. However, following each $NO^{\bullet}$ addition, the turnover period increased in duration, O_2 re-binding occurred more slowly (in part because of O_2 consumption), and recovery to the oxy state was clearly attenuated, an indication of GlbN deactivation via heme bleaching. This latter effect was exacerbated in GlbN-A (not shown). In comparison, myoglobin shows relatively little damage (Figure 4.6D). Thus, although GlbNs are capable of in vitro NOD activity, the high levels of heme damage observed upon repeated $NO^{\bullet}$ exposure casts doubt on a NOD function for GlbNs in *Synechococcus* cells.

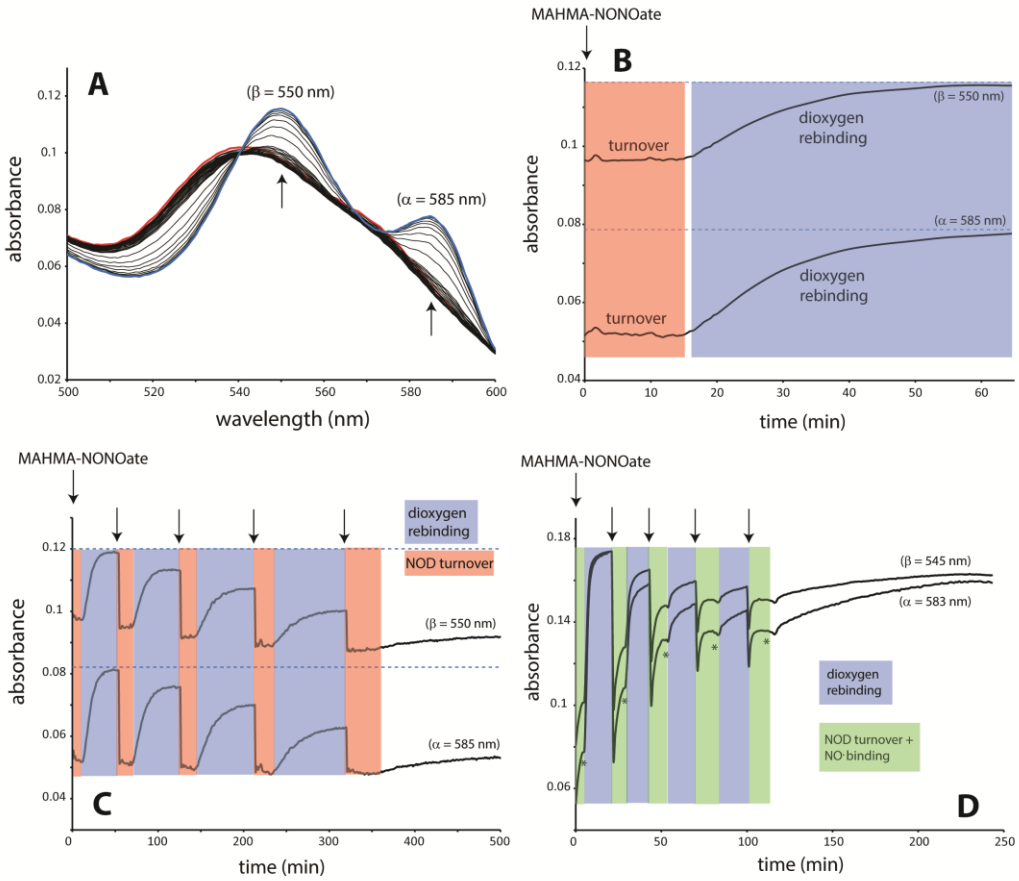
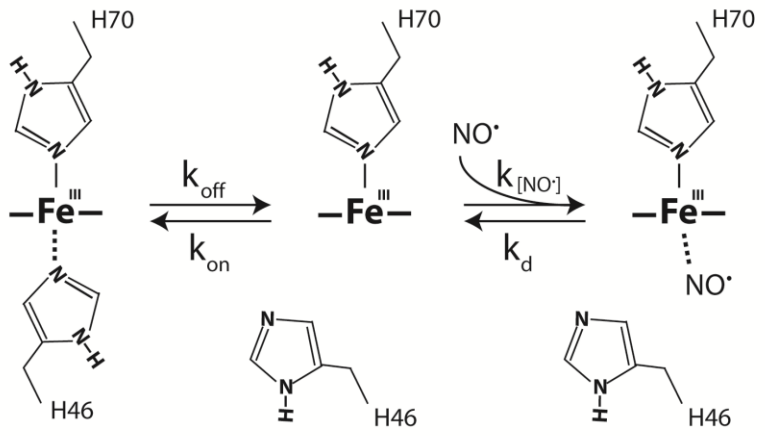


Figure 4.6. Single and multiple NO• addition NOD assays of GlbN and horse skeletal myoglobin. A) Zoom-in of α/β visible region of absorbance spectra. The α and β absorbance bands are highly sensitive to iron ligation status and differentiate Fe(II)-O_2 (blue spectrum), Fe(II) *bis*-histidine, Fe(III) *bis*-histidine (red spectrum) and NO• bound GlbN forms. B) A single addition of MAHMA-NONOate (marked with arrow) results in the immediate oxidation of Fe(II)-O_2 GlbN to Fe(III), NOD turnover, followed by slow recovery to the oxy state. C) The ability for GlbN to withstand repeated NO• exposure was tested with a multiple addition experiment. Arrows mark each MAHMA-NONOate addition; heme damage is clearly evident by the decay in α and β absorbance bands. D) In contrast, myoglobin is a known NOD and is capable of undergoing multiple NO• additions with relatively little damage. The additional phase marked with an asterisk corresponds to Fe(II) + NO• → Fe(II)-NO• binding, which competed with NOD in Mb, but was not observed in GlbNs.

The ability for GlbNs to bind NO• in the ferric state (Scheme 4.5) and undergo spontaneous autoreduction (Ferric reaction 4.3) was next examined.



Scheme 4.5. Ferric NO• binding scheme for *bis*-histidine GlbNs. The strong field ligand NO• has pM affinity for ferrous iron and can also bind heme in the ferric state (mM–μM affinity).

Unlike horse skeletal myoglobin and several other hemoglobins,[52,54] treatment of GlbN (or GlbN-A) with NO• donor (MAHMA-NONOate) lead to stable formation of the Fe(III)-NO• (GlbN-A, Figure 4.7).

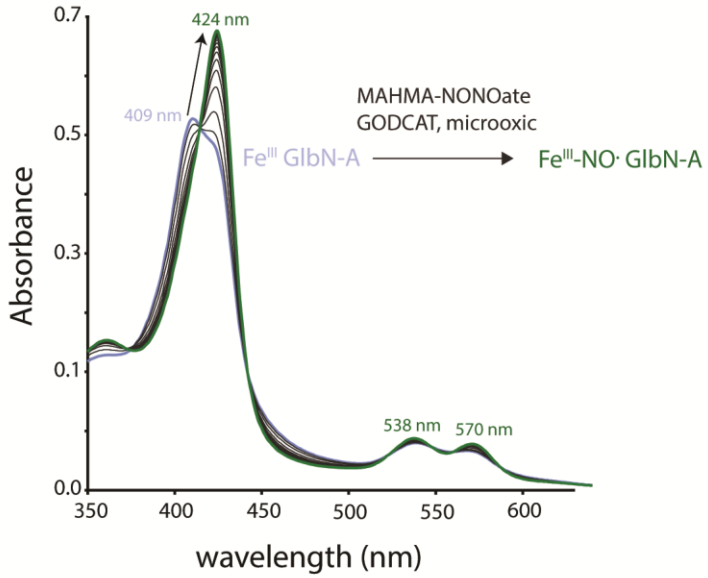


Figure 4.7. Stable formation of Fe(III)-NO• GlbN-A monitored by optical absorbance spectrophotometry. To limit background O_2 + NO• side reactions, the ~5 μM Fe(III) GlbN-A (gray spectrum) pH 7 solution was incubated with an enzymatic O_2 elimination system (D-(+)-glucose, glucose oxidase, catalase (GODCAT)) for ~10 min prior to NO• addition. Treatment of GlbN-A with ~1 mM MAHMA-NONOate resulted in complete formation of the Fe(III)-NO• GlbN-A complex (green spectrum).

I hypothesize that formation of the hydrogen bond network involving Tyr22, Gln43, and Gln47 stabilizes the Fe(III)-NO• state (Figure 4.1B) and protects against base-catalyzed reduction (Ferric reaction 4.3). NMR spectroscopy was used to interrogate the nature of the Fe(III)-NO• complexes of GlbN and GlbN-A. ^{1}H NMR data supported that both complexes were diamagnetic (S = 0), consistent with antiferromagnetic coupling between the Fe(III) (S=1/2) center and NO• (S = 1/2). The relative stability and diamagnetism of these complexes allowed for acquisition of two-dimensional ^{1}H-^{1}H NOESY and TOCSY experiments, and resolved heme resonances could be assigned. In the

diamagnetic Fe(III)-NO• state, a ring current from the aromatic porphyrin acts to shift in-plane heme protons (e.g., meso and methyl protons) several ppm downfield. Portions of NOESY and TOCSY spectra used for preliminary heme resonance assignment in Fe(III)-NO• GlbN are shown in Figure 4.8 (see Table 4.1 for heme ^{1}H chemical shifts of Fe(III)-NO• GlbN and Fe(III)-NO• GlbN-A).

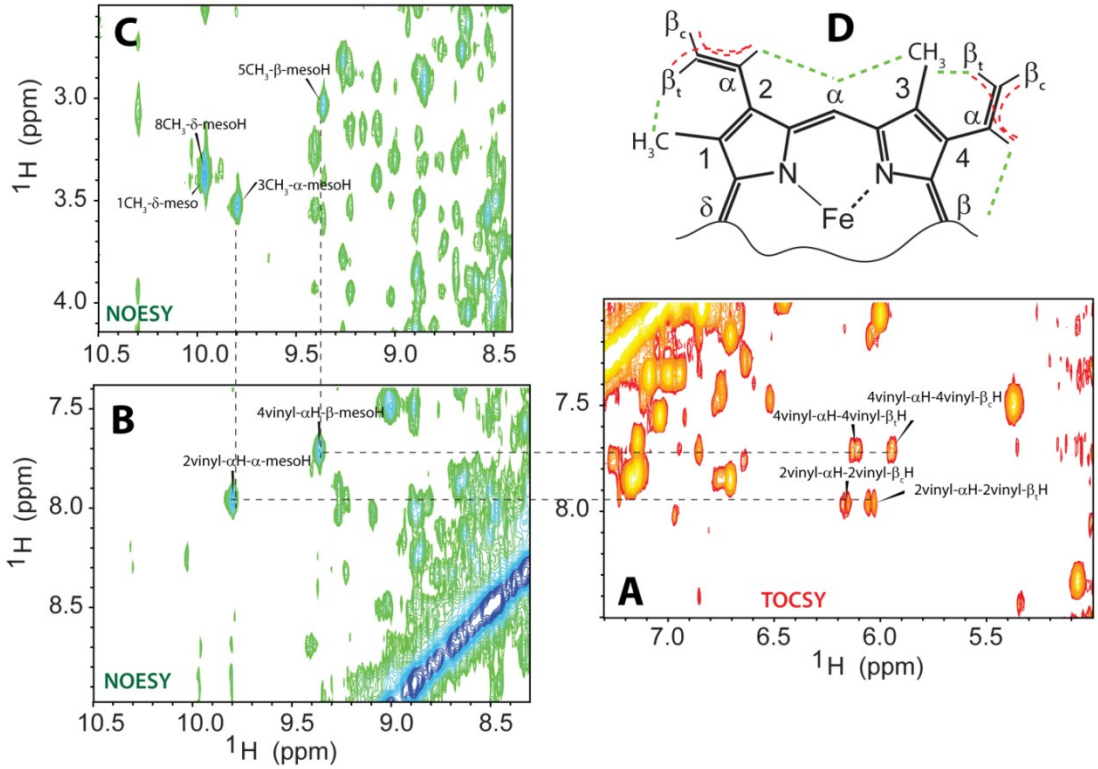


Figure 4.8. Portions of ^{1}H-^{1}H TOCSY and NOESY NMR spectra collected on a ~1 mM Fe(III)-NO• GlbN sample, pH 7, 298 K, in the presence of the GODCAT O_2 scavenging system. Fe(III)-NO• GlbN was produced by incubation of Fe(III) GlbN with 5-fold excess DPTA-NONOate (an NO• releasing agent with $t_{1/2}$ ~ 3-4 h at pH 7, 298 K). The NMR spectra were consistent with the formation of a persistent diamagnetic (Fe(II)-NO$^+$ GlbN) complex. GlbN-A also displayed stable Fe(II)-NO$^+$ binding. A) Heme vinyl Hα-Hβ_{cis}-Hβ_{trans} TOCSY correlations were identified by their characteristic chemical shifts and J-couplings. B) Strong vinylic Hα to heme meso proton NOEs indicate that both vinyls adopt a *trans* configuration. C) Heme meso to methyl proton NOEs facilitated heme assignments. D) Model of Fe(III)-NO• GlbN heme vinyl geometry.

Since both GlbN and GlbN-A were maintained in the ferric nitrosyl state on the order of hours, it is highly unlikely that the autoreduction (Ferric reaction 4.3) is relevant to their ability to protect *Synechococcus* cell from RNS/ROS.

We next examined whether the Fe(III)-NO$^{\bullet}$ GlbN-A and Fe(III)-NO$^{\bullet}$ GlbN complexes could undergo forced reduction to the Fe(II)-NO$^{\bullet}$ form. Although neither GlbN nor GlbN-A underwent detectable autoreduction, NO$^{\bullet}$ binding to GlbNs is expected to increase the heme redox potential and favor reduction given a suitably reducing environment. Dithionite (DT) is a relatively powerful reducing agent[72] and was first used to test for formation of the ferrous nitrosyl adducts. The Fe(III)-NO$^{\bullet}$ form of GlbN-A was prepared by treatment of Fe(III) GlbN-A with excess MAHMA-NONOate and monitored optically. Upon saturation with NO$^{\bullet}$, the sample was reduced by addition of 2 mM DT (Figure 4.9A). Within the dead time of the experiment (~15 s), rapid changes from the starting material (green spectrum) were observed and suggested that complete reduction to the Fe(II)-NO$^{\bullet}$ GlbN-A form (magenta spectrum) had occurred. An analogous experiment performed using an ^{15}N-labeled NMR sample of Fe(III)-NO$^{\bullet}$ GlbN-A. Figure 4.9B shows the initial ^{1}H-^{15}N HSQC spectrum of Fe(III)-NO$^{\bullet}$ GlbN-A (green peaks). Although isoelectronic with Fe(II)-CO GlbN-A, Fe(III)-NO$^{\bullet}$ GlbN-A gave rise to a distinct ^{1}H-^{15}N correlation map. Reduction of Fe(III)-NO$^{\bullet}$ GlbN-A with excess DT resulted in significant ^{1}H-^{15}N HSQC spectral changes (Figure 4.9B, magenta spectrum). The data support that complete formation of Fe(II)-NO$^{\bullet}$ GlbN-A had occurred. Interestingly, the ferrous nitrosyl spectrum

(magenta) was virtually identical to the previously assigned carbonmonoxy (Fe(II)-CO GlbN-A) spectrum.[14]

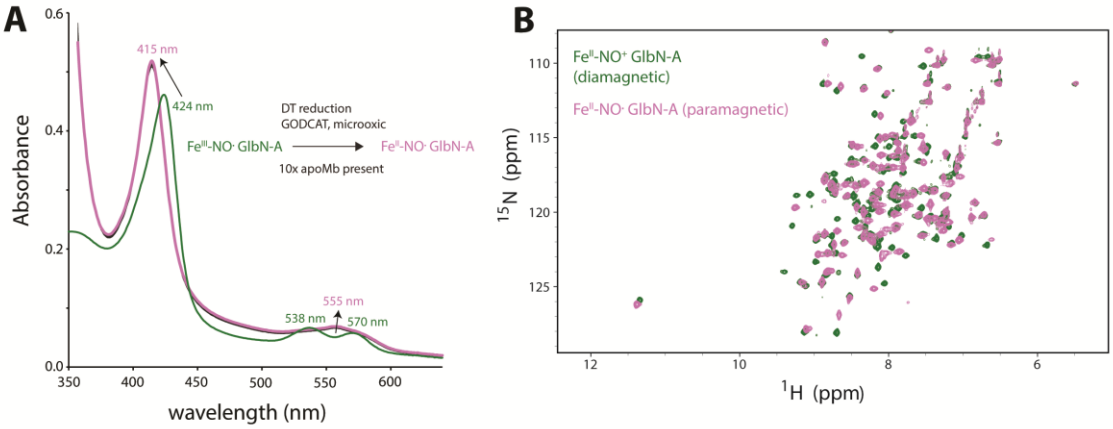


Figure 4.9. Forced reduction of Fe(III)-NO$^{\bullet}$ GlbN-A by DT. A) An ~4 μM Fe(III) GlbN-A pH 7.1 sample containing GODCAT was treated with 1 mM MAHMA-NONOate until complete formation of Fe(III)-NO$^{\bullet}$ (green spectrum). Addition of 2 mM DT causes rapid reduction of GlbN-A from Fe(III)-NO$^{\bullet}$ to Fe(II)-NO$^{\bullet}$ (magenta spectrum). The transition is characterized by blue shifting of the Soret from 424 to 415 nm and coalescence of the α/β region (555 nm). As expected for a protein with a covalently bound cofactor, heme transfer was not observed upon addition of 10-fold excess apomyoglobin (not shown). B) Overlay of ^{1}H-^{15}N HSQC spectra recorded on ^{15}N-labeled Fe(III)-NO$^{\bullet}$ (green spectrum) and Fe(II)-NO$^{\bullet}$ GlbN-A (magenta spectrum) forms (600 MHz, pH 7.1, 298 K). Overall the spectra are similar, supporting that both nitrosyl forms adopt the same overall fold. However, in Fe(II)-NO$^{\bullet}$ GlbN-A, the additional unpaired electron causes rapid paramagnetic relaxation for nearby NH nuclei and therefore broadens several signals beyond detection.

The reduction behavior of ferric nitrosyl GlbN-A was used as a point of reference for GlbN. As for GlbN-A, an Fe(III)-NO$^{\bullet}$ GlbN-A sample was produced by treatment of the ferric protein with excess MAHMA-NONOate and monitored optically. Following saturation, 2 mM DT was added to initiate reduction. Figure 4.10A shows representative results of the experiment. GlbN displayed

distinctive biphasic behavior compared to GlbN-A (Figure 4.9A). The first (rapid) phase was completed within ~3 min and was characterized by a marked decrease in absorbance occurring concomitantly with blue-shift (brown spectrum); this initial rapid phase was followed by a slow ($t_{1/2}$ ~ 1 h) recovery phase (purple spectrum: 418 nm Soret, 558 nm α/β). The low intensity of the intermediate (brown spectrum) and ~380 nm shoulder supports that formation of Fe(II)-NO• GlbN prevents ferrous heme crosslinking and instead causes rapid heme dissociation.

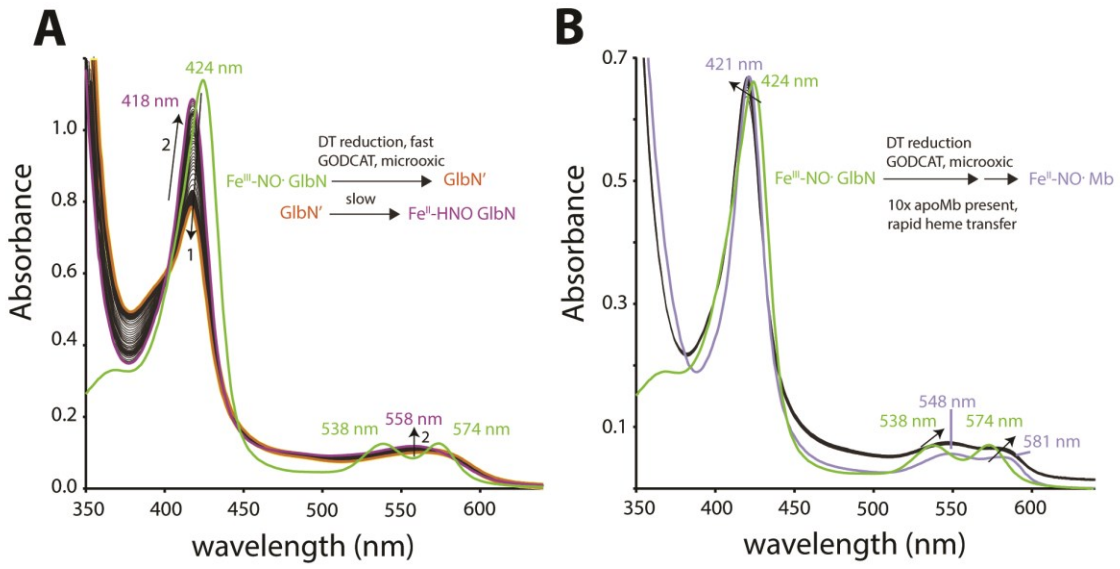
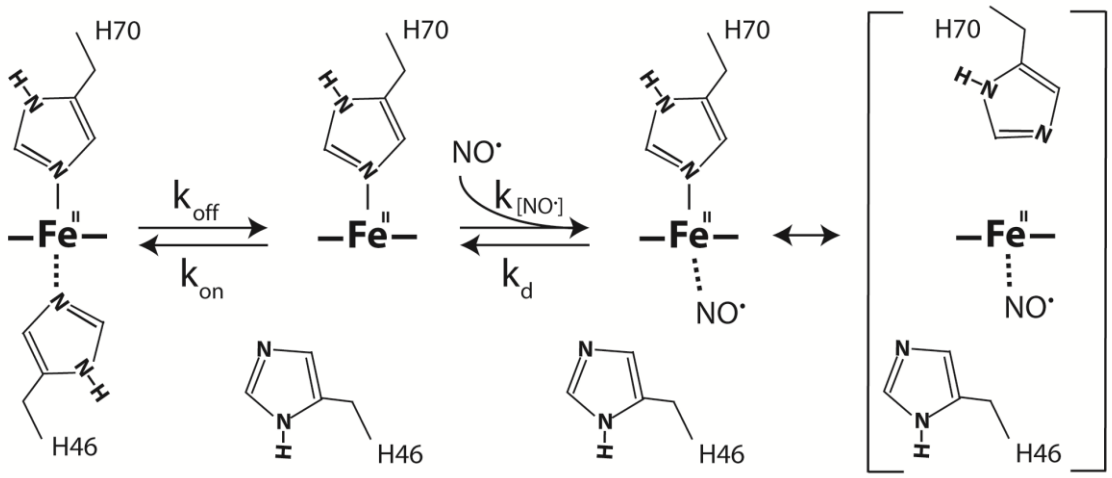


Figure 4.10. Forced reduction of Fe(III)-NO• GlbN by dithionite (DT). A) An ~10 μM Fe(III) GlbN (pH 7, GODCAT) sample was incubated with 800 μM MAHMA-NONOate until complete Fe(III)-NO• formation (green-yellow spectrum: 424 nm Soret, 574 nm α, 538 nm β). Sodium dithionite (2 mM) was added to initiate reduction. Unlike GlbN-A, a biphasic process ensued. B) To test the nature of the 1st phase, a 4 μM GlbN sample was treated as in A). However, following DT reduction a 10-fold excess of apomyoglobin was immediately mixed into the sample prior to spectral acquisition. In this case, complete formation of Fe(II)-NO• *myoglobin* occurred (421, 548, 581 nm species) during the 15 s dead time, an indication of rapid GlbN → apomyoglobin Fe(II)-NO• heme transfer.

NO• is well known for its strong negative *trans* effect (depicted in Scheme 4.6).[31,32]



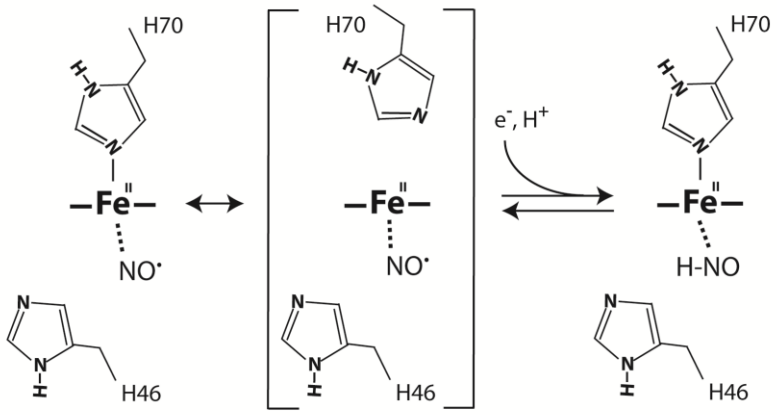
Scheme 4.6. Proposed mechanism for heme loss in Fe(II)-NO• GlbN. Unlike carbon monoxide, which typically exerts a moderate positive *trans* effect (CO binding strengthens the *trans* axial ligand), the short bond formed between NO• and ferrous iron causes a strong, negative trans effect[73] and weakening of the H70 Nε2-Fe bond. Complete scission of the proximal histidine is expected to lead to rapid Fe(II)-NO• heme dissociation as observed for GlbN in Figure 4.10. In contrast, all spectroscopic evidence (Figure 4.9) supports that H70 remains ligated to Fe(II)-NO• heme within the binding pocket of GlbN-A.

For ferrous heme, NO• can bind with pM affinity[74] and can cause deligation of the *trans* iron ligand (in GlbNs, H70). Proteins such as soluble guanylate cyclase use the *trans* effect to trigger a conformational change which activates its GTP cyclase activity by several hundred fold.[75] To test whether H70 de-coordination and heme loss was responsible for the first phase observed in Figure 4.10A, a heme transfer experiment was conducted. Specifically, Fe(III)-

$NO^{\bullet}$ GlbN was prepared as in Figure 4.10A. However, following DT addition and mixing (~15 s) a 10-fold molar excess of apomyoglobin was added prior to acquiring UV-visible spectra. Whereas apomyoglobin is colorless, Fe(II)-$NO^{\bullet}$ myoglobin has distinct visible absorbance features that can be used to report on heme transfer from GlbN. Figure 4.10B shows the results of DT reduction of Fe(III)-$NO^{\bullet}$ GlbN in the presence of 10x apomyoglobin. In contrast with Figure 4.10A, no initial decrease in absorbance was observed; instead, a spectrum corresponding to Fe(II)-$NO^{\bullet}$ myoglobin was immediately detected. Thus, these data support that the first phase detected upon reduction of Fe(III)-NO GlbN corresponds to H70 de-ligation and heme loss. In contrast, heme transfer experiments performed with ferric *bis*-histidine or ferrous carbonmonoxy GlbN result in only slow heme loss over a timescale of hours (data not shown). Comparison of Figure 4.9A and Figure 4.10A indicates that the histidine-heme crosslink protects GlbN-A from $NO^{\bullet}$-mediated heme release and points to a stabilization role for the PTM.

Figure 4.10 and Scheme 4.6 provide evidence and rationale for initial heme loss upon reduction of Fe(III)-$NO^{\bullet}$ GlbN to Fe(II)-$NO^{\bullet}$ GlbN. ^{1}H NMR was used to address the nature of the second phase and final product (Figure 4.10A, purple spectrum). In this experiment, the NO_2^- reductase activity of GlbN generated $NO^{\bullet}$ in situ (see Ferrous reaction 4.2). Figure 4.11 displays the ^{1}H NMR downfield region of H117A GlbN (incapable of forming His-heme PTM), treated with DT in the presence of NO_2^- and allowed to recover for > 1 h. A diamagnetic species with relatively sharp lines was observed, which is inconsistent with formation of Fe(II)-$NO^{\bullet}$ GlbN. Additionally, a ^{1}H signal reminiscent of nitrosyl

hydride (^{1}HNO) within Fe(II)-HNO myoglobin[76] was detected at ~14.8 ppm. To determine whether HNO was formed upon DT reduction of Fe(III)-NO$^\bullet$ H117A and wild-type GlbNs, an ^{15}N-chase experiment was performed. Here, unlabeled Fe(III)-NO$^\bullet$ GlbN was reduced in the presence of $^{15}NO_2^-$. Figure 4.11B shows the downfield region of the resulting ^{1}H NMR spectrum; clearly, the putative HNO signal had split into a doublet ($^1J_{NH}$ ~ 70 Hz), indicating that the ~14.8 ppm proton is attached to an ^{15}N nucleus. Moreover, since the protein was not labeled with ^{15}N, the splitting indicated that the H^{15}NO produced had to originate from $^{15}NO_2^-$. Thus, whereas many hemoglobins are capable of efficiently trapping HNO[77,78], GlbN (and H117A variant) has the ability to produce HNO from NO_2^- via NO$^\bullet$ (NO$^\bullet$ reductase, Scheme 4.7).



Scheme 4.7. Proposed nitric oxide reductase reaction for wild-type and H117A GlbNs. Dithionite reduction of Fe(III)-NO$^\bullet$ GlbN or Fe(III) *bis*-histidine GlbN in the presence of nitrite causes H70 deligation, heme loss and slow HNO formation along with heme re-binding. Only a relatively low level (~ < 30 %) of HNO could be trapped by GlbNs under the NMR conditions. In GlbN-A, heme loss is prevented and HNO production is completely extinguished.

The magnitude of the HNO $^{1}J_{NH}$-coupling was in excellent agreement with those previously determined for Fe(II)-HNO heme complexes.[76,78] The second, minor peak (^{1}H = 14.70 ppm, $^{1}J_{NH}$ ~ 70 Hz) likely corresponds to Fe(II)-HNO trapped by GlbN-A that had formed during the reduction procedure. It is noteworthy that no HNO and only Fe(II)-NO$^{\bullet}$ GlbN-A is detected when GlbN-A is subjected to DT reduction in the presence of NO_2^-.

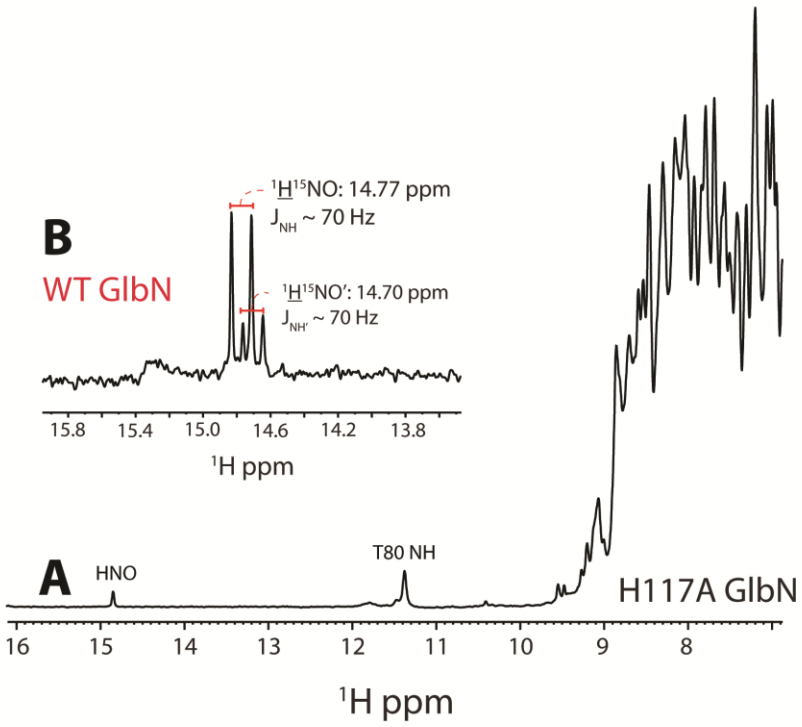


Figure 4.11. Downfield region of ^{1}H NMR spectra collected on Fe(II)-HNO GlbNs. H117A GlbN was used as an analogue of wild-type but is incapable of heme PTM. The ~1.3 mM ^{15}N H117A, pH 7.3 sample was treated with ~5 mM sodium nitrite (NO_2^-) and reduced with ~5 mM DT to initiate nitrite reductase activity and produce $NO^\bullet$ in situ. NMR data were recorded after > 1 h incubation (until recovery of the 418 nm, 558 nm species, see Figure 4.10A, purple spectrum). A) Presaturation and flipback-watergate ^{1}H NMR (^{15}N-decoupled) data collected on H117A after such treatment indicated the buildup of a diamagnetic species with most chemical shifts resonating between 10 and -1 ppm. The NH of T80, which donates a hydrogen bond to H83 Nδ1, is the furthest downfield-shifted protein ^{1}H signal. Additionally, there is a highly unusual downfield ^{1}H (~14.8 ppm) with weak intensity. B) An ^{15}N-chase experiment performed with unlabeled wild-type GlbN (400 μM), 5 mM ^{15}N-labeled nitrite, and dithionite demonstrates the assignment of the downfield ~14.8 ppm ^{1}H signal. In this case, the downfield peak (14.77 ppm) detected in ^{15}N-coupled ^{1}H spectra shows a $^1J_{NH}$ splitting of ~70 Hz. Thus, the ~14.8 ppm proton is attached to the nitrogen which originated from ^{15}N-labeled nitrite. Following nitrite reduction to $NO^\bullet$, an additional 1 e^- reduction and protonation must occur to form Fe(II)-HNO GlbN.

^{1}H-^{1}H NOESY and DQF-COSY spectra collected on Fe(II)-HNO GlbN allowed for partial heme assignments and indicated that both vinyl groups remained intact (see Table 4.1 for assigned ^{1}H shifts). Therefore upon $NO^{\bullet}$ or HNO binding, the spontaneous Fe(II) His-heme PTM was inhibited. Additionally, an NOE between HNO and heme 8-methyl group provided preliminary orientation information for the distal ligand and facilitates comparison with the detailed studies of HNO-myoglobin.[79]

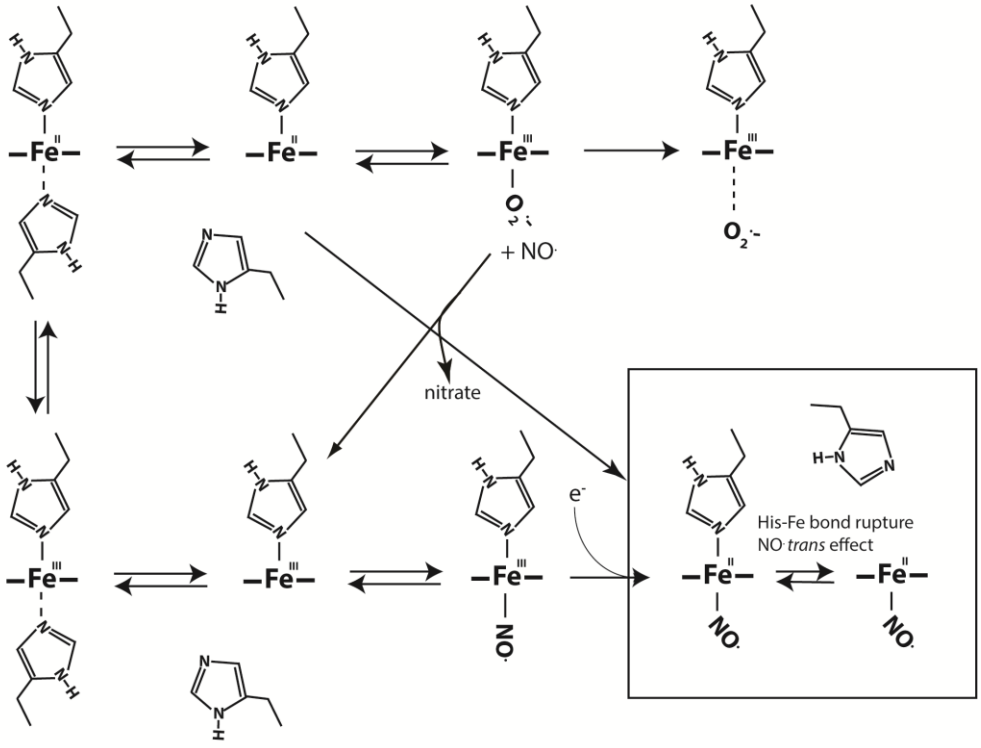
Discussion

NOD activity of GlbN and GlbN-A

An increasing number of studies have implicated hemoglobins as a potential detoxifier of biologically-produced $NO^{\bullet}$.[36,80,81] The current chapter provides in vitro support that both GlbN and GlbN-A are capable of NOD activity, but upon multiple exposures, undergo significant levels of heme damage. Unlike the proposed role for the heme covalent modifications in mammalian heme peroxidases,[82] the His-heme crosslink of GlbN does not appear to prevent heme damage during catalytic activity. This observation, along with the fact that GlbNs protect the cyanobacterial cell from ROS/RNS under microoxic conditions,[12] suggests an alternative role other than dedicated NOD enzyme. It is also possible that NOD-mediated damage of GlbN serves as a signal for RNS/ROS stress in synechococcal cells, since $NO^{\bullet}$ + $O2^{\bullet-}$ side reactions can lead to other PTMs such as tyrosine nitration.[83,84]

Fe(II)-NO• binding by GlbN and GlbN-A

GlbN and GlbN-A show marked differences with respect to their ability to bind NO• stably in the ferrous state. Fe(II) GlbN-A binds NO• readily, like many heme proteins.[74,85,86] On the other hand, if Fe(II) GlbN combines with NO• prior to spontaneous His-heme crosslinking, the heme group readily dissociates as demonstrated by the apo Mb scavenger experiment (Figure 4.10B). NO• mediated heme loss could potentially interfere with any enzymatic cycle involving ferrous iron, including NOD (Scheme 4.8).



Scheme 4.8. NO• binding may interrupt NOD activity in GlbN and cause heme loss. Formation of the His-heme PTM prevents this dissociation reaction.

Thus, it appears that the GlbN His-heme PTM is required for stable $NO^{\bullet}$ binding. Similarly, a stabilization role has been proposed for the analogous His-heme PTM in ferrous *Synechocystis* GlbN[87] (59% identical with *Synechococcus* GlbN). In addition to protecting against heme loss, crosslinking in *Synechococcus* GlbN-A completely extinguishes DT-mediated HNO formation. Why GlbN but not GlbN-A can generate HNO is not obvious, as they have similar Fe(III)/Fe(II) reduction potentials (in their *bis*-His forms).[11] Perhaps in the ligand bound state of GlbN, H70 de-ligation from the heme is a necessary requirement for $NO^{\bullet}$ reduction to HNO.

Nitrosyl hydride has gained considerable interest in recent years for its potential biomedical use.[88] Additionally, HNO exhibits fascinating ionization behavior, since its deprotonation also requires a change in spin-state (HNO: singlet, NO^{-}: triplet).[89] This effect can lead to unusually slow H^{+} transfer kinetics.[90] The mechanism of $NO^{\bullet}$ reduction to HNO by GlbN is unclear, since the direct reduction of $NO^{\bullet}$ to NO^{-} is highly unfavorable ($E^{\circ\prime} \sim -0.8$ V vs. NHE)[91] even when considering the use of dithionite as a strong reducing agent ($E^{\circ\prime} \sim -0.6$ V vs. NHE).[72] The direct observation of the Fe(II)-HNO proton in GlbN near neutral pH (7-8) supports an HNO $pK_a > 8.5$, in accordance with the low acidity of free HNO ($pK_a \sim 11$).[89,91] Overall, the GlbN observations suggest a proton-coupled reduction of $NO^{\bullet}$ to form HNO (Scheme 4.7). If HNO dissociates from GlbN, it may undergo dimerization in solution to form N_2O and H_2O, analogous to the products formed upon $NO^{\bullet}$ reduction in the fungal P450 $NO^{\bullet}$ reductases.[92] Future work could aim to determine the pK_a of HNO when bound to Fe(II) GlbN, and if possible, when bound to the Fe(III) form. Additional

relevant chemistry could include the reactions of HNO and NO^- with singlet and triplet O_2, or other ROS/RNS species such as H_2O_2, $NO^\bullet$ and peroxynitrite.

Conclusion

Although interesting differences exist between GlbN and GlbN-A with regards to their in vitro reactions with $NO^\bullet$, the functional role for His-heme PTM in GlbNs is still a mystery.

Table 4.1. Tentative heme and ligand ^{1}H chemical shifts for *Synechococcus* Fe(III)-NO$^{\bullet}$ GlbN, Fe(III)-NO$^{\bullet}$ GlbN-A and Fe(II)-HNO GlbN (pH 7.1, 298 K)

	Fe^{III}-$NO^{\cdot}$ GlbN	Fe^{II}-HNO GlbN	Fe^{III}-$NO^{\cdot}$ GlbN-A
Assignment	^{1}H (ppm)	^{1}H (ppm)	^{1}H (ppm)
1-CH_3	3.39	3.33	3.38
2-v-α	7.97	7.86	6.81
2-v-β_{cis}	6.15	5.69	n/a
2-v-β_{trans}	6.03	5.80	n/a
2-βCH_3	n/a	n/a	2.30
α-meso	9.80	9.22	9.69
3-CH_3	3.53	3.26	3.50
4-v-α	7.71	7.50	7.69
4-v-β_{cis}	5.94	5.55	5.84
4-v-β_{trans}	6.11	5.82	6.22
β-meso	9.36	8.77	9.46
5-CH_3	3.03	2.62	2.91
γ-meso	10.3	9.57	10.26
8-CH_3	3.33	3.10	3.29
δ-meso	9.96	9.35	9.91
HNO H117A	n/a	14.77 14.85	n/a

n/a: not applicable

Table 4.2. Absorbance maxima for *Synechococcus* GlbN and GlbN-A O_2 and $NO^{\bullet}$ bound forms

	GlbN	GlbN-A
Ligation state	γ,β,α (nm)	γ,β,α (nm)
Fe(II)-O_2	412, 550, 585	408, 550, 585
Fe(III)-$NO^{\bullet}$	425, 538, 574	424, 537, 571
Fe(II)- $NO^{\bullet}$	416?, 563?	415, 558
Fe(II)-HNO	418, 558	n/a

? indicates tentative values due to the contribution of free nitrosyl heme

References

1. Xu, Y. *et al.* Expression of genes in cyanobacteria: adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Methods Mol. Biol.* **684,** 273–293 (2011).
2. Zhu, Y. *et al.* Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Arch. Biochem. Biophys.* **504,** 86–99 (2010).
3. Scott, N. L. *et al.* Truncated hemoglobin from the cyanobacterium *Synechococcus* sp. PCC 7002: evidence for hexacoordination and covalent adduct formation in the ferric recombinant protein. *Biochemistry* **41,** 6902–6910 (2002).
4. Wenke, B. B., Lecomte, J. T. J., Héroux, A. & Schlessman, J. L. 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 (2014).
5. Vu, B. C., Jones, A. D. & Lecomte, J. T. J. Novel histidine-heme covalent linkage in a hemoglobin. *J. Am. Chem. Soc.* **124,** 8544–8545 (2002).
6. Nothnagel, H. J., Love, N. & Lecomte, J. T. J. The role of the heme distal ligand in the post-translational modification of *Synechocystis* hemoglobin. *J. Inorg. Biochem.* **103,** 107–116 (2009).
7. Nothnagel, H. J. *et al.* 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 (2011).

8. Barker, P. D. *et al.* Transmutation of a heme protein. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 6542–6546 (1993).
9. Daltrop, O., Allen, J. W. A., Willis, A. C. & Ferguson, S. J. In vitro formation of a c-type cytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7872–7876 (2002).
10. Daltrop, O. & Ferguson, S. J. In vitro studies on thioether bond formation between Hydrogenobacter thermophilus apocytochrome c(552) with metalloprotoporphyrin derivatives. *J. Biol. Chem.* **279,** 45347–45353 (2004).
11. Preimesberger, M. R., Pond, M. P., Majumdar, A. & Lecomte, J. T. J. 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 (2012).
12. Scott, N. L. *et al.* Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry* **49,** 7000–7011 (2010).
13. Vuletich, D. A., Falzone, C. J. & Lecomte, J. T. J. Structural and dynamic repercussions of heme binding and heme-protein cross-linking in *Synechococcus* sp. PCC 7002 hemoglobin. *Biochemistry* **45,** 14075–14084 (2006).
14. Pond, M. P., Majumdar, A. & Lecomte, J. T. J. Influence of heme post-translational modification and distal ligation on the backbone dynamics of a monomeric hemoglobin. *Biochemistry* **51,** 5733–5747 (2012).
15. Kamiya, K., Yamamoto, S., Shiraishi, K. & Oshiyama, A. Significant change in electronic structures of heme upon reduction by strong Coulomb repulsion between Fe d electrons. *J. Phys. Chem. B* **113,** 6866–6872 (2009).

16. Lin, J. T. & Stewart, V. Nitrate assimilation by bacteria. *Adv. Microb. Physiol.* **39,** 1–30, 379 (1998).
17. Pryor, W. A. & Squadrito, G. L. The chemistry of peroxynitrite: a product from the reaction of nitric oxide with superoxide. *Am. J. Physiol.* **268,** L699–722 (1995).
18. Beckman, J. S., Beckman, T. W., Chen, J., Marshall, P. A. & Freeman, B. A. Apparent hydroxyl radical production by peroxynitrite: implications for endothelial injury from nitric oxide and superoxide. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 1620–1624 (1990).
19. Beckman, J. S. & Koppenol, W. H. Nitric oxide, superoxide, and peroxynitrite: the good, the bad, and ugly. *Am. J. Physiol.* **271,** C1424–1437 (1996).
20. Radi, R. Peroxynitrite, a Stealthy Biological Oxidant. *J. Biol. Chem.* **288,** 26464–26472 (2013).
21. Milani, M., Pesce, A., Ouellet, H., Guertin, M. & Bolognesi, M. Truncated hemoglobins and nitric oxide action. *IUBMB Life* **55,** 623–627 (2003).
22. Su, J. & Groves, J. T. Mechanisms of peroxynitrite interactions with heme proteins. *Inorg. Chem.* **49,** 6317–6329 (2010).
23. De Marinis, E. *et al.* Catalytic peroxidation of nitrogen monoxide and peroxynitrite by globins. *IUBMB Life* **61,** 62–73 (2009).
24. Sadrzadeh, S. M., Graf, E., Panter, S. S., Hallaway, P. E. & Eaton, J. W. Hemoglobin. A biologic fenton reagent. *J. Biol. Chem.* **259,** 14354–14356 (1984).
25. Poulos, T. L. *et al.* The crystal structure of cytochrome c peroxidase. *J. Biol. Chem.* **255,** 575–580 (1980).

26. Poulos, T. L. & Kraut, J. The stereochemistry of peroxidase catalysis. *J. Biol. Chem.* **255,** 8199–8205 (1980).

27. Ortiz de Montellano, P. R. Catalytic sites of hemoprotein peroxidases. *Annu. Rev. Pharmacol. Toxicol.* **32,** 89–107 (1992).

28. Sakamoto, A. *et al.* Three distinct Arabidopsis hemoglobins exhibit peroxidase-like activity and differentially mediate nitrite-dependent protein nitration. *FEBS Lett.* **572,** 27–32 (2004).

29. Hakim, T. S., Sugimori, K., Camporesi, E. M. & Anderson, G. Half-life of nitric oxide in aqueous solutions with and without haemoglobin. *Physiol. Meas.* **17,** 267–277 (1996).

30. Liu, X. *et al.* Diffusion-limited reaction of free nitric oxide with erythrocytes. *J. Biol. Chem.* **273,** 18709–18713 (1998).

31. Traylor, T. G. & Sharma, V. S. Why NO? *Biochemistry* **31,** 2847–2849 (1992).

32. Decatur, S. M. *et al.* Trans effects in nitric oxide binding to myoglobin cavity mutant H93G. *Biochemistry* **35,** 4939–4944 (1996).

33. Gardner, P. R., Gardner, A. M., Martin, L. A. & Salzman, A. L. Nitric oxide dioxygenase: An enzymic function for flavohemoglobin. *Proc. Natl. Acad. Sci.* **95,** 10378–10383 (1998).

34. Gardner, A. M., Martin, L. A., Gardner, P. R., Dou, Y. & Olson, J. S. Steady-state and transient kinetics of Escherichia coli nitric-oxide dioxygenase (flavohemoglobin). The B10 tyrosine hydroxyl is essential for dioxygen binding and catalysis. *J. Biol. Chem.* **275,** 12581–12589 (2000).

35. Gardner, P. R. Nitric oxide dioxygenase function and mechanism of flavohemoglobin, hemoglobin, myoglobin and their associated reductases. *J. Inorg. Biochem.* **99,** 247–266 (2005).
36. Gardner, P. R. Hemoglobin: a nitric-oxide dioxygenase. *Scientifica* **2012,** 683729 (2012).
37. Ascenzi, P., Bolognesi, M., Milani, M., Guertin, M. & Visca, P. Mycobacterial truncated hemoglobins: from genes to functions. *Gene* **398,** 42–51 (2007).
38. Pathania, R., Navani, N. K., Gardner, A. M., Gardner, P. R. & Dikshit, K. L. Nitric oxide scavenging and detoxification by the Mycobacterium tuberculosis haemoglobin, HbN in Escherichia coli. *Mol. Microbiol.* **45,** 1303–1314 (2002).
39. Igarashi, J., Kobayashi, K. & Matsuoka, A. A hydrogen-bonding network formed by the B10-E7-E11 residues of a truncated hemoglobin from Tetrahymena pyriformis is critical for stability of bound oxygen and nitric oxide detoxification. *J. Biol. Inorg. Chem.* **16,** 599–609 (2011).
40. Stewart, J. J. & Coyne, K. J. Analysis of raphidophyte assimilatory nitrate reductase reveals unique domain architecture incorporating a 2/2 hemoglobin. *Plant Mol. Biol.* **77,** 565–575 (2011).
41. Shiva, S. *et al.* Deoxymyoglobin Is a Nitrite Reductase That Generates Nitric Oxide and Regulates Mitochondrial Respiration. *Circ. Res.* **100,** 654–661 (2007).
42. Gladwin, M. T. & Kim-Shapiro, D. B. The functional nitrite reductase activity of the heme-globins. *Blood* **112,** 2636–2647 (2008).

43. Sturms, R., DiSpirito, A. A. & Hargrove, M. S. Plant and Cyanobacterial Hemoglobins Reduce Nitrite to Nitric Oxide under Anoxic Conditions. *Biochemistry* **50,** 3873–3878 (2011).
44. Tiso, M. *et al.* Human neuroglobin functions as a redox-regulated nitrite reductase. *J. Biol. Chem.* **286,** 18277–18289 (2011).
45. Tiso, M., Tejero, J., Kenney, C., Frizzell, S. & Gladwin, M. T. Nitrite Reducatase Activity of Non-Symbiotic hemoglobins from *Arabidopsis thaliana*. *Biochemistry* **51,** (2012).
46. Smagghe, B. J., Trent, J. T. & Hargrove, M. S. NO dioxygenase activity in hemoglobins is ubiquitous in vitro, but limited by reduction in vivo. *PloS One* **3,** e2039 (2008).
47. Ilari, A., Bonamore, A., Farina, A., Johnson, K. A. & Boffi, A. The X-ray structure of ferric Escherichia coli flavohemoglobin reveals an unexpected geometry of the distal heme pocket. *J. Biol. Chem.* **277,** 23725–23732 (2002).
48. Bonamore, A. & Boffi, A. Flavohemoglobin: structure and reactivity. *IUBMB Life* **60,** 19–28 (2008).
49. Weiland, T. R., Kundu, S., Trent, J. T., Hoy, J. A. & Hargrove, M. S. Bis-histidyl hexacoordination in hemoglobins facilitates heme reduction kinetics. *J. Am. Chem. Soc.* **126,** 11930–11935 (2004).
50. Simonneaux, G. & Bondon, A. Mechanism of Electron Transfer in Heme Proteins and Models: The NMR Approach. *Chem. Rev.* **105,** 2627–2646 (2005).

51. Dixon, D. W., Hong, X. & Woehler, S. E. Electrostatic and steric control of electron self-exchange in cytochromes c, c551, and b5. *Biophys. J.* **56,** 339–351 (1989).
52. Fernandez, B. O. & Ford, P. C. Nitrite catalyzes ferriheme protein reductive nitrosylation. *J. Am. Chem. Soc.* **125,** 10510–10511 (2003).
53. Fernandez, B. O., Lorkovic, I. M. & Ford, P. C. Mechanisms of ferriheme reduction by nitric oxide: nitrite and general base catalysis. *Inorg. Chem.* **43,** 5393–5402 (2004).
54. Addison, A. W. & Stephanos, J. J. Nitrosyliron(III) hemoglobin: autoreduction and spectroscopy. *Biochemistry* **25,** 4104–4113 (1986).
55. Rassaf, T. *et al.* Nitrite Reductase Function of Deoxymyoglobin Oxygen Sensor and Regulator of Cardiac Energetics and Function. *Circ. Res.* **100,** 1749–1754 (2007).
56. Ascenzi, P. *et al.* Nitrite-reductase and peroxynitrite isomerization activities of Methanosarcina acetivorans protoglobin. *PloS One* **9,** e95391 (2014).
57. He, C. & Knipp, M. Formation of nitric oxide from nitrite by the ferriheme b protein nitrophorin 7. *J. Am. Chem. Soc.* **131,** 12042–12043 (2009).
58. Englander, S. W., Calhoun, D. B. & Englander, J. J. Biochemistry without oxygen. *Anal. Biochem.* **161,** 300–306 (1987).
59. Hayashi, A., Suzuki, T. & Shin, M. An enzymic reduction system for metmyoglobin and methemoglobin, and its application to functional studies of oxygen carriers. *Biochim. Biophys. Acta* **310,** 309–316 (1973).
60. Hrabie, J. A., Klose, J. R., Wink, D. A. & Keefer, L. K. New nitric oxide-releasing zwitterions derived from polyamines. *J. Org. Chem.* **58,** 1472–1476 (1993).

61. Keefer, L. K., Nims, R. W., Davies, K. M. & Wink, D. A. in *Methods Enzymol.* (ed. Lester Packer) **268,** 281–293 (Academic Press, 1996).
62. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 (1995).
63. Goddard, T. D. & Kneller, D. G. *SPARKY 3*.
64. Hvitved, A. N., Trent, J. T., Premer, S. A. & Hargrove, M. S. Ligand binding and hexacoordination in *synechocystis* hemoglobin. *J. Biol. Chem.* **276,** 34714–34721 (2001).
65. Couture, M. *et al.* Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC6803 reveal a unique distal heme pocket. *Eur. J. Biochem. FEBS* **267,** 4770–4780 (2000).
66. Falzone, C. J., Christie Vu, B., Scott, N. L. & Lecomte, J. T. J. The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state. *J. Mol. Biol.* **324,** 1015–1029 (2002).
67. Trent, J. T., Kundu, S., Hoy, J. A. & Hargrove, M. S. Crystallographic analysis of *synechocystis* cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. *J. Mol. Biol.* **341,** 1097–1108 (2004).
68. Hoy, J. A., Kundu, S., Trent, J. T., Ramaswamy, S. & Hargrove, M. S. The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. *J. Biol. Chem.* **279,** 16535–16542 (2004).
69. Pesce, A. *et al.* A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO J.* **19,** 2424–2434 (2000).

70. Milani, M. *et al.* Cyanide binding to truncated hemoglobins: a crystallographic and kinetic study. *Biochemistry* **43,** 5213–5221 (2004).

71. Ignarro, L. J., Fukuto, J. M., Griscavage, J. M., Rogers, N. E. & Byrns, R. E. Oxidation of nitric oxide in aqueous solution to nitrite but not nitrate: comparison with enzymatically formed nitric oxide from L-arginine. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 8103–8107 (1993).

72. Mayhew, S. G. The Redox Potential of Dithionite and SO−2 from Equilibrium Reactions with Flavodoxins, Methyl Viologen and Hydrogen plus Hydrogenase. *Eur. J. Biochem.* **85,** 535–547 (1978).

73. Kharitonov, V. G., Sharma, V. S., Magde, D. & Koesling, D. Kinetics of nitric oxide dissociation from five- and six-coordinate nitrosyl hemes and heme proteins, including soluble guanylate cyclase. *Biochemistry* **36,** 6814–6818 (1997).

74. Tsai, A.-L., Berka, V., Martin, E. & Olson, J. S. A 'sliding scale rule' for selectivity among NO, CO, and O_2 by heme protein sensors. *Biochemistry* **51,** 172–186 (2012).

75. Poulos, T. L. Soluble guanylate cyclase. *Curr. Opin. Struct. Biol.* **16,** 736–743 (2006).

76. Lin, R. & Farmer, P. J. The HNO Adduct of Myoglobin: Synthesis and Characterization. *J. Am. Chem. Soc.* **122,** 2393–2394 (2000).

77. Sulc, F., Immoos, C. E., Pervitsky, D. & Farmer, P. J. Efficient trapping of HNO by deoxymyoglobin. *J. Am. Chem. Soc.* **126,** 1096–1101 (2004).

78. Kumar, M. R. *et al.* Nitrosyl hydride (HNO) as an O2 analogue: long-lived HNO adducts of ferrous globins. *Biochemistry (Mosc.)* **48,** 5018–5025 (2009).

79. Sulc, F., Fleischer, E., Farmer, P. J., Ma, D. & La Mar, G. N. 1H NMR structure of the heme pocket of HNO-myoglobin. *J. Biol. Inorg. Chem.* **8,** 348–352 (2003).

80. Li, B.-R. *et al.* Rhodobacter sphaeroides haem protein: a novel cytochrome with nitric oxide dioxygenase activity. *Biochem. Soc. Trans.* **36,** 992–995 (2008).

81. Johnson, E. A. *et al.* 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 (2014).

82. Huang, L., Wojciechowski, G. & Ortiz de Montellano, P. R. Role of heme-protein covalent bonds in mammalian peroxidases. Protection of the heme by a single engineered heme-protein link in horseradish peroxidase. *J. Biol. Chem.* **281,** 18983–18988 (2006).

83. Reiter, C. D., Teng, R. J. & Beckman, J. S. Superoxide reacts with nitric oxide to nitrate tyrosine at physiological pH via peroxynitrite. *J. Biol. Chem.* **275,** 32460–32466 (2000).

84. Sawa, T., Akaike, T. & Maeda, H. Tyrosine nitration by peroxynitrite formed from nitric oxide and superoxide generated by xanthine oxidase. *J. Biol. Chem.* **275,** 32467–32474 (2000).

85. Brucker, E. A., Olson, J. S., Ikeda-Saito, M. & Phillips, G. N. Nitric oxide myoglobin: crystal structure and analysis of ligand geometry. *Proteins* **30,** 352–356 (1998).

86. Copeland, D. M., West, A. H. & Richter-Addo, G. B. Crystal structures of ferrous horse heart myoglobin complexed with nitric oxide and nitrosoethane. *Proteins Struct. Funct. Bioinforma.* **53,** 182–192 (2003).
87. Hoy, J. A., Smagghe, B. J., Halder, P. & Hargrove, M. S. Covalent heme attachment in Synechocystis hemoglobin is required to prevent ferrous heme dissociation. *Protein Sci. Publ. Protein Soc.* **16,** 250–260 (2007).
88. Fukuto, J. M. *et al.* The Physiological Chemistry and Biological Activity of Nitroxyl (HNO): The Neglected, Misunderstood, and Enigmatic Nitrogen Oxide. *Chem. Res. Toxicol.* **18,** 790–801 (2005).
89. Shafirovich, V. & Lymar, S. V. Nitroxyl and its anion in aqueous solutions: spin states, protic equilibria, and reactivities toward oxygen and nitric oxide. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7340–7345 (2002).
90. Shafirovich, V. & Lymar, S. V. Spin-Forbidden Deprotonation of Aqueous Nitroxyl (HNO). *J. Am. Chem. Soc.* **125,** 6547–6552 (2003).
91. Bartberger, M. D. *et al.* The reduction potential of nitric oxide (NO) and its importance to NO biochemistry. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 10958–10963 (2002).
92. Daiber, A., Shoun, H. & Ullrich, V. Nitric oxide reductase (P450nor) from Fusarium oxysporum. *J. Inorg. Biochem.* **99,** 185–193 (2005)

Chapter 5.

Facile heme vinyl posttranslational modification in a hemoglobin

Matthew R. Preimesberger, Belinda B. Wenke, Lukas Gilevicius, Matthew P. Pond, and Juliette T.J. Lecomte

Reproduced with permission from Biochemistry (2013) 52 (20):3478–3488

DOI: 10.1021/bi400289e Publication Date (Web): April 22, 2013

Author contributions:

MRP: designed research, assisted with protein preparation, performed NMR data acquisition and analysis, performed and analyzed UV-visible experiments and wrote the paper

BBW: prepared protein samples

LG: assisted with protein preparation, performed UV-visible experiments

MPP: wrote Python program to parse PDB for specific heme-protein distances

JTJL: designed research, analyzed data, wrote the paper

Abbreviations: DT, sodium dithionite; GlbN, *Synechocystis* sp. PCC 6803 hemoglobin without heme posttranslational modification; GlbN-A, GlbN with a bond between His117 Nε2 and heme 2-Cα; GlbN-B, GlbN with a bond between His79 Nε2 and heme 4-Cα; GlbN-AB, GlbN with a bond between His117 Nε2 and heme 2-Cα and a bond between His79 Nε2 and heme 4-Cα; lr, long-range; Mb, myoglobin.

Abstract

Iron-protoporphyrin IX, or *b* heme, is utilized as such by a large number of proteins and enzymes. In some cases, notably the *c*-type cytochromes, this group undergoes a posttranslational covalent attachment to the polypeptide chain adjusting the physico-chemical properties of the holoprotein. The hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 (GlbN), contrary to the archetypical hemoglobin, modifies its *b* heme covalently. The posttranslational modification links His117, a residue that does not coordinate the iron, to the porphyrin 2-vinyl substituent and forms a hybrid *b*/*c* heme. The reaction is an electrophilic addition that occurs spontaneously in the ferrous state of the protein. This apparently facile type of heme modification has been observed only in two cyanobacterial GlbNs. To explore the determinants of the reaction, we examined the behavior of *Synechocystis* GlbN variants containing a histidine at position 79, which is buried against the porphyrin 4-vinyl substituent. We found that L79H/H117A GlbN bound the heme weakly, but nevertheless formed a cross-link between His79 Nε2 and the heme 4-Cα. In addition to this linkage, the single variant L79H GlbN also formed the native His117–2-vinyl bond yielding an unprecedented *bis*-alkylated protein adduct. The ability to engineer the doubly modified protein indicates that the histidine–heme modification in GlbN is robust and could be engineered in different local environments. The rarity of the histidine linkage in natural proteins, despite the ease of reaction, is proposed to stem from multiple sources of negative selection.

Introduction

Hemoproteins are remarkable for both the chemistry that they perform and the chemistry that they avoid. The exquisite control exerted by the polypeptide chain on the high intrinsic reactivity of ferrous iron-protoporphyrin IX (*b* heme, Figure 5.1A) is well known and is often illustrated by contrasting proteins that favor reversible ligand binding (e.g., hemoglobin), O_2 activation and O transfer (e.g., cytochrome P450), or electron transport (e.g., cytochrome b_5).[1] The nature of the axial ligands to the iron, most commonly, histidine, methionine, and cysteine, is a primary determinant of essential heme properties such as reduction potential and capacity for electron transfer. Reactivity also extends to and is modulated by covalent attachment of the heme to the protein matrix, a characteristic of *c*-type cytochromes[2,3] and many mammalian peroxidases.[4,5]

Multiple functional roles are adopted by members of the hemoglobin superfamily;[6] as a result, these proteins are well suited for a focused inspection of heme chemistry diversification over evolutionary time. Known hemoglobins all utilize an unmodified *b* heme. Exceptions to this rule are found in the Group I 2/2 hemoglobins from the cyanobacteria *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. Specifically, these two proteins (GlbNs) undergo a spontaneous posttranslational modification (PTM) by which a non-coordinating histidine near the C-terminus (His117) reacts irreversibly with the heme 2-vinyl substituent to form a His117 Nε2–Cα linkage (Figure 5.1B).[7]

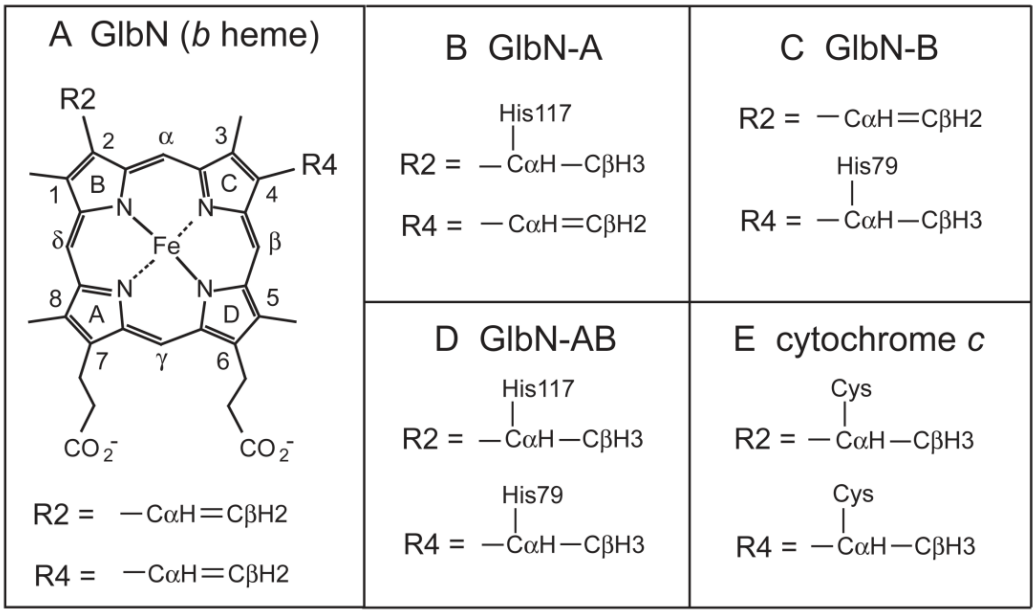


Figure 5.1. Chemical structure of *b* heme and covalently modified hemes discussed in the text. A, the *b* heme has vinyl groups at the 2 and 4 positions. Protoporphyrin IX has a pseudo C_{2v} axis of symmetry passing through the α–γ meso positions. Heme orientational isomerism arises from a 180° rotation about this axis. B, the heme in GlbN-A has a bond between His117 Nε2 and the 2-Cα. C, the heme in GlbN-B has a bond between His79 Nε2 and the heme 4-Cα. D, the novel heme in GlbN-AB has the linkages found in both GlbN-A and GlbN-B. E, the *c* heme in *c*-type cytochromes has thioether linkages at the 2-Cα and 4-Cα positions through conserved cysteines.

In *Synechococcus*, there is evidence that the linkage is generated in vivo under microoxic growth conditions,[8] and the expectation is that the posttranslationally modified protein is functional in *Synechococcus* and *Synechocystis*. According to the current view, both GlbNs have a physiological role in reactive nitrogen species stress response.[8,9]

The most common heme–protein covalent linkage is the thioether bond found in *c*-type cytochromes,[2,3] where cysteines (generally two, occasionally one)

play a role similar to that of the reactive histidine of GlbN (Figure 5.1E). In multiheme *c*-type cytochromes, covalent heme attachment is thought to allow for dense packing of the prosthetic groups and their controlled orientation.[10] In GlbN and mono-heme *c*-type cytochromes, however, the necessity and consequences of the heme–protein linkage are not well understood.[8,11-14] Among proposed possibilities for GlbN are stabilization of the holoprotein[15,16] and heme sequestration.[17] The attachment of the cofactor may also lower the constraints on protein stability during evolutionary selection and allow for a broad tuning of functional properties such as electron transfer and reduction potential.[12,18] Modification of redox properties is apparent in other heme PTMs[5] and can contribute in GlbN as well.[16,19] A characterization of the chemistry underlying the formation of the GlbN linkage and its roles will therefore have broad implications for describing the evolution of heme proteins.

Previously, we have exposed several features of the in vitro PTM mechanism (Scheme S5.1 of the Supporting Information).[19,20] The reaction occurs in the ferrous state and is initiated by protonation of the heme 2-vinyl Cβ by His117. Rapid nucleophilic attack of the heme 2-vinyl Cα by the now neutral histidine follows, generating an adduct with R stereochemistry at the Cα. In the past, we explored the reactivity determinants that may be specific to GlbN. For example, in the absence of exogenous ligands, GlbN His46 (distal) and His70 (proximal) are axial ligands to the heme iron, but replacement of the distal histidine (His46) with a non-coordinating residue does not prevent the modification.[21] In addition, reaction in wild-type GlbN and its distal variants occurs when cyanide is bound to the ferric state prior to reduction. His46 replacement and cyanide binding are perturbations that permit formation of a product in which the

minor heme orientational isomer (related to the major isomer by a 180° flip about the α–γ meso axis, Figure 5.1A) reacts to generate a His117 adduct to the vinyl group on pyrrole C (4-vinyl).[20,21] Thus, the modification displays a degree of insensitivity to heme packing geometries and distal ligation.

The documented mechanistic features and relative insensitivity of reaction to amino acid replacement suggest that it should be possible to engineer the histidine–vinyl modification in different protein contexts. Achieving this goal would not only generate a deeper understanding of the PTM, but also offer guidance in the design of stable proteins containing a non-displaceable heme. We begin an exploration of the modification by focusing on *Synechocystis* sp. PCC 6803 GlbN and probing the 4-vinyl reactivity with an engineered histidine. As we will show, the successful introduction of the histidine–heme linkage provides insight into heme chemistry and prompts a discussion of the rarity of the linkage in natural heme proteins.

Materials and Methods

Materials. Horse skeletal myoglobin (Mb, >95% pure), horse heart cytochrome *c* (97%), sodium hydrosulfite (sodium dithionite, DT ~85%), monobasic potassium phosphate, D_2O (99.9% ^2H), and $^{15}NH_4Cl$ (98% ^{15}N) were purchased from Sigma. Bovine hemin chloride (98%) and 2-butanone were obtained from Alfa-Aesar. Pyridine, dibasic sodium phosphate, and Tris-base were obtained from Fischer, urea and EDTA from EM Science, and $^{14}NH_4Cl$ from VWR. The chromatographic materials used to purify GlbNs were: G-50 (fine)

size-exclusion resin from Amersham BioScience, DEAE anion exchange resin from GE BioScience, and G-25 Sephadex desalting material from Sigma. Amicon Ultra (Millipore, MWCO 10 kDa) centrifugal filter units were used for protein concentration and buffer exchange.

Protein production and purification. A pET3c plasmid encoding H117A or wild-type *Synechocystis* GlbN was mutated by the QuikChange Site-Directed Mutagenesis PCR method (Qiagen, Valencia, CA) using primers purchased from IDT (Coralville, IA). Following Dpn1 digestion of the parent vector, the PCR product was used to transform DH5α *E. coli* cells for plasmid preparation. Mutant pET3c plasmids encoding the single L79H or double L79H/H117A replacements in GlbN were purified using the Qiaprep spin miniprep kit (Qiagen, Valencia, CA) and positively confirmed by sequencing of the mutant genes (GENEWIZ, Inc. South Plainfield, NJ). These plasmids were used to transform BL21(DE3) *E. coli* for protein production.

Cells were grown in M9 minimal medium containing either $^{14}NH_4Cl$ or $^{15}NH_4Cl$ as the sole nitrogen source. Variant GlbN overexpression and purification were accomplished as previously described for wild-type GlbN[20,22] with minor modifications. Because both L79H GlbN variants exhibit significantly attenuated heme affinity (vide infra), the heme addition step was completed last in order to avoid heme loss during DEAE anion exchange purification. Additionally, a higher strength salt gradient (0–1 M, NaCl) was used to elute variant apoproteins from the anion exchange column. Finally, purified (> 95% by SDS-PAGE) variant GlbNs were exchanged from chromatography buffer (50 mM Tris, 5 mM EDTA, pH 8.2) into storage buffer

(~1 mM phosphate, pH ~7.5). Proteins prepared in this manner were either immediately examined for histidine–heme reactivity or lyophilized for long term storage.

The UV-visible spectra of freshly reconstituted ferric L79H/H117A and L79H GlbNs are presented in Figure S5.1 of the Supporting Information. The Soret maxima are located at 413 nm (L79H/H117A GlbN) and 414 nm (L79H GlbN). Both proteins exhibit Q-bands at ~539 nm (~563 nm shoulder). These features are similar to wild-type GlbN (Soret band: 410 nm, Q bands: ~546 nm) and are consistent with low-spin *bis*-histidine complexes at pH 7.1 and room temperature. Both L79H GlbN variants reconstituted using this method displayed $Abs_{(Soret)}/Abs_{(280\ nm)}$ ratios of approximately 4–5 (Figure S5.1 of the Supporting Information), similar to wild-type GlbN preparations. Concentrations were estimated optically with wild-type coefficients of ε_{278} = 7.4 $mM^{-1}cm^{-1}$ for the apoprotein [22] and ε_{410} = 100 $mM^{-1}cm^{-1}$ for unreacted ferric GlbNs[23].

L79H/H117A and L79H GlbN reduction reactions. Histidine addition to a vinyl group causes a small blue shift in the reduced spectrum that can be monitored optically. Concentrated stock GlbN solutions were diluted into 100 mM phosphate buffer, pH 7.2, to yield ~10 μM GlbN, and reduction reactions were initiated by addition of freshly prepared DT (~2 mM final concentration). Manual mixing dead times were 15–20 s. For NMR investigation, concentrated (~1 mM) GlbN samples were treated with five-fold molar excess DT for 2–3 h, followed by $K_3[Fe(CN)_6]$ oxidation to the ferric state, and G-25 desalting into the desired buffer. In some cases, centrifugal filter units were used to re-

concentrate and further exchange the sample into D_2O. Final protein NMR sample concentrations ranged from 1–5 mM. Apomyoglobin (apoMb) was prepared by heme extraction from the acidified horse skeletal Mb (Sigma) with cold butanone,[24] followed by extensive dialysis of the soluble apoprotein against 1 mM phosphate. The amount of remaining Mb was estimated to be < 5% based on UV-visible absorbance.

Pyridine hemochromogen assay. As a rapid test for heme modifications, the hemochromogen assay[25] was performed as previously reported.[20] DT-reacted samples were typically reduced for 1–2 h to allow for complete conversion prior to assaying. The results were interpreted by comparison with published spectra (α and β band maxima).[26,27]

Optical spectroscopy. Absorbance spectra were collected on a Varian Cary50 for monitoring the kinetics of heme modification reactions. Spectra were collected from 640 to 350 nm, in 1-nm steps, using a 0.1-s averaging time. The hemochromogen assay results were read on an Aviv 14-DS spectrophotometer, equipped with a Peltier device set to 25 °C, from 600 to 500 nm, in 0.5 nm steps (0.5-nm slit width), using an averaging time of 1 s. The presented pyridine hemochrome spectra are the average of three consecutive scans. The same Aviv 14-DS instrument was used for thermal heme loss experiments. Temperature-dependent spectra were completed in triplicate and collected from 700 to 280 nm, in 1-nm steps (1-nm slit width), using a 0.2-s averaging time. Spectra were recorded on ~10 μM GlbN samples, in 250 mM degassed phosphate buffer pH 7.3, from 25 to 91 °C, in 3-°C steps, each following a 5-min thermal equilibration period. Because all samples exhibit

irreversible transitions, variable temperature data were interpreted qualitatively.

NMR spectroscopy. All data were collected at 600 MHz on Bruker AVANCE or AVANCE-II spectrometers equipped with cryoprobes. Probe temperature was calibrated using methanol.[28] NMR experiments used to establish the heme pocket environment and covalent structure of the variant heme modifications were as previously reported[15,20] and included ^{1}H 1-D, ^{1}H-^{1}H DQF-COSY, ^{1}H-^{1}H TOWNY TOCSY,[29] water presaturation ^{1}H-^{1}H NOESY, water-elimination Fourier transform (WEFT) ^{1}H-^{1}H NOESY,[30] ^{1}H-^{15}N HSQC, ^{1}H-^{15}N long range (lr) HMQC,[31] natural abundance ^{1}H-^{13}C HMQC and natural abundance band-selective ^{1}H-^{13}C HSQC with or without ^{1}H decoupling. Data were processed using either TopSpin 2.1 (Bruker BioSpin, Rheinstetten, Germany) or NMRPipe.[32] Processed spectra were analyzed using the program Sparky 3.[33] ^{1}H chemical shifts were referenced to DSS through the temperature-corrected water line (4.76 ppm at 298 K); ^{13}C and ^{15}N chemical shifts were referenced indirectly using Ξ ratio.[34]

Database analysis. A list of PDB files[35] was obtained from the RCSB site[36] using HEM as a ligand keyword. The 95% similarity filter option was applied, and proteins containing a thioether bridge to a HEM ligand were discarded. If an artificial variant was selected in the set, it was removed and replaced by the corresponding wild-type structure. If a wild-type structure was not available, the protein was eliminated from the set. The resultant dataset was composed of 341 structures, with PDB ID listed in Table S5.3 of the Supporting Information. The set contained 28% globins, 23% cytochromes P450, and a variety of other

cytochromes, catalases, peroxidases, etc. The BioPython module[37] was used to identify non-coordinating residues with a Cα within 7 Å and Cβ within 6 Å of a HEM vinyl Cα or propionate Cα and to count the total number of residues. The propensity P_i^v of amino acid of type i to satisfy the vinyl (v) distance criteria was estimated using

$$P_i^v = \frac{N_i^v / \sum_j N_j^v}{N_i / \sum_j N_j}$$

where N_i^v represents the number of amino acids of type i identified near a vinyl in the set, and N_i represents the total number of amino acids of type i in the heme domains of the set. The index j covers all amino acids.

Results

The crystal structure of *Synechocystis* sp. PCC 6803 GlbN with heme covalently attached (GlbN-A, PDB ID: 1RTX,[38] Figure S5.2 of the Supporting Information) was used to identify a position appropriate for the placement of a histidine for reaction with the heme 4-vinyl. Of the few residues located in the proximity of the 4-vinyl substituent, semi-conserved Leu79 was selected because its orientation is analogous to that of His117 with respect to the 2-vinyl (Figure 5.2). Furthermore, position 79 resides in the FG loop, which may tolerate some extent of backbone relaxation.

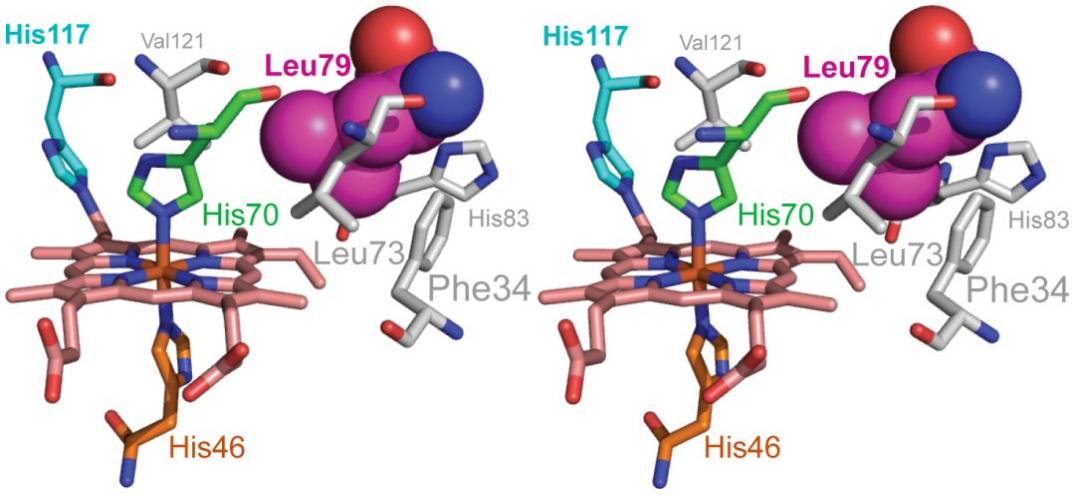


Figure 5.2. Stereo view (wall-eyed) of the proximal GlbN heme pocket (PDB ID: 1RTX). Leu79 is shown in purple spheres. The position of Leu79 with respect to the heme 4-vinyl group is similar to that of His117 (cyan sticks) with respect to the reacted 2-vinyl group. The side chains of Leu73, Phe34, His83 and Val121 (grey sticks) pack near Leu79. The axial iron ligands are shown as green sticks (proximal, His70) and orange sticks (distal, His46).

Properties of L79H/H117A GlbN. The apoprotein of GlbN partitions primarily into inclusion bodies. HoloGlbN preparation therefore includes a heme titration step, typically performed before purification by anion exchange chromatography.[22] When this protocol was applied to the variant, most of the heme dissociated from the apoprotein during passage through the column. According to optical data, increasing or decreasing the pH from neutrality led to the increase of high-spin heme characteristics (shoulder at ~375 nm, charge transfer band at ~620 nm, data not shown). Because the H117A replacement has no such consequences on its own,[39] we concluded that ferric heme affinity was significantly compromised by the L79H replacement. To circumvent this issue, heme was added in stoichiometric amount to the apoprotein after the chromatographic steps.

The resistance of L79H/H117A GlbN to thermally-induced heme loss was monitored by absorbance spectroscopy (Figure S5.3 of the Supporting Information). As the temperature is raised above 25 °C, a charge-transfer band emerges at ~610 nm providing evidence for deligation of His46 and the population of a high-spin complex. The Soret band decreases gradually, and eventually the spectrum changes into that of free ferric heme. Under the conditions assayed (~10 μM protein, pH 7.3, 250 mM phosphate buffer) the thermal transition is irreversible, and thermodynamic parameters characterizing the process were not obtained. Heme loss occurs with low cooperativity and an apparent midpoint of ~52 °C, ~20 °C lower than wild-type GlbN.[22]

^{1}H NMR spectroscopy was used to characterize ferric L79H/H117A GlbN prior to treatment with the reducing agent. The spectrum, shown in Figure 5.3C, can be compared to that of wild-type GlbN without histidine–heme PTM (Figure 5.3A).

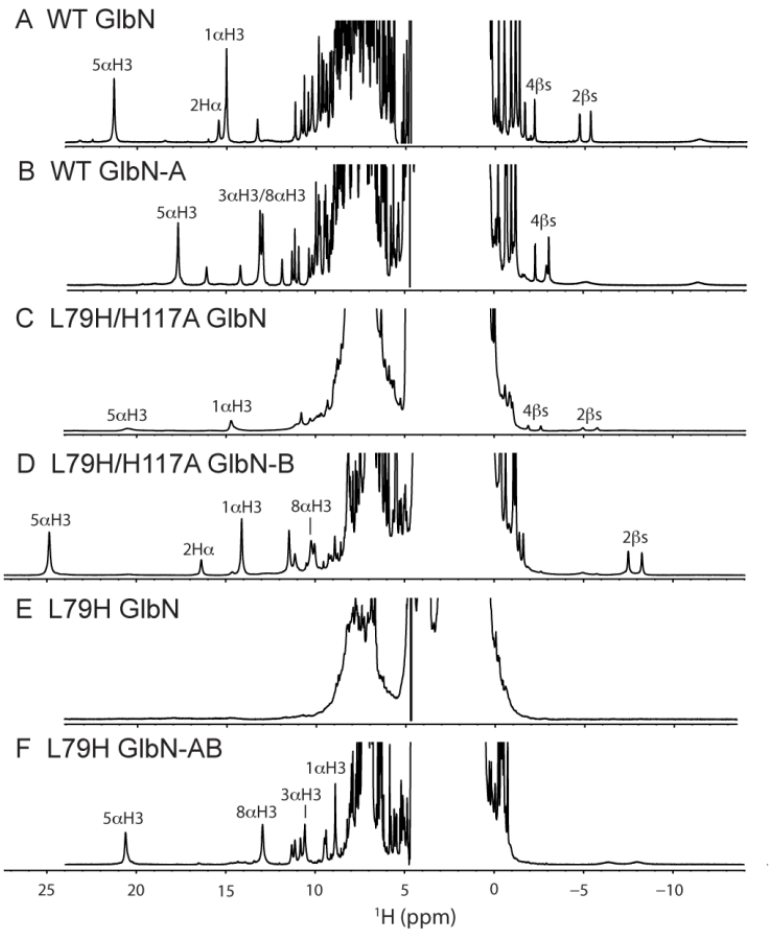


Figure 5.3. ^{1}H NMR spectra of various ferric GlbNs. A, wild-type GlbN and B, wild-type GlbN-A are included for reference (pH 7.2, 298 K). C, L79H/H117A GlbN (pH 7.3, 298 K). D, L79H/H117A GlbN-B (pH* 7.0, 298 K, 99.9 % D_2O). E, L79H GlbN (pH 7.3, 298 K). F, L79H GlbN-AB (pH* 7.61, 313 K, 99.9 % D_2O). The complexes are paramagnetic (S = 1/2) with *bis*-histidine iron coordination; in C and E a contribution of high spin (S = 5/2) and apoprotein species is likely. Selected heme assignments are as indicated (see Tables S1 and S2 of the Supporting Information for chemical shift values).

Although the chemical shifts of resolved resonances are similar, L79H/H117A GlbN exhibits much broader lines than the wild-type protein across the whole spectrum. The excess line width is attributed to exchange among various paramagnetic species: partially folded apoprotein with non-

specifically bound heme and *bis*-histidine holoprotein. NMR data detect the largest population of low-spin species at neutral pH in agreement with the optical data. The study of ferric L79H/H117A GlbN was not pursued owing to its inferior NMR spectral properties.

Heme modification in L79H/H117A GlbN monitored optically. The response of L79H/H117A GlbN to reduction with DT was monitored as a function of time by absorption spectroscopy in the visible range (Figure 5.4) as was done for the wild-type protein.[20] In a first set of experiments, ~10 μM L79H/H117A GlbN was treated with a 200-fold molar excess DT in 100 mM phosphate (pH 7.1), and spectra were collected from 350 to 640 nm for several hours. Reduction, which is essentially complete during the manual mixing dead time (~15 s), is manifested in the Soret maximum of approximately 426 nm, α band at 560 and β band at 531 nm (Figure 5.4A, red trace). Interestingly, further spectral changes are observed, and after one hour, the sample exhibits Soret, α, and β bands at 424, 558, and 528 nm, respectively, all blue-shifted compared to the initial ferrous state (Figure 5.4A, cyan trace). Also of note is the initial decrease in Soret intensity, because reduced *bis*-histidine GlbNs generally have larger α, β and Soret extinction coefficients than in the corresponding oxidized state.

A second set of experiments was performed to identify the nature of the kinetic phases and the species involved. The L79H/H117A variant was again treated with 200-fold molar excess DT and mixed manually for ~15 s, but in this case the reduction step was immediately followed by addition of a two-fold molar excess horse skeletal apomyoglobin (apoMb) as a heme scavenger.

Within seconds of apoMb addition, an absorption band characteristic of deoxyMb[27] appears in the spectrum (Figure 5.4B), and no obvious spectral component corresponding to ferrous GlbN remains. This demonstrates that the fast reduction phase results in a protein containing a non-covalently attached heme, and that heme dissociation from this reduced state occurs rapidly.

In contrast, when L79H/H117A GlbN was incubated with reducing agent for 1 h to produce the species with a blue-shifted optical spectrum (Figure 5.4A, cyan trace), oxidized, and then re-treated with excess DT, reduction led rapidly and directly to the blue-shifted spectrum. Addition of apoMb at this stage does not perturb the spectrum and confirms that the heme is no longer transferrable (Figure 5.4C). These observations are consistent with the fast kinetic phase corresponding to reduction of the *b* heme and the slow phase reflecting its irreversible attachment to the protein.

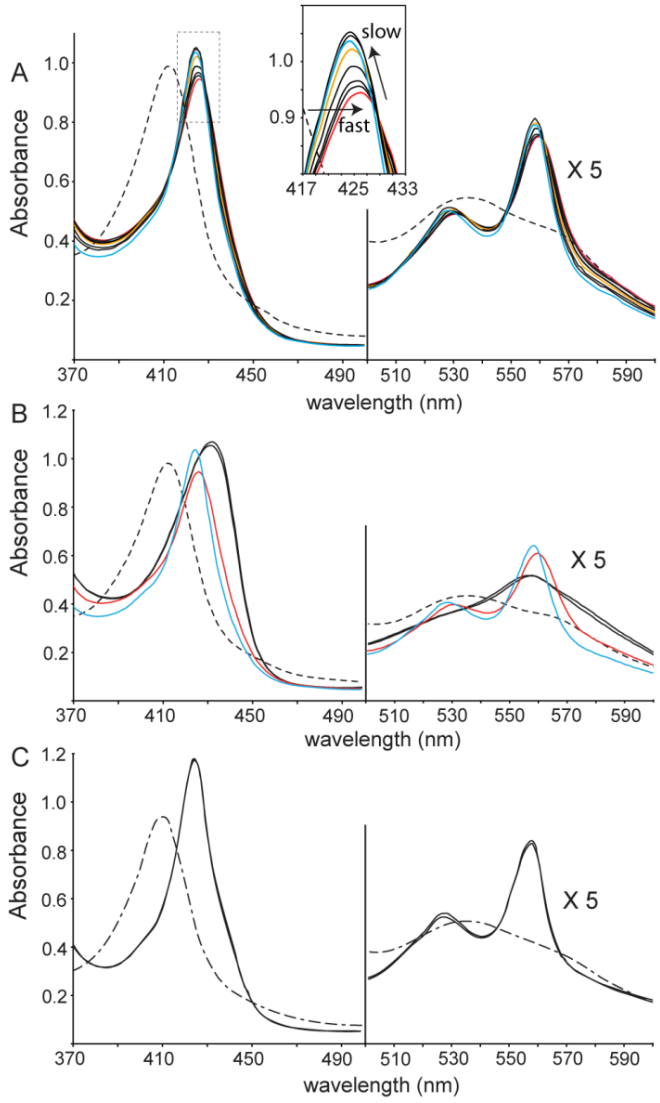


Figure 5.4. Time dependence of the absorbance spectrum of L79H/H117A GlbN after reduction with DT. Data were collected in the presence and absence of the heme scavenger horse skeletal apoMb. Conditions were ~10 μM GlbN, 100 mM phosphate, pH ~7.1, 2 mM DT, room temperature, dead time: 15–20 s. A, ferric GlbN (dashed line); intermediate ferrous form appearing in the dead time (red line); spectrum of the product generated after 10 min (orange line) and 1 h (cyan line). The inset magnifies the Soret region. B, the same experiment as in A, except that 20 μM apoMb was mixed in immediately following reduction by DT. The initial spectrum (dashed line) converted rapidly to that of deoxyMb (spectra shown after 15 s and 10 min, black solid lines). C, as in B, but with ferric GlbN-B as starting material (dash-dotted line). No heme transfer to apoMb was observed (15 s and 10 min spectra shown, black solid lines).

To inspect the integrity of the heme group, we first used the pyridine hemochromogen assay.[25] This method eliminates the influence of tertiary structure on optical properties by denaturing the protein in basic pyridine. Subsequent reduction yields a *bis*-pyridine ferrous heme complex with absorption maxima characteristic of various heme modifications. Figure S5.4A of the Supporting Information presents the pyridine hemochrome spectra of untreated and DT-reacted L79H/H117A GlbN. As expected, the untreated variant yields the spectrum of *bis*-pyridine ferrous protoporphyrin IX (α and β band maxima at 557 and 525 nm, respectively). The reacted variant displays blue-shifted bands (553 and 523 nm), within error of those observed for wild-type GlbN-A[40] and confirming the saturation of one vinyl group. Henceforth, the product of the DT treatment is denoted as L79H/H117A GlbN-B.

Characterization of ferric GlbN-B by NMR spectroscopy. Several adducts are a priori consistent with the results above, involving the vinyl Cα as in GlbN-A or Cβ as in the cytochrome *c* maturation protein CcmE,[41] and histidine Nε2 (GlbN-A) or Nδ1 (CcmE). Other heme modifications, caused by side reactions with DT byproducts, are also possible. Homo- and heteronuclear NMR data were collected to determine the nature of the heme–protein linkage.

The ^{1}H NMR spectrum of ferric L79H/H117A GlbN-B is shown in Figure 5.3D. Compared to the starting material (Figure 5.3C), the distinct chemical shifts confirm that a reaction has gone to completion. In addition, the resolved lines in GlbN-B are sharp and their chemical shift is consistent with a well-folded, low-spin, *bis*-histidine complex such as GlbN-A (Figure 5.3B). The small

number of paramagnetically shifted lines also suggest that a single product has been obtained.

In addition to the engineered His79 in the FG turn, L79H/H117A GlbN contains five histidines: the proximal (His70) and distal (His46) axial ligands, His33, His77, and His83 (Figure S5.2 of the Supporting Information). To examine these side chains, we used a uniformly ^{15}N-labeled sample of ferric L79H/H117A GlbN-B and collected ^{15}N band-selective ^{1}H-^{15}N long-range (lr) HMQC spectra (Figure 5A). The assignments of Hδ2, Hε1, ^{15}Nδ1 and ^{15}Nε2 in His83, His33, and His77 are readily transferred from wild-type GlbN-A.[7] The ^{1}H-^{15}N lr-HMQC spectrum exhibits an unusual pair of cross peaks indicating that an aromatic nitrogen (^{15}N at 205.5 ppm) is covalently connected to two sets of protons resonating upfield, at 2.88 and −1.04 ppm (Figure 5A). Because signals from the iron axial ligands (His46 and His70) are efficiently relaxed by paramagnetism and not detected in this experiment,[7] the unusual spin system is attributed to a moiety either part of a modified His79 or connected to it.

^{1}H–^{1}H NOESY, ^{1}H–^{1}H DQF-COSY, ^{1}H–^{1}H TOCSY, and natural abundance ^{1}H–^{13}C HMQC spectra were collected to assign the heme ^{1}H and ^{13}C resonances. Four heme CαH$_3$ signals (Figure S5.5 of the Supporting Information) are identified by their intensity, downfield-shifted ^{1}H resonances, and upfield-shifted ^{13}C resonances.[15] DQF-COSY signals at 16.42, −7.45, and −8.21 ppm correspond to the α, β(*trans*) and β(*cis*) protons of one vinyl group, respectively (Figure S5.6A of the Supporting Information). The β protons exhibit strong NOEs to a heme CαH$_3$ (δ(^{1}H) = 14.17 ppm), itself in contact with another heme CαH$_3$ (δ(^{1}H) = 10.28 ppm, Figure S5.6B-D of the Supporting Information). The

latter heme methyl shows NOEs to a set of four *J*-coupled protons belonging to a heme propionate. These connectivities assign the heme 1-CαH$_3$, 8-CαH$_3$, and 7-propionate (Figure 5.1A) and demonstrate that GlbN-B contains an intact 2-vinyl group. NOEs between the 2-vinyl CβH(*trans*) and the 1-CαH$_3$ (Figure S5.6C of the Supporting Information), and between the 2-vinyl Hα and the heme α-meso proton (not shown), define the orientation of the vinyl group in the heme plane as "*trans*."[42]

The heme 3-CαH$_3$ (δ(^{1}H) = 6.87 ppm) exhibits a strong NOE to the signal observed at δ(^{1}H) = –1.04 ppm in the ^{1}H-^{15}N lr-HMQC spectrum (Figure 5.5A and Figure S5.6I of the Supporting Information). This proton or set of protons is *J*-correlated to proton(s) resonating at –1.36 ppm (Figure 5.5B). ^{1}H-^{13}C cross peaks corresponding to the –1.36, –1.04 ppm pair are resolved (Figure 5.5D) so that their multiplicity can be determined with a ^{13}C band-selective ^{1}H-^{13}C HSQC experiment collected without ^{1}H decoupling in the indirect dimension (Figure S5.7 of the Supporting Information). The signal at –1.04 ppm is a quartet, whereas the signal at –1.36 ppm is a doublet, both with $^1J_{CH}$ ~ 130 Hz. Thus, the pattern indicates a CαH–CβH$_3$ system, and the proximity to the 3-CαH$_3$ confirms the modification of the 4-vinyl group. The remaining heme signals are assigned as usual with a combination of ^{1}H and ^{13}C data.

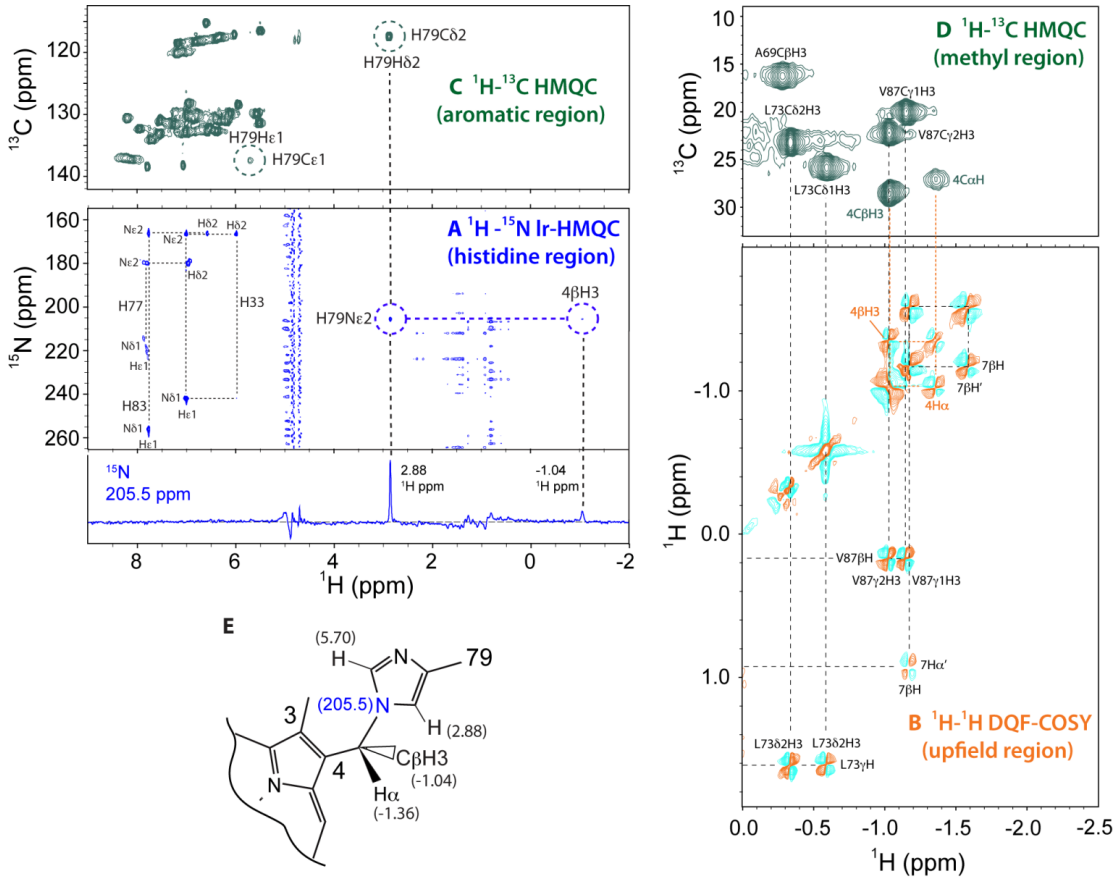


Figure 5.5. Identification of the heme modification in ferric L79H/H117A GlbN-B. A, ^{1}H-^{15}N lr HMQC highlighting an unusual pair of proton signals (2.88, –1.04 ppm) with a common ^{15}N shift at 205.5 ppm. B, portion of the ^{1}H-^{1}H DQF-COSY data showing a *J*-connectivity between the protons at –1.04 ppm and –1.36 ppm. Signals corresponding to Val87, Leu73, and 7-propionate spin systems are also labeled. C, aromatic region of the natural abundance ^{1}H-^{13}C HMQC spectrum identifying the signal at 2.88 ppm (^{1}H shift) as His79 Hδ2 through its directly attached Cδ2 (117.5 ppm). The His79 CεH1 correlation is also circled. D, methyl region of the natural abundance ^{1}H-^{13}C HMQC spectrum showing resolved ^{13}C shifts for the 4CβH$_3$–CαH pair. E, structure of posttranslational modification consistent with the data. ^{1}H-^{1}H and ^{1}H-^{13}C data were acquired on ~ 5 mM ferric L79H/H117A GlbN-B at pH* 7.04, 298 K, 99.9 % D_2O; ^{1}H-^{15}N data were acquired on a ~1 mM ^{15}N-labeled GlbN-B sample at pH 7.62, 298 K in H_2O.

The last step in the structure determination is the characterization of the modified His79 and its bond to the heme. According to the ^{1}H-^{13}C HMQC spectrum, the proton at 2.88 ppm (lr correlation to a ^{15}N at 205.5 ppm, Figure 5.5A) is directly bonded to a ^{13}C resonating at 114.8 ppm (Figure 5.5C). This shift is consistent with a histidine Cδ2[43] and it assigns the ^{15}N signal as the Nε2 of His79. Furthermore, the ^{1}H–^{1}H TOCSY spectrum contains a connectivity between 2.88 ppm and 5.70 ppm, assigning the latter proton to His79 Hε1 (data not shown); accordingly the His79 CεH1 direct correlation is present in the aromatic region of ^{1}H-^{13}C HMQC data (Figure 5.5C, $\delta(^{13}C)$ = 137.6 ppm). The PTM covalent structure is therefore as depicted in Figure 5.5E. Strong dipolar contacts are observed between His79 Hδ2 and the 4-CαH–CβH_3 moiety (Figure S5.6K of the Supporting Information), and His79 Hε1 is in contact with both the 3-CαH_3 (not shown) and 4-CβH_3 substituents (Figure S5.6I of the Supporting Information). These effects are consistent with a conformation in which His79 Hε1 faces the heme 3-CαH_3 and His79 Hδ2 faces the heme 4-CαH/5-CαH_3 side of heme pyrrole C. Heme and reacted histidine chemical shifts are listed in Table S5.1 of the Supporting Information.

The composite NMR data establish that His79 reacted with the 4-vinyl substituent to form a Markovnikov adduct with R stereochemistry. The addition is analogous to the wild-type PTM and consistent with a reduction-driven reaction in which 4-Cβ protonation is followed by nucleophilic attack linking His79 Nε2 to the heme 4-Cα.[20]

Heme environment in ferric GlbN-B. Several Tyr and Phe side chains line the GlbN-B heme pocket and are in contact with the cofactor (Figure S5.6E-I of

the Supporting Information). In addition, contacts between upfield-shifted aliphatic groups and heme (Figure S5.6I-L of the Supporting Information) orient the modified heme group unambiguously in the cavity of L79H/H117A GlbN-B. These effects are consistent with the crystallographic model of wild-type GlbN-A (PDB ID: 1RTX[38]); no major structural change occurs within the heme pocket upon covalent attachment via the nonnative His79 linkage (Figure S5.6M of the Supporting Information). This is an important feature of GlbN-B because, on the basis of geometry alone, a histidine at position 117 should be able to react with the 2-vinyl in this already modified protein. The hybrid *b*/*c* heme, however, is expected to have electronic properties different from those of a *b* heme.[12] Formation of histidine–heme linkages at both vinyl groups would provide additional clues about the mechanism and the robustness of the modification. We therefore attempted reaction with the single variant, L79H GlbN.

Properties of L79H GlbN. L79H GlbN loses ferric heme when passed through the anion exchange column and requires reconstitution with heme after apoprotein purification. The ^{1}H spectrum of ferric L79H GlbN, shown in Figure 5.3E, does not present obvious hyperfine-shifted lines and exhibits little chemical shift dispersion. The thermal response (data not shown) is within error of that of the double variant (Figure S5.3 of the Supporting Information).

Heme modification in L79H GlbN monitored optically. Reduction of L79H GlbN as described for L79H/H117A GlbN yields qualitatively similar results. In the deadtime of the experiment, a ferrous spectrum is obtained with maxima near 426, 560, and 531 nm (Figure S5.8 of the Supporting Information,

red trace) and reduced intensity. A slow phase then leads to a second ferrous species with recovered Soret, α and β bands at 420, 555 (~550 sh), and 524 (~530 sh) nm, respectively (Figure S5.8 of the Supporting Information, purple trace). These blue shifts were twice as large as observed in the same experiment with L79H/H117A GlbN (Figure 5.4A), but occurred over a similar timescale (~1 h).

The result of the hemochromogen assay performed on reacted L79H GlbN is shown in Figure S5.4B of the Supporting Information, along with that of cytochrome *c* for reference. The α and β bands of the *bis*-pyridine ferrous complex are detected at 550 and 520 nm. These values are similar to those of mesoheme and *c* heme (Figures S4B and S4C of the Supporting Information), chromophores in which both the 2 and 4 substituents are saturated. Henceforth, the product of the L79H GlbN reaction is referred to as GlbN-AB.

Characterization of ferric GlbN-AB by NMR spectroscopy. The NMR strategy described above for L79H/H117A GlbN-B was applied to ferric L79H GlbN-AB. The one-dimensional ^{1}H spectrum (Figure 5.3F) exhibits sharp lines and indicates the formation of a specific product. The ^{1}H-^{15}N lr-HMQC spectrum (Figure S5.9 of the Supporting Information) contains signals arising from all five non-axial histidines. Resonances corresponding to His83, His33, and His77 can be assigned by comparison to L79H/H117A GlbN-B data. Additionally, a pair of cross peaks with common ^{15}N shift (209.9 ppm) and ^{1}H shifts of 3.42 and –0.57 ppm provides initial evidence that His79 reacted as in GlbN-B. By elimination, His117 is responsible for the remaining three signals, at $\delta(^{15}N)$ = 248.1 ppm with two connectivities ($\delta(^{1}H)$ = 6.84 and 6.16 ppm), and

$\delta(^{15}N)$ = 250.4 ppm with one connectivity ($\delta(^{1}H)$ = 6.84 ppm). The ^{15}N shifts are comparable to those observed for His117 in wild-type GlbN-A[7] and are indicative of ^{15}N atoms without an attached proton.

The ^{1}H-^{13}C data identify the four heme CαH_3 resonances with characteristic upfield ^{13}C shifts (Figure S5.10 of the Supporting Information). The reacted 4-substituent in GlbN-AB has ^{1}H and ^{13}C shifts similar to those in GlbN-B (Figure S5.11 of the Supporting Information). The 4-CβH_3 ^{1}H (–0.57 ppm) displays a strong NOE to the heme 3-CαH_3 group ($\delta(^{1}H)$ = 10.44 ppm). Other NOEs from the 4-CαH–CβH_3 unit to the heme and protein are similar to those observed in L79H/H117A GlbN-B and establish the orientation of the reacted His79 and R stereochemistry at 4-Cα (Figure 5.6).

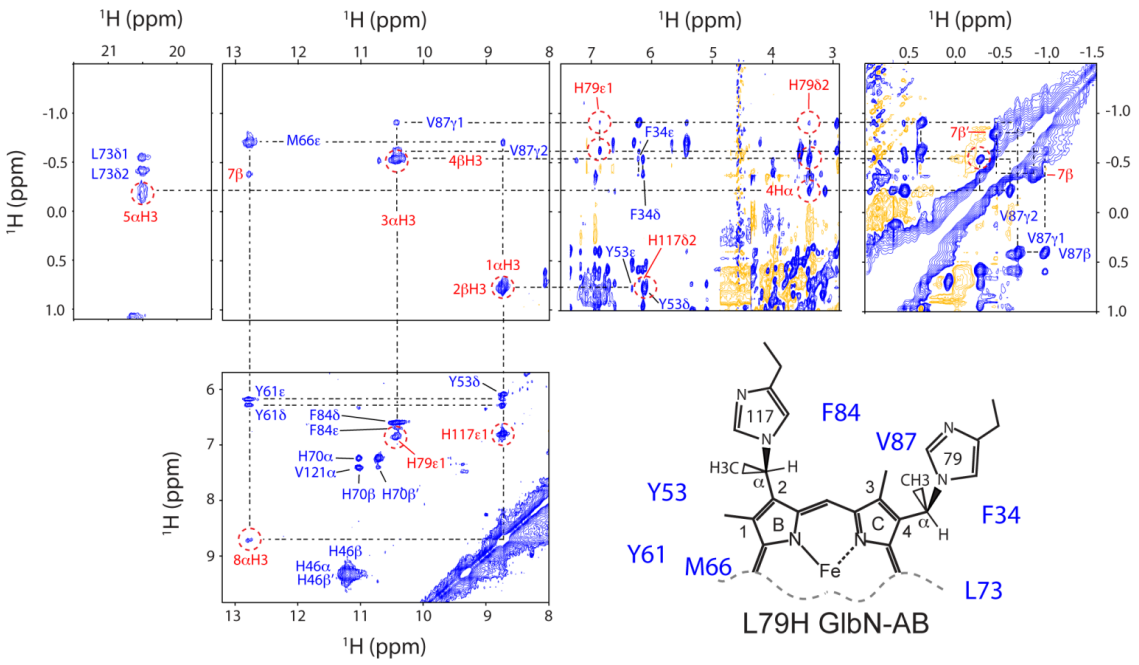


Figure 5.6. NOE contacts exhibited at the 2 and 4 sites of modification within L79H GlbN-AB. As observed in L79H/H117A GlbN-B (Figure S5.6I,K of the Supporting Information), the heme 4-βH_3 signal (–0.57 ppm) exhibits strong cross peaks with His79 Hδ2, Phe34 ring protons and the heme 4-Hα and 3-αH_3. The 4-Hα resonance is in dipolar contact with His79 Hδ2, Phe34 Hδs (weak) and the heme 5-αH_3 group. On pyrrole B, the heme 2-βH_3 signal displays strong interactions to His117 Hδ2, the aromatic protons of Tyr53, and the heme 1-αH_3, itself in contact with Tyr53, Tyr61, Met66 εH_3, His117 Hε1, and the heme 8-αH_3. Signals derived from the modified heme are highlighted in red. The set of dipolar connectivities are consistent with R stereochemistry at both heme 2- and 4-substituents (bottom right, Figure 5.6). Ferric L79H GlbN-AB NOESY data (τ_{mix} = 50 ms) were acquired on a ~5 mM sample at pH* 7.61, 313 K, in 99.9 % D_2O.

Vinyl groups are conspicuous in ^{1}H–^{1}H DQF-COSY data of ferric *bis*-histidine GlbNs. The downfield shifted CαH exhibits two strong upfield cross peaks, one with a large $^3J_{HH}$ (~12–15 Hz) corresponding to the *trans* proton. No spin system consistent with an intact vinyl moiety was detected in the spectrum of ferric L79H GlbN-AB. Signals corresponding to His117 Cδ2 and Cε1 were

identified in the ^{1}H-^{13}C HMQC spectrum (Figure S5.12 of the Supporting Information), along with a far downfield-shifted methyl group ($\delta(^{13}C)$ = 61.8 ppm, Figure S5.11 of the Supporting Information), which was absent from GlbN-B spectra, but present in wild-type GlbN-A and assigned to the 2-CβH_3. The attached protons (0.74 ppm) display strong NOEs to the heme 1-CαH_3 group ($\delta(^{1}H)$ = 8.75 ppm) and His117 Hδ2. In addition, NOEs involving the 1-CαH_3 and His117 Hε1 are as in wild-type GlbN-A and establish the presence of the linkage between His117 Nε2 and the heme 2-Cα (Figure 5.6). Heme and reacted histidine chemical shifts are listed in Table S5.2 of the Supporting Information.

We conclude that two histidine–heme covalent linkages can readily be created in GlbN. The reaction required no enzyme, and a unique product is obtained with complete stereospecificity. To our knowledge alkylation of two histidines by a heme group has not been reported. This novel modification parallels the double thioether modification observed in the majority of *c*-type cytochromes.

Frequency of histidine presence near a heme vinyl group. An analysis of 341 three-dimensional structures of proteins with at most 95% sequence similarity and containing a *b* heme (Table S5.3 of the Supporting Information) was performed to identify residues with a backbone Cα within 7 Å of a vinyl Cα and a side chain Cβ within 6 Å of the same vinyl Cα. These arbitrary thresholds were chosen to capture His117 and Leu79 in GlbN-A and be otherwise permissive considering that the distance between a histidine Cα and the tip of its imidazole ring is shorter than 5 Å in any rotameric state.

When the number of hits with vinyls (excluding iron-ligating residues) is normalized for frequency of occurrence in the set, hydrophobic residues are found to be preferred in the order Val > Phe > Ile > Leu (Figure S5.13 of the Supporting Information). Thr is also preferred, along with Met and Tyr, two residues that can be oxidized and form heme adducts.[5,44] Also unsurprising is the depletion of ionizable residues. The order is Asp ~ Lys > Glu > Cys > Arg > His, Asp and Lys being the rarest. Interestingly, each of the low propensity side chains (except Arg) is documented to react covalently with the porphyrin macrocycle in various cytochromes and peroxidases.[4,5,44] Of direct relevance to this study are 40 histidines found near vinyl groups. A majority of the imidazole rings are involved in strong interactions (helix N-capping or salt bridge to a propionate) and none are poised for vinyl addition in the $\chi 1$ rotameric state observed in the structures or alternative $\chi 1$ rotameric states. When the same distance criteria are used with propionate instead of vinyl Cα atoms, His and Arg have high propensities whereas Cys, Glu, and Asp are strongly avoided.

Discussion

Structural and chemical criteria for histidine–vinyl addition in GlbN. There are obvious geometric requirements for the histidine–heme modification to occur. The histidine Nε2 atom must be able to reach a vinyl Cα atom, and the local structure must be such that sp^3 bond angles can be established at that carbon. In our search for suitable positions in the X-ray structures of

GlbN and proteins such as myoglobin and cytochrome b_5, we found the distance and angle criteria are difficult to satisfy simultaneously without global structural adjustments.

Position 79 in the X-ray structure of GlbN-A is an exception; a histidine modeled in place of the wild-type leucine fulfills both requirements with few clashes. However, only the *trans* $\chi 1$ rotameric state can form product. This restrictive geometry forces the polar imidazole ring in contact with the 4-vinyl group and protein hydrophobic groups, unfavorable interactions that weaken the affinity of the apoprotein for ferric heme. High pH conditions do not improve binding, which leads us to conclude that histidine desolvation and steric hindrance are the primary causes of variant holoprotein instability. Low pH conditions also reduce heme affinity, likely because histidine protonation is energetically disfavored in the non-polar environment of the holoprotein. The quality of the NMR spectra did not allow for the determination of the ionization constant of His79.

The contrasting distribution of histidines near vinyls and propionates illustrates clearly the influence of electrostatics (Figure S5.13 of the Supporting Information) on heme packing. In natural *b* heme proteins, the effects of introducing a bulky polar or charged residue in or near a heme pocket can presumably be moderated with compensatory mutations and the luxury of an evolutionary timescale. In this light, the depletion of histidines near heme vinyl groups may be caused by physico-chemical properties other than charge and polarity. The imidazole ring is an excellent ligand for ferric and ferrous heme iron. By removing histidines from heme binding sites, the risk of misligation is

minimized during folding. The facile histidine–heme PTM in the L79H GlbN variants suggests that reactions leading to aberrant covalent attachment of chromophores are another potential source of pressure eliminating histidines near vinyls. In these two properties, ligation and reactivity, histidines resemble cysteines, which are thought to be preferentially avoided in *c*-type cytochrome sequences to ensure the formation of desired and homogeneous holoproteins.[12,45]

GlbN-A, GlbN-B, and GlbN-AB are chemically analogous and form under similar reducing conditions. We therefore hypothesize that the electrophilic addition mechanism proposed for GlbN-A is conserved in GlbN-B and GlbN-AB. Accordingly, the nucleophilic histidine (His117, His79) attacks a vinyl with its Nε2 atom, the common R stereochemistry apparently stemming from a protonated vinyl intermediate oriented in *trans* as in the ferric starting material. In support of the reactive residue delivering the proton, ferrous H117C GlbN is able to form the thioether linkage of a *c*-type cytochrome, linking Cys117 Sγ to the heme 2-Cα,[20] but H117S GlbN is not modified in the reduced state (data not shown), consistent with the high energetic cost of deprotonating the serine hydroxyl group. Additionally, since water does not compete efficiently with the side chain nucleophile and does not add to the vinyl group in any of the studied proteins, other geometrical requirements are inferred, for example close proximity of the two reactive partners once heme protonation has occurred. A concerted protonation and attack mechanism can also account for the above observations.

Optical data indicate that under similar conditions, crosslinking in both L79H/H117A and L79H GlbN occurs on a timescale several orders of magnitude slower than in the wild-type protein[20] (Figure 5.4A and Figure S5.8 of the Supporting Information). Rather than intrinsic difficulty in protonating the vinyl group, heme scavenging experiments (Figure 5.4B) and ^{1}H NMR spectra point to reduced affinity for the ferric and ferrous heme, i.e., shorter residence time in a reaction competent state, as responsible for the observed deceleration. In L79H GlbN, the apparent monophasic conversion to product (Figure S5.8 of the Supporting Information) is consistent with the first crosslinking event being rate determining and the second occurring faster than the first. Whether His117 or His79 initially attacks the heme cannot be distinguished with the current data. Importantly, the production of a double crosslink indicates that addition of an imidazole to one vinyl group does not inhibit reaction at the second vinyl group. It is especially clear from L79H GlbN that high heme affinity is not required for a stereospecific posttranslational modification.

Histidine–heme linkages versus cysteine–heme linkages. Unlike the cysteine thiol reactant and thioether product of *c*-type cytochromes, the imidazole ring is titratable at physiological pH in its unmodified *and* modified forms. In GlbN, the kinetics of the PTM are dictated by the competing rates of vinyl protonation and acid-induced ferrous heme dissociation: between pH 9.5 and 6.0, the reaction is accelerated several orders of magnitude,[19] but too low a pH causes heme loss. The properties of the histidine crosslinked product, which are under functional constraint, may depend on pH through a coupling of iron redox state and the modified histidine pK_a. Depending on local solvent

accessibility, the alkylated histidine may remain ionizable at physiological pH, favoring charge heterogeneity at the heme active site. Thus, a histidine-modified heme may be undesirable in electron transfer proteins such as the *c*-type cytochromes.

In natural proteins, the preference for cysteine–heme attachment is likely to stem not only from the properties of the product but also from the versatility of the thiol group. Cysteine can react with the heme via reductive and oxidative mechanisms[13] and has additional options of deprotonation, methylation, phosphorylation, disulfide bridging (redox sensing), nitrosylation, and numerous other oxidative modifications.[46] This high reactivity enables multiple modes of regulation, some of which are critical in cytochrome *c* maturation.[13] In contrast, the opportunities for regulation of histidine–vinyl chemistry are more limited, principally to iron redox state, pH, and histidine PTMs such as phosphorylation and methylation.

Because a cysteine at position 117 can serve as the nucleophile in GlbN,[20] and cysteine/histidine interchangeability has been documented in the heme chaperone CcmE,[47,48] the advantages of a histidine–heme bond over a cysteine–heme bond should also be considered. Under high oxidative and nitrosative stress, conditions that many cyanobacteria encounter, undesirable cysteine modifications may interfere with specific covalent heme attachment.[49] We proposed that the histidine–heme linkage is a solution to this problem in *Synechococcus* sp. PCC 7002.[8] Furthermore, cysteine tends to react with non-heme molecules containing vinyl groups, as demonstrated by spontaneous attachment of biliverdin IXα to phytochromes in vitro[50,51] or the reactive

ethylidines found in common cyanobacterial chromophores such as phycocyanobilin.[52] A histidine may be preferred over a cysteine in GlbN to limit these types of side reactions.

In conclusion, the present study highlights that heme vinyl groups are capable of spontaneous reaction with histidines to yield products similar to *c*-type cytochromes. This chemistry can occur even in highly destabilized proteins. In addition to providing opportunities for a comparative study of cysteine- and histidine-modified hemes and the consequences of such PTMs, histidine–heme attachment holds promises for the rational production of stable, artificial proteins containing a covalently bound heme. Interesting prospects include the design of redox protein switches responding to the reacted histidine's ionization state.

Acknowledgements

We thank Christopher Falzone for careful reading of the manuscript and Ananya Majumdar for assistance with the NMR experiments.

References

(1) Lippard, S. J., and Berg, J. M. (1994) *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA.

(2) Salemme, F. R. (1977) Structure and function of cytochromes *c*. *Annu. Rev. Biochem. 46*, 299-329.

(3) Pettigrew, G. W., and Moore, G. R., (Eds.) (1991) *Cytochromes c: Evolutionary, Structural, and Physicochemical Aspects*, Springer-Verlag, Berlin.

(4) Colas, C., and Ortiz de Montellano, P. R. (2003) Autocatalytic radical reactions in physiological prosthetic heme modification. *Chem. Rev. 103*, 2305-2332.

(5) Battistuzzi, G., Stampler, J., Bellei, M., Vlasits, J., Soudi, M., Furtmuller, P. G., and Obinger, C. (2011) Influence of the covalent heme-protein bonds on the redox thermodynamics of human myeloperoxidase. *Biochemistry 50*, 7987-7994.

(6) Vinogradov, S. N., and Moens, L. (2008) Diversity of globin function: enzymatic, transport, storage, and sensing. *J. Biol. Chem. 283*, 8773-8777.

(7) Vu, B. C., Jones, A. D., and Lecomte, J. T. J. (2002) Novel histidine-heme covalent linkage in a hemoglobin. *J. Am. Chem. Soc. 124*, 8544-8545.

(8) Scott, N. L., Xu, Y., Shen, G., Vuletich, D. A., Falzone, C. J., Li, Z., Ludwig, M., Pond, M. P., Preimesberger, M. R., Bryant, D. A., and

Lecomte, J. T. J. (2010) Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry 49*, 7000-7011.

(9) Smagghe, B. J., Trent, J. T., 3rd, and Hargrove, M. S. (2008) NO dioxygenase activity in hemoglobins is ubiquitous in vitro, but limited by reduction in vivo. *PLoS ONE 3*, e2039.

(10) Barker, P. D., and Ferguson, S. J. (1999) Still a puzzle: why is haem covalently attached in *c*-type cytochromes? *Structure Fold. Des. 7*, R281-R290.

(11) Allen, J. W., Barker, P. D., Daltrop, O., Stevens, J. M., Tomlinson, E. J., Sinha, N., Sambongi, Y., and Ferguson, S. J. (2005) Why isn't `standard' heme good enough for *c*-type and d_1-type cytochromes? *Dalton Trans.*, 3410-3418.

(12) Bowman, S. E., and Bren, K. L. (2008) The chemistry and biochemistry of heme *c*: functional bases for covalent attachment. *Nat. Prod. Rep. 25*, 1118-1130.

(13) Kranz, R. G., Richard-Fogal, C., Taylor, J. S., and Frawley, E. R. (2009) Cytochrome *c* biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. *Microbiol. Mol. Biol. Rev. 73*, 510-528.

(14) Stevens, J. M., Uchida, T., Daltrop, O., and Ferguson, S. J. (2005) Covalent cofactor attachment to proteins: cytochrome *c* biogenesis. *Biochem. Soc. Trans. 33*, 792-795.

(15) Vu, B. C., Vuletich, D. A., Kuriakose, S. A., Falzone, C. J., and Lecomte, J. T. J. (2004) Characterization of the heme-histidine cross-link in cyanobacterial hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. *J. Biol. Inorg. Chem. 9*, 183-194.

(16) Hoy, J. A., Smagghe, B. J., Halder, P., and Hargrove, M. S. (2007) Covalent heme attachment in *Synechocystis* hemoglobin is required to prevent ferrous heme dissociation. *Protein Sci. 16*, 250-260.

(17) Pond, M. P., Majumdar, A., and Lecomte, J. T. J. (2012) Influence of heme post-translational modification and distal ligation on the backbone dynamics of a monomeric hemoglobin. *Biochemistry 51*, 5733-5747.

(18) Simonneaux, G., and Bondon, A. (2005) Mechanism of electron transfer in heme proteins and models: the NMR approach. *Chem. Rev. 105*, 2627-2646.

(19) Preimesberger, M. R., Pond, M. P., Majumdar, A., and Lecomte, J. T. J. (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.

(20) Nothnagel, H. J., Preimesberger, M. R., Pond, M. P., Winer, B. Y., Adney, E. M., and Lecomte, J. T. J. (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.

(21) Nothnagel, H. J., Love, N., and Lecomte, J. T. J. (2009) The role of the heme distal ligand in the post-translational modification of *Synechocystis* hemoglobin. *J. Inorg. Biochem. 103*, 107-116.

(22) Lecomte, J. T. J., Scott, N. L., Vu, B. C., and Falzone, C. J. (2001) Binding of ferric heme by the recombinant globin from the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochemistry 40*, 6541-6552.

(23) Scott, N. L., and Lecomte, J. T. J. (2000) Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803. *Protein Sci. 9*, 587-597.

(24) Teale, F. W. J. (1959) Cleavage of heme-protein link by acid methylethylketone. *Biochim. Biophys. Acta 35*, 543.

(25) 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.

(26) Antonini, E., Brunori, M., Caputo, A., Chiancone, E., Rossi-Fanelli, A., and Wyman, J. (1964) Studies on the structure of hemoglobin. III. Physicochemical properties of reconstituted hemoglobins. *Biochim. Biophys. Acta 79*, 284-292.

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

(28) Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., and Skelton, N. J. (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, San Diego, USA.

(29) Kadkhodaei, M., Hwang, T. L., Tang, J., and Shaka, A. J. (1993) A simple windowless mixing sequence to suppress cross-relaxation in TOCSY experiments. *J. Magn. Reson. A 105*, 104-107.

(30) Inubushi, T., and Becker, E. D. (1983) Efficient detection of paramagnetically shifted NMR resonances by optimizing the WEFT pulse sequence. *J. Magn. Reson. 51*, 128-133.

(31) Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci. 2*, 543-558.

(32) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277-293.

(33) Goddard, T. D., and Kneller, D. G. (2006) SPARKY 3. *University of California, San Francisco*.

(34) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135-140.

(35) Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol. 112*, 535-542.

(36) http:/www.rcsb.org/

(37) Hamelryck, T., and Manderick, B. (2003) PDB parser and structure class implemented in Python. *Bioinformatics 19*, 2308-2310.

(38) Hoy, J. A., Kundu, S., Trent, J. T., 3rd, Ramaswamy, S., and Hargrove, M. S. (2004) The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. *J. Biol. Chem. 279*, 16535-16542.

(39) Vu, B. C., Nothnagel, H. J., Vuletich, D. A., Falzone, C. J., and Lecomte, J. T. J. (2004) Cyanide binding to hexacoordinate cyanobacterial hemoglobins: Hydrogen bonding network and heme pocket rearrangement in ferric H117A *Synechocystis* Hb. *Biochemistry 43*, 12622-12633.

(40) Scott, N. L., Falzone, C. J., Vuletich, D. A., Zhao, J., Bryant, D. A., and Lecomte, J. T. J. (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.

(41) Lee, D., Pervushin, K., Bischof, D., Braun, M., and Thöny-Meyer, L. (2005) Unusual heme-histidine bond in the active site of a chaperone. *J. Am. Chem. Soc. 127*, 3716-3717.

(42) Marzocchi, M. P., and Smulevich, G. (2003) Relationship between heme vinyl conformation and the protein matrix in peroxidases. *J. Raman Spect. 34*, 725-736.

(43) http://www.bmrb.wisc.edu/

(44) Pearson, A. R., Elmore, B. O., Yang, C., Ferrara, J. D., Hooper, A. B., and Wilmot, C. M. (2007) The crystal structure of cytochrome P460 of *Nitrosomonas europaea* reveals a novel cytochrome fold and heme-protein cross-link. *Biochemistry 46*, 8340-8349.

(45) Sawyer, E. B., and Barker, P. D. (2012) Continued surprises in the cytochrome *c* biogenesis story. *Protein Cell 3*, 405-409.

(46) Lecomte, J. T. J., and Falzone, C. J. (2013) Protein structure: Unusual covalent bonds. In *eLS*, John Wiley & Sons, Ltd, Chichester, UK.

(47) Aramini, J. M., Hamilton, K., Rossi, P., Ertekin, A., Lee, H. W., Lemak, A., Wang, H., Xiao, R., Acton, T. B., Everett, J. K., and Montelione, G. T. (2012) Solution NMR structure, backbone dynamics, and heme-binding properties of a novel cytochrome *c* maturation protein CcmE from *Desulfovibrio vulgaris*. *Biochemistry 51*, 3705-3707.

(48) Goddard, A. D., Stevens, J. M., Rao, F., Mavridou, D. A., Chan, W., Richardson, D. J., Allen, J. W., and Ferguson, S. J. (2010) c-Type cytochrome biogenesis can occur via a natural Ccm system lacking CcmH, CcmG, and the heme-binding histidine of CcmE. *J. Biol. Chem. 285*, 22882-22889.

(49) Barker, P. D., Ferrer, J. C., Mylrajan, M., Loehr, T. M., Feng, R., Konishi, Y., Funk, W. D., MacGillivray, R. T., and Mauk, A. G. (1993) Transmutation of a heme protein. *Proc. Natl. Acad. Sci. U. S. A. 90*, 6542-6546.

(50) Lamparter, T., Michael, N., Mittmann, F., and Esteban, B. (2002) Phytochrome from *Agrobacterium tumefaciens* has unusual spectral properties and reveals an N-terminal chromophore attachment site. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11628-11633.

(51) Borucki, B., Seibeck, S., Heyn, M. P., and Lamparter, T. (2009) Characterization of the covalent and noncovalent adducts of Agp1 phytochrome assembled with biliverdin and phycocyanobilin by circular dichroism and flash photolysis. *Biochemistry 48*, 6305-6317.

(52) Scheer, H., and Zhao, K. H. (2008) Biliprotein maturation: the chromophore attachment. *Mol. Microbiol. 68*, 263-276.

Supporting Information for

Facile Heme Vinyl Posttranslational Modification in a Hemoglobin

Matthew R. Preimesberger, Belinda B. Wenke, Lukas Gilevicius, Matthew P. Pond, and Juliette T.J. Lecomte

Scheme S5.1. Proposed reduction-driven heme crosslinking mechanism in wild-type GlbNs consistent with current experimental evidence.[1,2] Following reduction, the 2-vinyl group undergoes rate-determining labile proton addition at Cβ to generate a transient α-carbenium ion. His117 Nε2 attacks the Cα electrophile to generate the resulting Markovnikov adduct (R stereochemistry). If the reaction is drawn as previously,[3] the intermediate closely resembles the reactive ethylidines of phycocyanobilin.

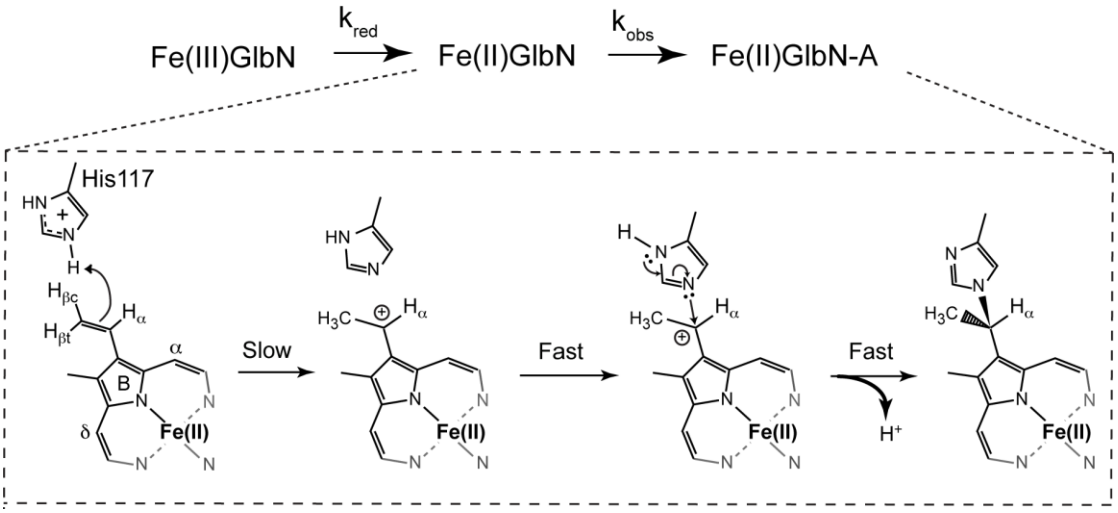


Figure S5.1. Optical absorbance spectra of freshly reconstituted ferric L79H/H117A GlbN (dark dashed line) and L79H GlbN (grey dot-dot-dashed line). Near neutral pH, both GlbNs (~10 μM protein, room temperature) have spectral properties comparable to those of the wild-type GlbN and are consistent with the predominant formation of low-spin (S = ½) *bis*-histidine complexes. In each variant GlbN, adjustment of the pH to more basic or acidic values resulted in an increase in free hemin signature (the shoulder below 400 nm increases and a charge transfer band emerges above 600 nm).

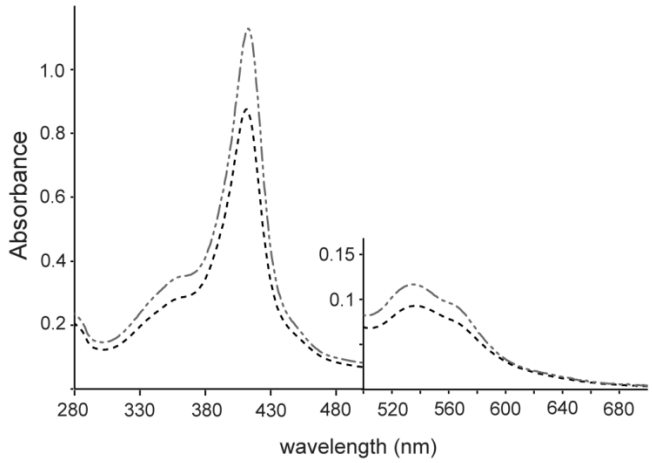


Figure S5.2. Ribbon diagram of wild-type *Synechocystis* GlbN-A (PDB ID: 1RTX). Helices are labeled A–H. Note the absence of the D helix and the kink in the H helix, which is divided into H and H' portions. The modified heme, along with the proximal (H70), distal (H46), and reactive (H117) histidines, are presented as grey sticks. The heme in this structure was modeled as planar, but related crystal structures exhibit heme ruffling (PDB ID: 2HZ1, 2HZ2, 2HZ3, 4I0V). The figure was prepared with MOLSCRIPT.[4]

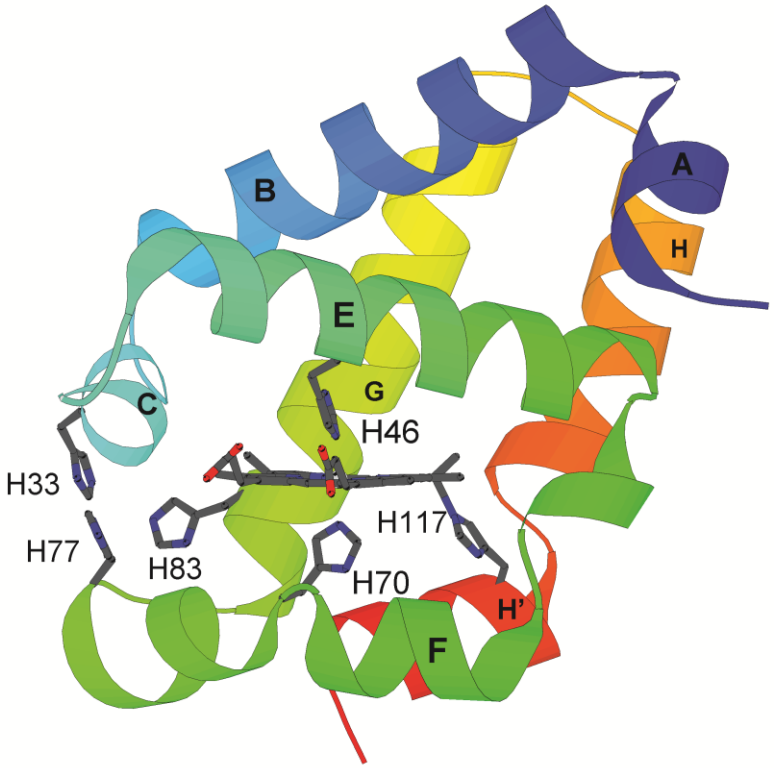


Figure S5.3. Optical absorbance spectra of ferric L79H/H117A GlbN collected as a function of temperature. At 25 °C (red trace), the double variant exhibits a profile consistent with a low-spin complex. Upon increasing the temperature, the Soret band decays in intensity and a charge transfer band emerges at ~610 nm, suggesting a low-spin to high-spin transition. As the sample is heated further, the spectrum shifts to the blue and adopts a profile resembling free hemin (brown trace). (Inset) Normalized absorbance at 410 nm as a function of temperature (L79H/H117A GlbN: black diamonds, wild-type GlbN: grey dashed lines[5]). The double variant lacks a native baseline and exhibits a broad transition with an apparent midpoint about 20 °C lower than wild-type GlbN. All transitions are irreversible and were not analyzed quantitatively. Sample conditions: ~10 μM GlbN, pH ~7.3.

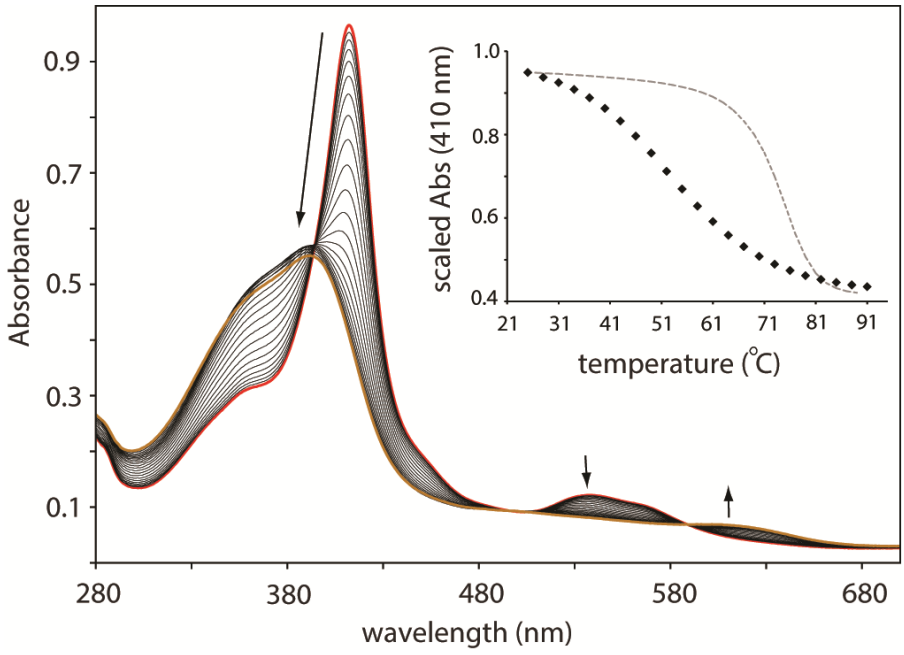
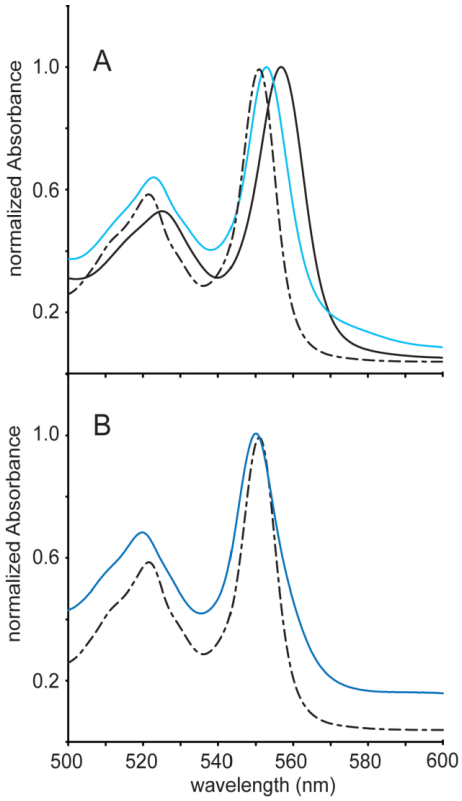


Figure S5.4. A. Pyridine hemochrome visible spectra of the dithionite-reacted (GlbN-B, cyan solid line) and unreacted (GlbN, black solid line) L79H/H117A variant overlaid with horse heart cytochrome *c* data (dash-dash-dotted line). **B.** Pyridine hemochrome data demonstrating the spectral similarity between the dithionite-treated L79H variant (GlbN-AB, blue solid line) and cytochrome *c* (dash-dash-dot line). **C.** Pyridine hemochrome absorbance maxima for different heme types. Note the trend between α and β band absorbance maxima and extent of heme vinyl saturation. All spectra were normalized to an α band max of 1.0.



C

Heme type	β band (nm)	α band (nm)	Ref
b heme (protoheme from HbA)	528	555	[6]
GlbN (*b* heme)	526	557	[7], this work
GlbN-A	522	552	[7]
GlbN-B	523	553	this work
GlbN-AB	520	550	this work
cytochrome c	521	551	this work
hematoheme	520	551	[6]
mesoheme	518	547	[6]

Figure S5.5. Identification of heme $C\alpha H_3$ and ligating histidine $C\beta H_2$ resonances in the natural abundance ^{1}H-^{13}C HMQC spectrum of ferric L79H/H117A GlbN-B (far-upfield ^{13}C region). Heme 2-vinyl and 6-propionate $C\alpha H$ correlations are also labeled. The second His70 $C\beta H$ resonance is present but very broad and not visible at this contour level. Sample conditions: ~5 mM GlbN-B, 20 mM phosphate, pH^{*} = 7.04, 99.9 % D_2O, 298 K.

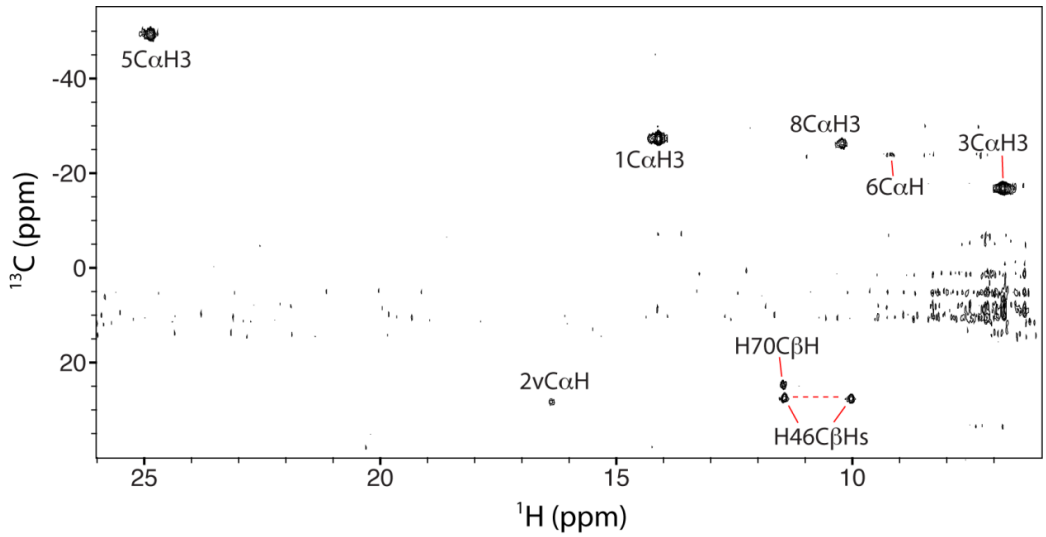
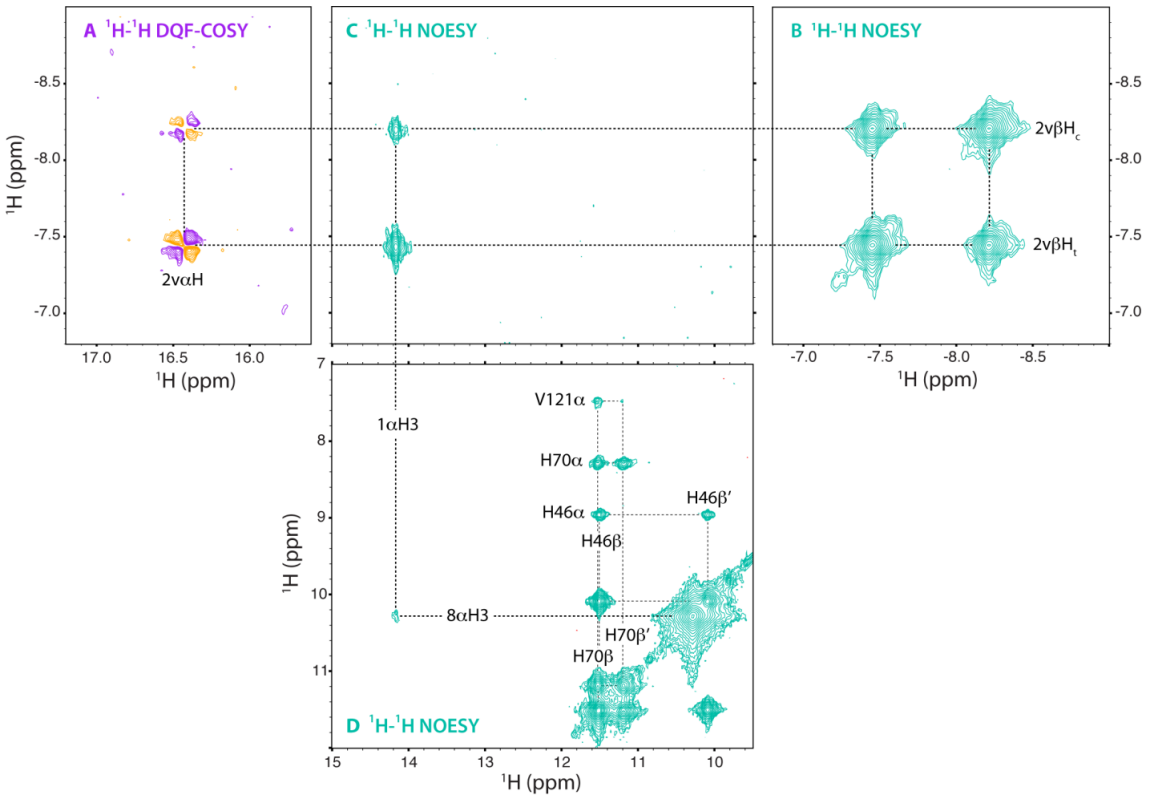
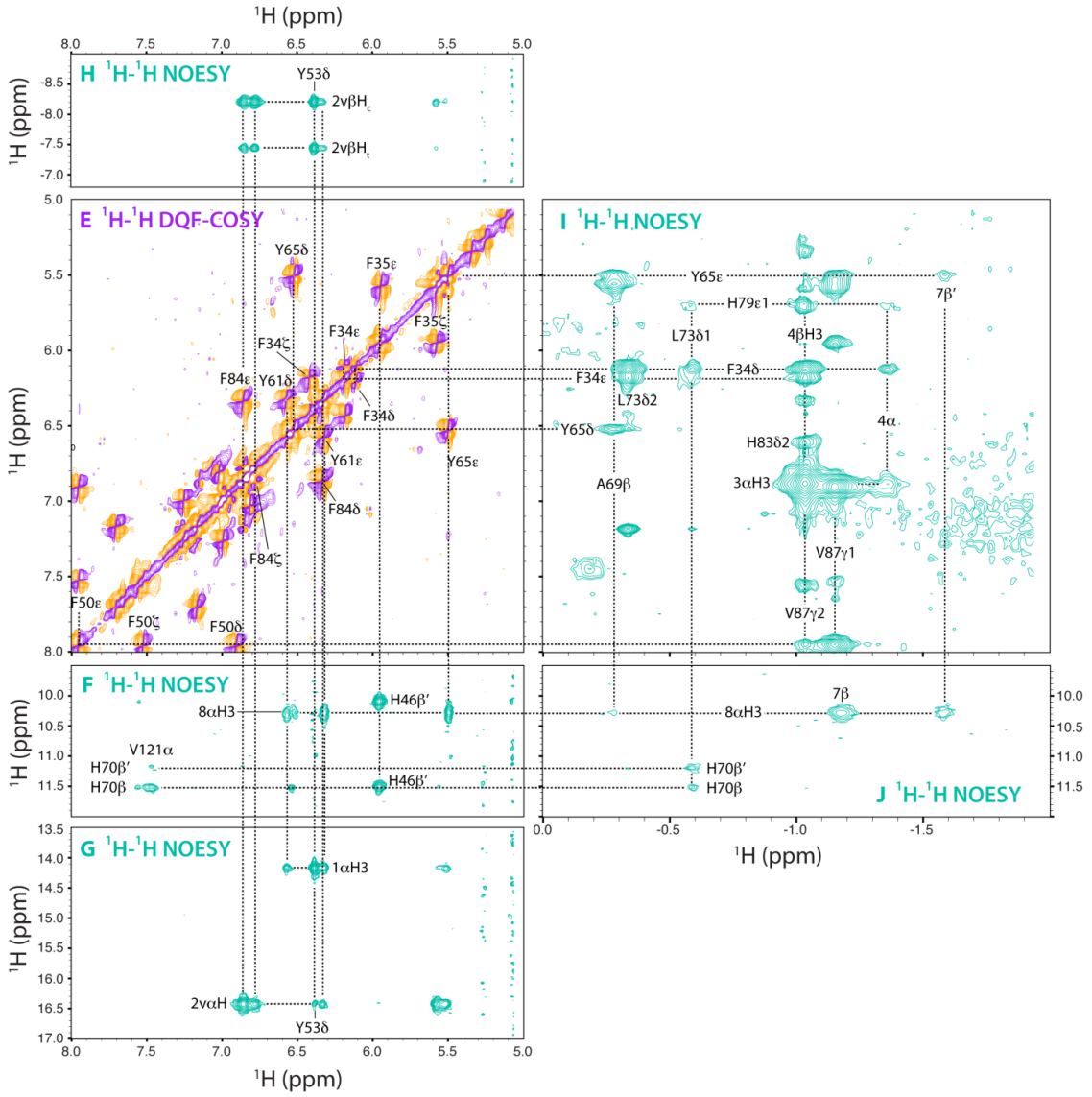
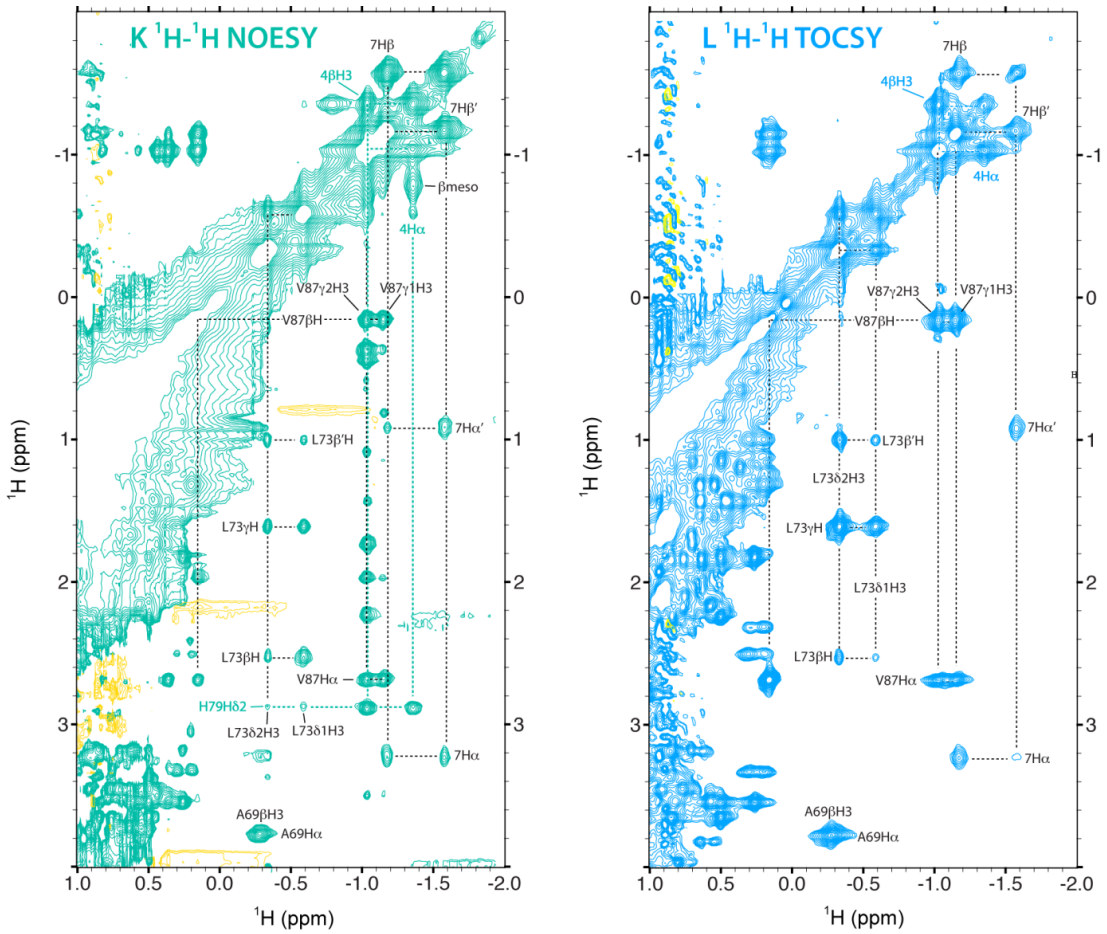


Figure S5.6. (Next three pages) Composite ^{1}H-^{1}H NMR spectra (DQF-COSY, NOESY, TOCSY) used for the characterization of the ferric L79H/H117A GlbN-B heme pocket. The displayed interaction network supports that the modified heme is embedded within GlbN in the "major" orientation as in the wild-type protein, but in this case the site of modification involves the heme 4-substituent and engineered His79. **A.** Note the DQF-COSY correlation pattern derived from the intact heme 2-vinyl group. **B.** The β protons of the 2-vinyl exhibit strong NOEs to the 1-αH_3 (shown in **C**), which is itself in contact with the 8-αH_3 (shown in **D**). These dipolar connectivities uniquely define the A-B pyrrole side of the heme prosthetic group. **E.** DQF-COSY data showing assigned Tyr and Phe side chains that are located within the GlbN-B heme pocket. **F.** The 8-αH_3 has expected NOEs to Tyr65 and Tyr61. **G.** The 1-αH_3 contacts Tyr61 and Tyr53; and the 2-vinyl protons show dipolar interactions with Tyr53 and Phe84 (shown in **G** and **H**). On the opposite side of the heme α-γ meso axis, NOEs between Phe34, Leu73, His70 and the reacted His79 (shown in **F, I** and **J)** establish that reaction occurs on the proximal side of the heme as depicted in Figure 5.2 of the main text. Overall, NOEs expected from the wild-type GlbN-A crystal structure (PDB ID: 1RTX) are observed in ferric L79H/H117A GlbN-B. For example, the newly generated heme 4-βH_3 group was in dipolar contact with the heme 3-αH_3, Phe34 ring protons and His83 Hδ2, in addition to the reacted His79 (shown in **I** and **K**). The 5-αH_3 group exhibited wild-type-like NOEs to Phe34 δHs and Leu73 $\delta 1H_3$ and $\delta 2H_3$ (data not shown). NOESY spectra also confirmed that the 3-αH_3 group packed near the His79 Hε1 and the side chain of Val87 (shown in **I**). **L.** Upfield region of the TOCSY spectrum with heme and paramagnetically-shifted signals from side chains such as Val87, Ala69, and Leu73 indicated. Sample conditions were as in Figure S5. **M.** Stereo triptych of the proximal heme environment in wild-type GlbN (PDB ID: 1RTX). Residues within dipolar contact of the cofactor are included, with darkened Cα atoms to guide the eye. The NOEs detected in both GlbN-B and GlbN-AB are consistent with the heme pocket geometry seen in the X-ray structure. In viewing this figure, the left pair is wall-eyed, and His70 should appear above the page.


A ^{1}H-^{1}H DQF-COSY
2vαH
C ^{1}H-^{1}H NOESY
B ^{1}H-^{1}H NOESY
2vβH$_c$
2vβH$_t$
^{1}H (ppm)
1αH3
8αH3
V121α
H70α
H46α
H46β
H46β′
H70β′
H70β
D ^{1}H-^{1}H NOESY


H ^{1}H-^{1}H NOESY
Y53δ
2νβH$_c$
2νβH$_t$
E ^{1}H-^{1}H DQF-COSY
Y65δ
F35ε
F35ζ
F34ε
F34ζ
F84ε
Y61δ
F34δ
Y61ε
Y65ε
F84δ
F84ζ
F50ε
F50ζ
F50δ
I ^{1}H-^{1}H NOESY
Y65ε
H79ε1
7β′
L73δ1
4βH3
F34ε
F34δ
L73δ2
Y65δ
4α
H83δ2
A69β
3αH3
V87γ1
V87γ2
F ^{1}H-^{1}H NOESY
8αH3
H46β′
V121α
H70β′
H70β
G ^{1}H-^{1}H NOESY
1αH3
2ναH
Y53δ
J ^{1}H-^{1}H NOESY
7β
8αH3
H70β′
H70β
^{1}H (ppm)


K ¹H-¹H NOESY
L ¹H-¹H TOCSY
7Hβ
4βH3
7Hβ′
βmeso
4Hα
V87γ2H3
V87γ1H3
V87βH
7Hα′
L73β′H
L73γH
L73βH
V87Hα
H79Hδ2
L73δ2H3
L73δ1H3
7Hα
A69βH3
A69Hα
¹H (ppm)

M

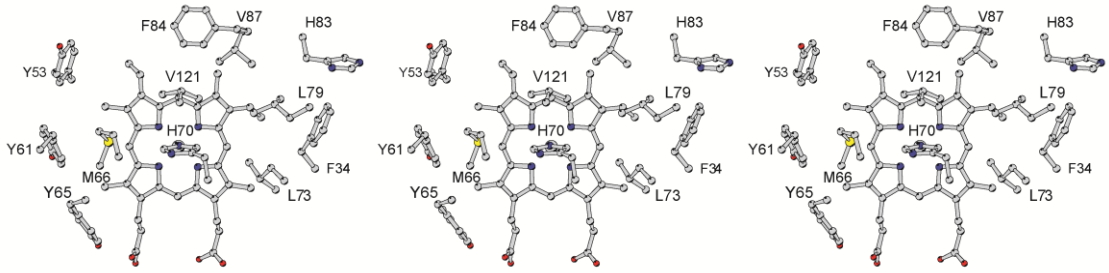

F84
V87
H83
Y53
V121
L79
H70
Y61
M66
F34
Y65
L73

Figure S5.7. Natural abundance ^{13}C band-selective ^{1}H-^{13}C HSQC (^{1}H-coupled) data indicate that the heme 4-vinyl group has reacted to generate a CαH-CβH_3 moiety in L79H/H117A GlbN-B. For orientation, this spectrum can be compared with Figure 5.5D of the text. Sample conditions were as in Figure S5.

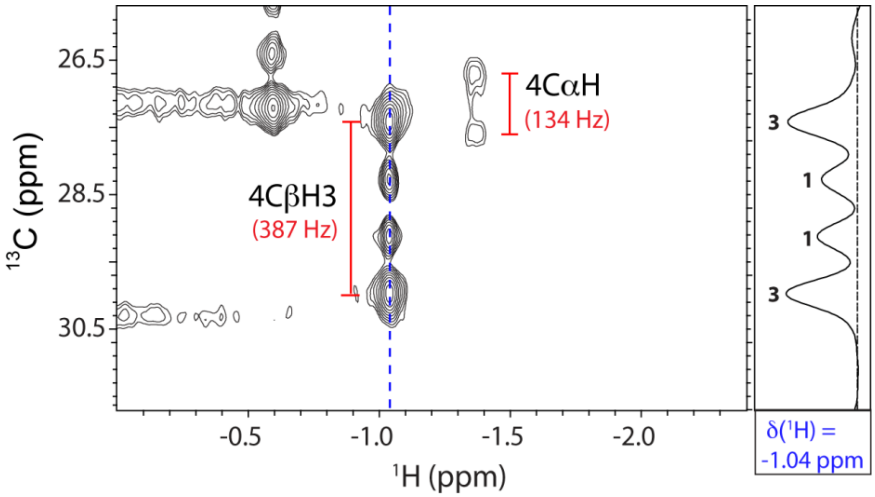


Table S5.1: Heme and reacted histidine assignments in ferric L79H/H117A GlbN-B

^{1}H chemical shifts were referenced through the water resonance at 4.76 ppm and extracted from NOESY data acquired at pH* 7.04, 99.9% D_2O and 298 K. ^{13}C Chemical shifts were extracted from ^{1}H-^{13}C HMQC data collected under the same conditions and referenced indirectly.[8]

Assignment	^{1}H	^{13}C
1-CαH$_3$	14.17	–27.2
2-CαH	16.42	28.4
2-vinyl CβH$_{cis}$	–8.21	–
2-vinyl CβH$_{trans}$	–7.45	–
2-CβH$_3$	n/a	n/a
α-meso	0.71	–
3-CαH$_3$	6.87	–16.8
4-CαH	–1.36	27.2
4-vinyl CβH$_{cis}$	n/a	n/a
4-vinyl CβH$_{trans}$	n/a	n/a
4-CβH$_3$	–1.04	28.7
β-meso	–0.78	–
5-CαH$_3$	24.91	–49.3
6-CαH	9.23	–
6-CαH’	7.45	–
6-CβH	0.20	108.1
6-CβH’	1.31	108.1
γ-meso	–0.17	–
7-CαH	3.23	–
7-CαH’	0.92	–

7-CβH	–1.18	81.6
7-CβH’	–1.58	81.6
8-CαH_3	10.28	–26.3
δ-meso	–	–
H79 Hδ2	2.88	117.5
H79 Hε1	5.70	137.6

Figure S5.8. UV-visible spectral changes following dithionite (DT) reduction of L79H GlbN. The ferric L79H GlbN starting material is represented by the grey dot-dot-dashed line. Within the manual mixing dead time (15–20 sec), the spectrum converts to an intermediate ferrous form (red solid line) with band maxima similar to the L79H/H117A double variant intermediate shown in Figure 5.4A of the text (red line). This rapid reduction phase is followed by a slow second phase in which the Soret, α, and β bands undergo a nearly monophasic blueshift to a second ferrous form. The slow phase likely corresponds to sequential vinyl modifications forming L79H GlbN-AB (purple spectrum collected 1 h after DT addition). The inset magnifies the Soret region for clarity. The right inset plots the Soret absorbance at various wavelengths as a function of time. Under the same conditions, these spectral changes occur over a timescale similar to that of the L79H/H117A GlbN (Figure 5.4A).

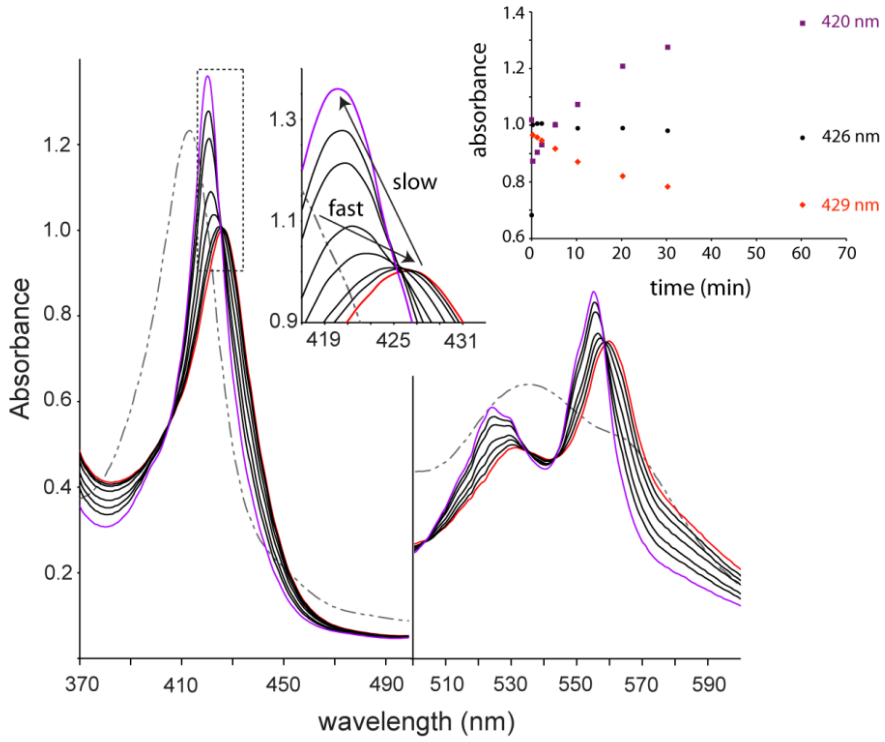


Figure S5.9. Identification of non-axial histidines within ferric L79H GlbN-AB. A spin system corresponding to a modified His117 (red labels) is observed in the ^{1}H-^{15}N lr HMQC spectrum of L79H GlbN-AB (see Figure 5.5A in the text for comparison with ferric L79H/H117A GlbN-B). Note the similar chemical shift behavior exhibited by His83, His33, His77 and reacted His79 in L79H GlbN-AB and L79H/H117A GlbN-B. Sample conditions: ~ 1 mM ^{15}N GlbN-AB, pH* 7.1, 99.9% D_2O, 313 K. The inset presents the same data at a lower contour level.

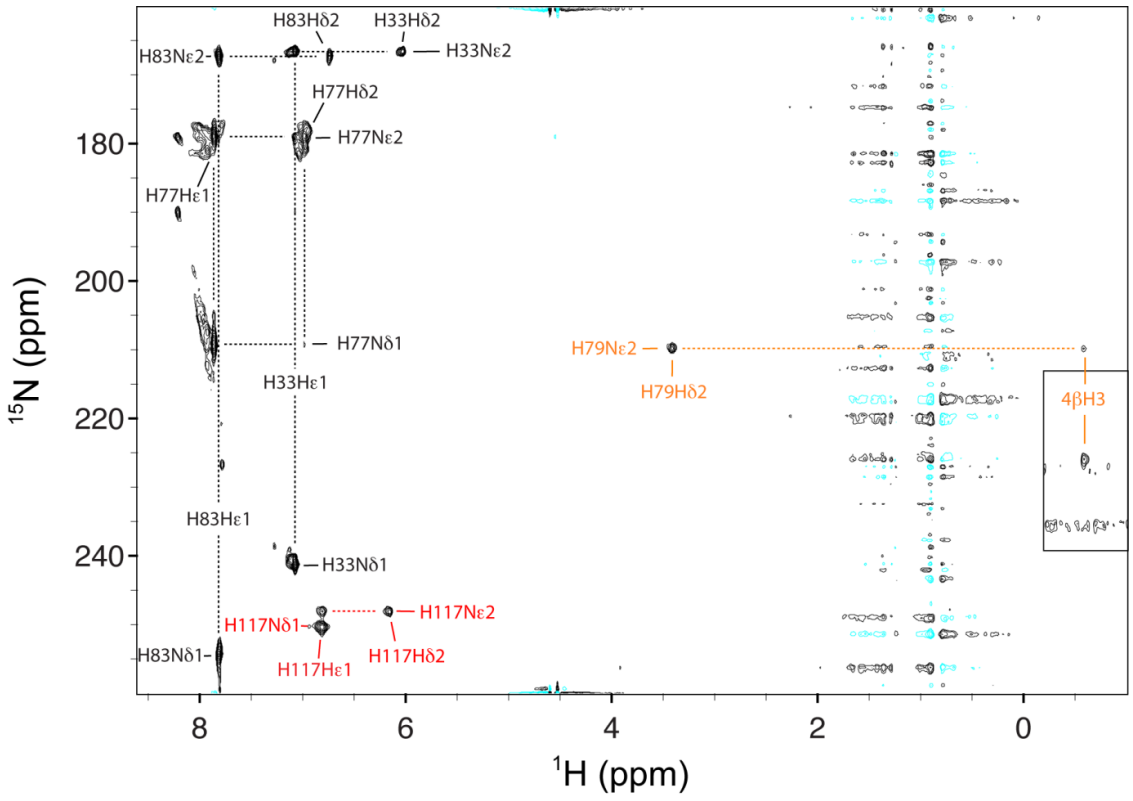


Figure S5.10. Resonances corresponding to heme Cα methyl groups and the proximal and distal coordinating His CβHs are identified in the natural abundance ^{1}H-^{13}C HMQC spectrum of ferric L79H GlbN-AB (far-upfield ^{13}C region). The heme 6-propionate CαH signals are also labeled. For comparison with GlbN-B, see Figure S5. Sample conditions: ~5 mM GlbN-AB, pH* = 7.61, 99.9 % D_2O, 313 K.

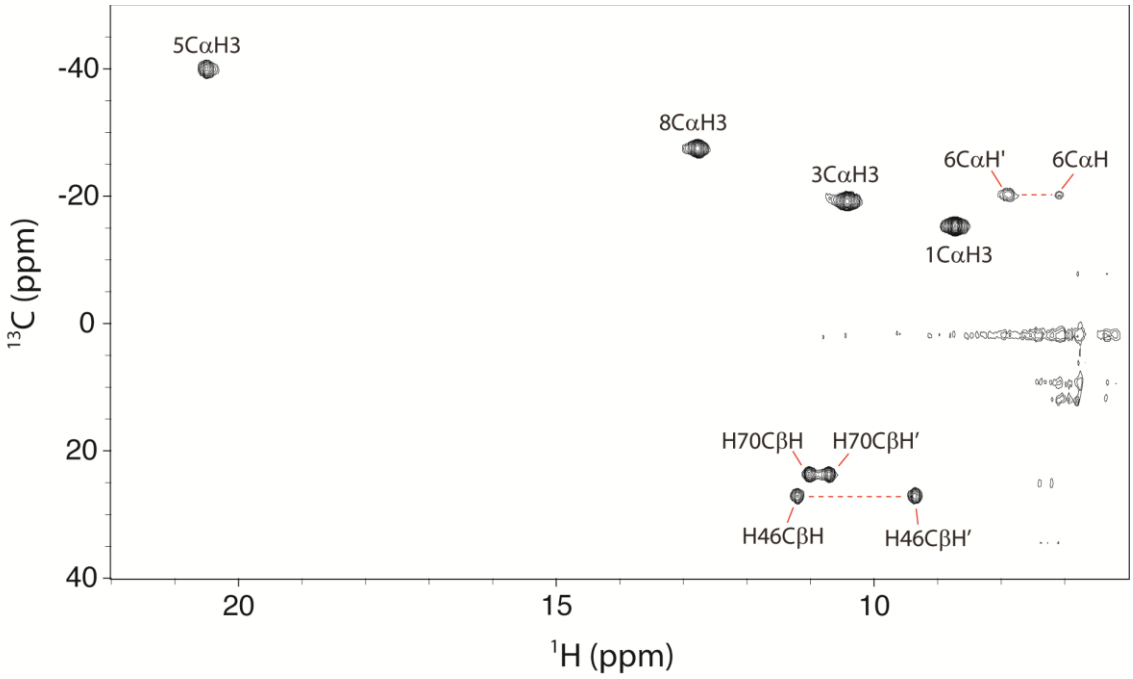


Figure S5.11. Natural-abundance ^{1}H-^{13}C HMQC methyl region comparison of (**A**) ferric L79H/H117A GlbN-B (green peaks, 298 K) with (**B**) ferric L79H GlbN-AB (black peaks, 313 K) with select chemical shift assignments indicated. Both GlbN-B and GlbN-AB exhibit the 4-CαH–CβH$_3$ unit, but an additional downfield-shifted methyl ($\delta(^{13}C)$ = 61.8 ppm) corresponding to a modified 2-CβH$_3$ group is observed in GlbN-AB (analogous to that detected in wild-type GlbN-A[9]). L79H GlbN-AB sample conditions were as in Figure S5.10; L79H GlbN-B sample conditions were as in Figure S5.

The differences in chemical shifts are due to both changes in structure and temperature effects. At room temperature, the proton shifts of the 4-CαH and 4-CβH$_3$ overlap in GlbN-AB. At 313 K, the peaks are resolved and a *J*-connectivity is observed in DQF-COSY data.

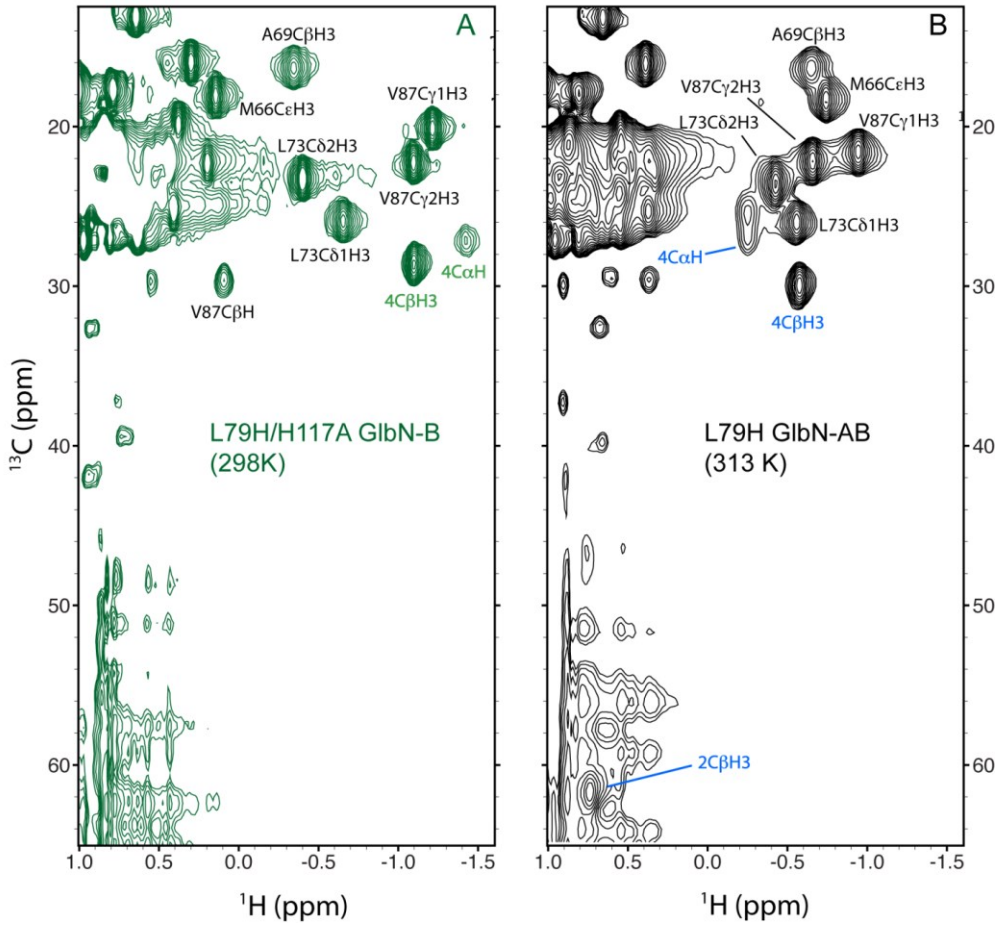


Figure S5.12. Natural-abundance ^{1}H-^{13}C HMQC aromatic region of ferric L79H GlbN-AB (red contours, 298 K) overlaid onto ferric L79H/H117A GlbN-B (grey contours, 298 K), with histidine assignments indicated. In both proteins, signals derived from the reacted histidines (His79, His117) are of weaker intensity, probably owing to efficient paramagnetic relaxation. Note that the His79 Cε1H signal in L79H GlbN-AB is weak but observable at 313 K (see outer box, which shows the relevant spectral region at low contour).

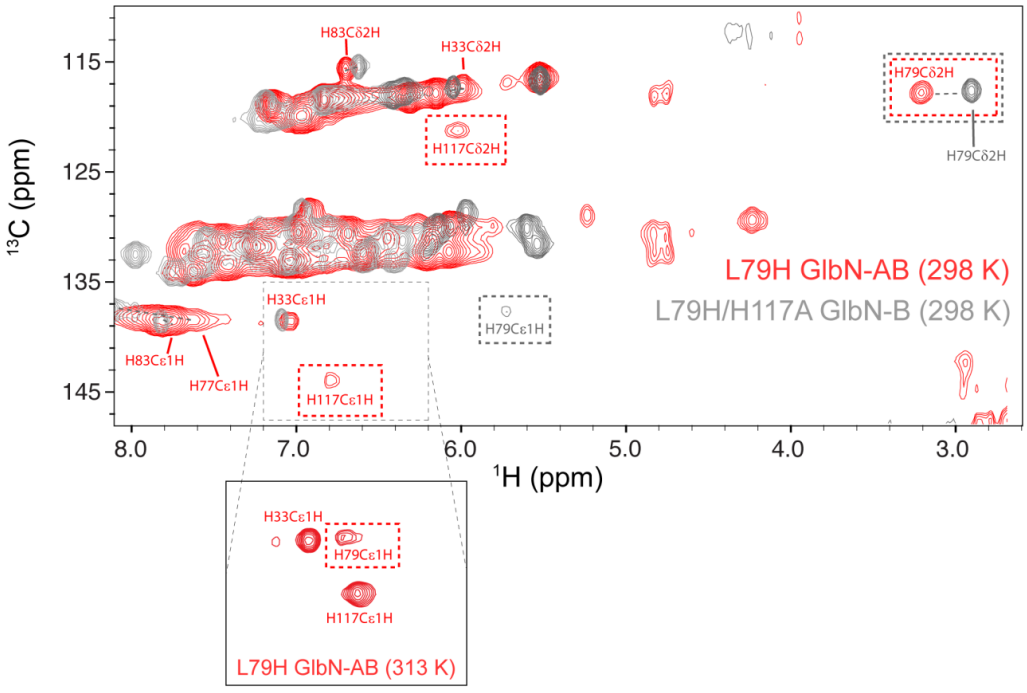


Table S5.2: Heme and reacted histidine assignments in ferric L79H GlbN-AB

^{1}H chemical shifts were referenced through the temperature-corrected water resonance at 4.58 ppm and extracted from NOESY data acquired at pH* 7.61, 99.9% D_2O and 313 K. ^{13}C Chemical shifts were extracted from ^{1}H-^{13}C HMQC data collected under the same conditions and referenced indirectly.[8]

Assignment	^{1}H	^{13}C
1-CαH$_3$	8.75	–15.2
2-CαH	1.92	6.1
2-vinyl CβH$_{cis}$	n/a	n/a
2-vinyl CβH$_{trans}$	n/a	n/a
2-CβH$_3$	0.74	61.8
α-meso	1.93	–
3-CαH$_3$	10.44	–19.0
4-CαH	–0.24	27.1
4-vinyl CβH$_{cis}$	n/a	n/a
4-vinyl CβH$_{trans}$	n/a	n/a
4-CβH$_3$	–0.57	30.1
β-meso	–	–
5-CαH$_3$	20.51	–39.8
6-CαH	7.15	–20.1
6-CαH’	7.88	–20.1
6-CβH	0.09	99.5
6-CβH’	0.64	99.5
γ-meso	–	–
7-CαH	4.01	–10.0
7-CαH’	1.81	–10.0

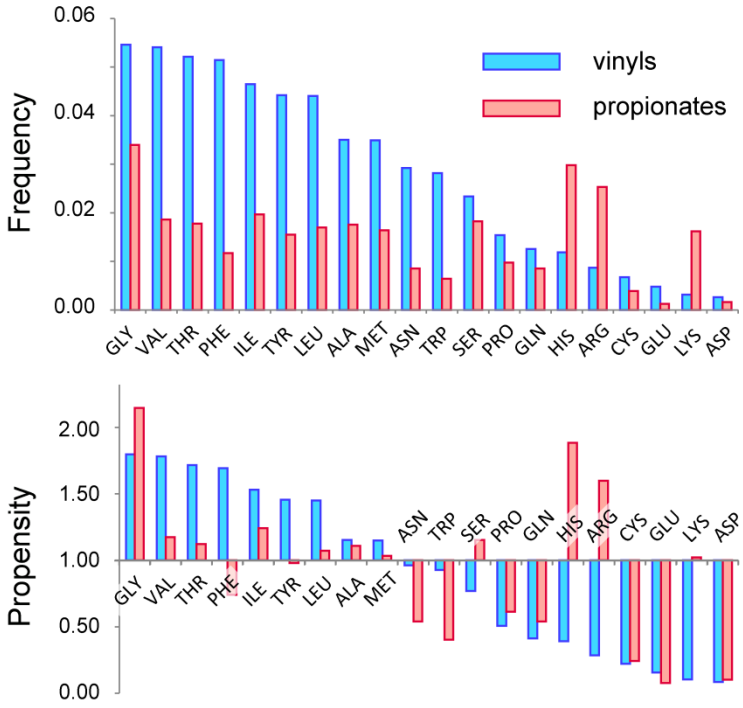
7-CβH	–0.41	81.3
7-CβH’	–0.83	81.3
8-C H_3	12.8	–27.3
δ-meso	3.08	–
H79 Hδ2	3.42	117.9
H79 Hε1	6.88	138.2
H117 Hδ2	6.16	121.2
H117 Hε1	6.84	143.2

Table S5.3: PDB ID of the 341 structures used in estimating the propensity of amino acids near a heme vinyl group. Only the heme domains were analyzed.

1A4E	1HLB	1OR4	1X3X	2IMQ	2VYW	3B98	3KS0	3QZ1
1A4F	1HLM	1OUT	1X46	2ISA	2VZM	3BCQ	3L4D	3R9B
1A6M	1I3D	1PA2	1X8Q	2ITF	2VZW	3BJ1	3LB2	3RGP
1APX	1IO7	1PEE	1X9F	2IVF	2W0B	3BKN	3LD6	3RIV
1ASH	1IT2	1PO5	1YHU	2J0P	2W31	3BOM	3LF5	3RTL
1B0B	1ITH	1Q16	1YMB	2J2M	2W3G	3BUJ	3LGN	3RWL
1B80	1IW0	1Q1F	1Z8O	2JJN	2WDQ	3CLI	3LXI	3S79
1BGP	1IYN	1Q4G	1ZBY	2KII	2WIY	3CQV	3M5Q	3S8F
1CG5	1IZO	1Q5D	1ZOY	2LHB	2WM5	3CZH	3MDM	3SDH
1CPT	1J77	1QGJ	1ZRT	2NNJ	2WTG	3D1K	3MGX	3SIK
1CQX	1JBQ	1QHU	256B	2NOX	2WX2	3D4X	3MK7	3SWJ
1CXY	1JEB	1QPA	2A06	2NW8	2WY4	3DAN	3MKB	3SWZ
1CYO	1JF4	1QPW	2BK9	2NWB	2XBK	3DAX	3MVC	3TBG
1D0C	1JFB	1QWL	2BKM	2O09	2XKI	3DBG	3MZS	3TGU
1D7B	1JL7	1R9O	2BMM	2O6P	2XKR	3DY5	3N5W	3TKT
1D8U	1KBI	1RTE	2BS2	2OIF	2XQ1	3E4W	3N9Y	3TYW
1DGF	1KQF	1S1F	2C0K	2ORT	2XYK	3E6I	3NC3	3UBC
1DK0	1LFK	1S67	2CIW	2OYY	2Y3Q	3E7G	3NER	3UOI
1DLW	1LHT	1S69	2D0T	2PBJ	2Y5N	3EJ6	3NN1	3UT2
1DLY	1LLP	1SCH	2D5X	2Q7A	2YIU	3EJB	3NT1	3VOL
1DNU	1M54	1SI8	2DC3	2Q8P	2Z36	3EMM	3NU1	3VP5
1DO9	1M7S	1SJ2	2DKK	2QSS	2Z3T	3FH9	3NV5	3VRB
1ECA	1M7V	1SK7	2E3B	2QU0	2Z6F	3FMU	3NXU	3VRG

1EMY	1M85	1SOX	2FDV	2R50	2ZBX	3FVB	3O0R	3VXJ
1EUE	1MBA	1SPG	2FKZ	2R79	2ZDO	3GAS	3O72	4A5G
1EW0	1MBS	1SY7	2FLQ	2R7A	2ZFB	3GDJ	3OFT	4AM5
1FFT	1MJ4	1TU9	2FMY	2R80	2ZS0	3GVY	3OO3	4CAT
1FHF	1MJT	1U4H	2G5G	2RAO	2ZWU	3GW9	3P3O	4DNJ
1FHJ	1MWC	1U5U	2GDM	2RFB	3A15	3H8T	3PC3	4DO1
1FSL	1MWV	1UB2	2H88	2RGZ	3A4G	3HCN	3PM0	4E6K
1FT9	1N40	1UED	2HBG	2UUQ	3ABA	3HDL	3PT8	4ESA
1GCV	1N45	1UX8	2HI4	2V7K	3ABB	3HF4	3Q3U	4FB2
1GVH	1N5U	1V4X	2HQ2	2VCF	3AQ5	3HX9	3QL6	4G1V
1GWE	1N97	1V9Y	2I96	2VE3	3AT5	3HYU	3QM9	4G2C
1GWI	1NGK	1WE1	2IBJ	2VEB	3AWM	3IA8	3QNS	4G7L
1H58	1NR6	1WMU	2IG3	2VHB	3AYF	3IBD	3QPI	4GQS
1H97	1ODO	1WOV	2IIZ	2VV6	3B4X	3IVY	3QQQ	4HRT
1HDS	1OJ6	1X3K	2IJ2	2VXH	3B6H	3K9V	3QQR	

Figure S5.13: Top: Frequency of occurrence of the 20 amino acids near heme vinyl Cα atoms (blue) or propionate Cα atoms (red). To be counted, the residue must have its Cα within 7 Å and Cβ within 6 Å of the target (glycines were used as such and treated with the Cα distance criterion only). The residues collected in this fashion are not necessarily in contact with the heme group. Bottom: propensity of occurrence as defined in the text. In both plots, the values were ranked according to the vinyl values.



Supporting Information References

(1) Nothnagel, H. J., Preimesberger, M. R., Pond, M. P., Winer, B. Y., Adney, E. M., and Lecomte, J. T. J. (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.

(2) Preimesberger, M. R., Pond, M. P., Majumdar, A., and Lecomte, J. T. J. (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.

(3) Kranz, R. G., Richard-Fogal, C., Taylor, J. S., and Frawley, E. R. (2009) Cytochrome *c* biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. *Microbiol. Mol. Biol. Rev. 73*, 510-528.

(4) Kraulis, P. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 946-950.

(5) Lecomte, J. T. J., Scott, N. L., Vu, B. C., and Falzone, C. J. (2001) Binding of ferric heme by the recombinant globin from the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochemistry 40*, 6541-6552.

(6) Antonini, E., Brunori, M., Caputo, A., Chiancone, E., Rossi-Fanelli, A., and Wyman, J. (1964) Studies on the structure of hemoglobin. III. Physicochemical properties of reconstituted hemoglobins. *Biochim. Biophys. Acta 79*, 284-292.

(7) Scott, N. L., Falzone, C. J., Vuletich, D. A., Zhao, J., Bryant, D. A., and Lecomte, J. T. J. (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.

(8) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135-140.

(9) Vu, B. C., Vuletich, D. A., Kuriakose, S. A., Falzone, C. J., and Lecomte, J. T. J. (2004) Characterization of the heme-histidine cross-link in cyanobacterial hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. *J. Biol. Inorg. Chem. 9*, 183-194

Chapter 6.

Introduction of a covalent histidine-heme linkage in a hemoglobin: A promising tool for protein engineering

Selena L. Rice[1], Matthew R. Preimesberger[1], Eric A. Johnson, and Juliette T.J. Lecomte

Reproduced with permission from J Inorg Biochem (2014) pii: S0162-0134(14)00251-7. doi: 10.1016/j.jinorgbio.2014.09.009. [Epub ahead of print]

[1]These two authors contributed equally

Author contributions:

MRP: designed research, performed NMR data acquisition and structural analysis, wrote the paper

SLR: designed research, prepared protein samples, performed UV-visible spectroscopic experiments and analysis, conducted ECL assay, wrote the paper

EAJ: developed and performed ECL assay

JTJL: designed research, analyzed data, wrote the paper

Abbreviations: 1D, one-dimension; CtrHb, heme domain of *Chlamydomonas eugametos* hemoglobin L1637, DSS, 4,4-dimethyl-4-silapentane-1-sulfonic acid; DT, sodium dithionite; ECL, enhance chemiluminescence; Hb, hemoglobin; lr, long-range; PAGE, polyacrylamide electrophoresis; PTM, post-translational modification; SDS, sodium dodecyl sulfate; WT, wild-type

Abstract

The hemoglobins of the cyanobacteria *Synechococcus* and *Synechocystis* (GlbNs) are capable of spontaneous and irreversible attachment of the *b* heme to the protein matrix. The reaction, which saturates the heme 2-vinyl by addition of a histidine residue, is reproduced in vitro by preparing the recombinant apoprotein, adding ferric heme, and reducing the iron to the ferrous state. Spontaneous covalent attachment of the heme is potentially useful for protein engineering purposes. Thus, to explore whether the histidine–heme linkage can serve in such applications, we attempted to introduce it in a test protein. We selected as our target the heme domain of *Chlamydomonas eugametos* LI637 (CtrHb), a eukaryotic globin that exhibits less than 50% sequence identity with the cyanobacterial GlbNs. We chose two positions, 75 in the FG corner and 111 in the H helix, to situate a histidine near a vinyl group. We characterized the proteins with gel electrophoresis, absorbance spectroscopy, and NMR analysis. Both T111H and L75H CtrHbs reacted upon reduction of the ferric starting material containing cyanide as the distal ligand to the iron. With L75H CtrHb, nearly complete (> 90%) crosslinking was observed to the 4-vinyl as expected from the X-ray structure of wild-type CtrHb. Reaction of T111H CtrHb also occurred at the 4-vinyl in a 60% yield, indicating a preference for the flipped heme orientation in the starting material. The work suggests that the His–heme modification will be applicable to the design of proteins with a non-dissociable heme group.

Introduction

Proteins requiring a *b* heme at their active site hold this essential molecule with multiple interactions. Protein–heme contacts, heme exposure to solvent, coordination of protein residues to the heme iron [1] and apoprotein structure [2] all contribute to the energetics of heme binding. As a consequence, the heme affinity of a given protein depends on solution conditions and properties such as the redox state of the iron [3] and the nature of its exogenous ligands [4, 5]. These last two features can dominate heme affinity. For example, soluble guanylate cyclase, which has an axial histidine ligand, loses heme upon oxidation from the ferrous to the ferric state [6], whereas Dap1p, which has an axial tyrosine ligand, exhibits enhanced heme loss upon reduction from the ferric state to the ferrous state [7]. Various enzymes experience heme loss upon exposure to nitric oxide ($NO^{\bullet}$) [8].

In part because of the toxicity of free heme, enzymes, gas sensors, and gas or small molecule carriers that use a *b* heme have evolved to maintain sufficient affinity as they visit the various ligation and redox states required by their physiological role. Uncontrolled heme loss occurs rarely. However, the structural features dictating heme affinity and reactivity are intimately intertwined. Thus, heme retention can be problematic when modifying native heme proteins [9, 10] or designing artificial ones [11]. In those instances, optimization for the desired chemical properties may compromise heme binding and vice versa.

Covalent attachment of the heme to the protein matrix offers a versatile

solution to heme loss. Such post-translational modification (PTM) grants several advantages aside from heme retention, notably correct heme positioning and increased thermodynamic stability [12]. In addition, the PTM allows for a departure from the typical composition of a *b* heme cavity, i.e., the introduction of polar residues and weak axial ligands. The heme group, shown in Figure 6.1a, offers several possible points for crosslinking to the protein, the most common of which are the vinyl substituents. In *c* cytochromes, reductive reaction with cysteines produces thioether bridges [13]. The formation of these linkages in the cell is controlled by dedicated enzymatic machineries that ensure the correct localization of the apoprotein prior to heme binding and modification [14, 15]. This complicated process can be harnessed to prepare artificial cytochromes [16, 17], and thus far the molecular biology procedure to producing the thioether linkages surpasses non-enzymatic methods [18, 19]. For purposes of creating new proteins with non-dissociable heme and studying the effects of heme crosslinking, a chemical approach avoiding the use of cysteines would be advantageous. We have shown previously that the globins from two cyanobacteria (GlbNs from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002) can undergo enzyme-free heme attachment to a histidine, and we now explore the suitability of this reaction for protein engineering applications.

The heme modification in GlbN is the alkylation of histidine Nε2 by a heme vinyl Cα as shown in Figure 6.1b.

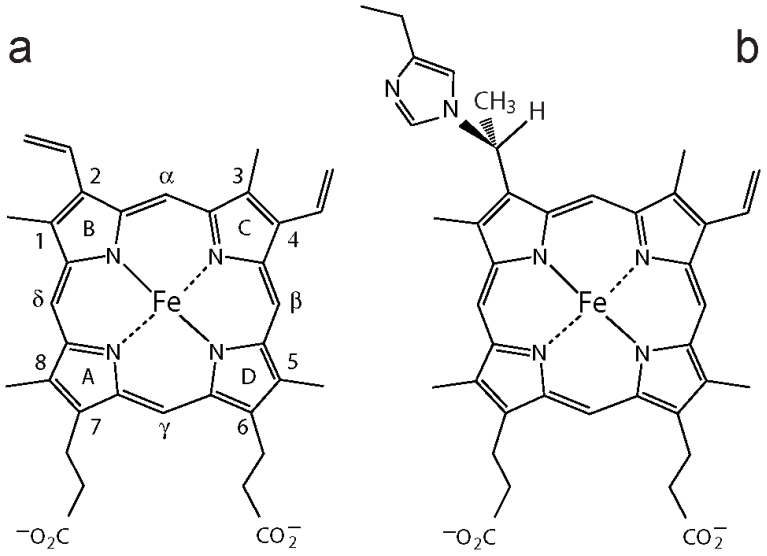


Figure 6.1. (a) The structure and nomenclature of the *b* heme (iron-protoporphyrin IX). Heme orientational isomerism involves a 180° rotation about the α–γ axis. (b) The heme PTM present in GlbN. His117, in the H helix, adds to the 2-vinyl group. Addition can also occur to the 4-vinyl group when the heme is in the flipped orientation [20, 21].

Several studies have defined the conditions under which the reaction occurs and have amounted to a plausible mechanism [20, 21]. Following reduction to the ferrous state, the key steps are the protonation of the heme vinyl Cβ followed by (or in concert with) a nucleophilic attack by a neutral histidine. In wild-type (WT) ferrous GlbNs, the modification reaches completion in a few seconds at neutral pH and is accelerated by decreasing pH as long as acid denaturation is avoided. We have monitored the reaction in variants of GlbN and proposed that the main determinant of linkage formation is sterics, i.e., whether a productive orientation of the vinyl and histidine ring can be adopted. This is evident in the preparation of a protein containing the WT crosslink (H117 to heme 2-vinyl) and an engineered crosslink (L79H to heme 4-vinyl) [22].

Additionally, we have noted that certain ligands to the ferrous state such as the π-acids CO [21], $NO^{\cdot}$ and O_2 (see Chapter 4) inhibit the reaction, presumably because of their propensity for a redistribution of electron density of the ferrous heme (e.g., formation of Fe(III)–$O_2^{\cdot-}$ from Fe(II)–O_2, [23]) disfavoring vinyl protonation.

The consequences of the linkage for the reactivity of the heme group have been under investigation as well. A variant of GlbN unable to form the PTM loses the ferrous heme more readily than the ferric heme, which has led to the proposition that the linkage is a physiological necessity [24]. In WT GlbNs, the PTM enhances substantially the thermal stability of the holoprotein [25]. On the other hand, the PTM appears to have only moderate effects on the structure of GlbN and its chemical properties such as electron self-exchange rate, redox potential [26], susceptibility to H_2O_2-mediated degradation [21], backbone dynamics [25, 27], and response to pressure [28]. In view of these observations, the histidine–heme linkage appears as a mostly benign modification, well suited for protein engineering purposes. However, the general utility of the linkage must be demonstrated by introduction into a protein different from the two original cyanobacterial globins.

The heme domain of *Chlamydomonas eugametos* LI637, CtrHb, was chosen as a validation protein for several reasons. (1) CtrHb has been the focus of spectroscopic [23, 29, 30] and structural [31] studies. The WT protein is therefore already well characterized. (2) Sequence identity with *Synechococcus* and *Synechocystis* GlbNs (Figure 6.2) is a modest ~46% with differences distributed throughout the structure, including the heme cavity. (3) In GlbN,

the WT reactive histidine is in the H helix (His117, at helical position "H16") and the engineered reactive histidine is in the FG corner (His79). A superimposition of the three-dimensional structures of cyanomet (ferric state with cyanide bound) CtrHb and post-translationally modified cyanomet *Synechocystis* GlbN (hereafter GlbN-A) (Figure 6.2) reveals that CtrHb has two-residue deletions near the 79 and 117 sites of GlbN. (4) Unlike both GlbNs, CtrHb does not have a histidine that can reversibly coordinate the iron on the distal side.

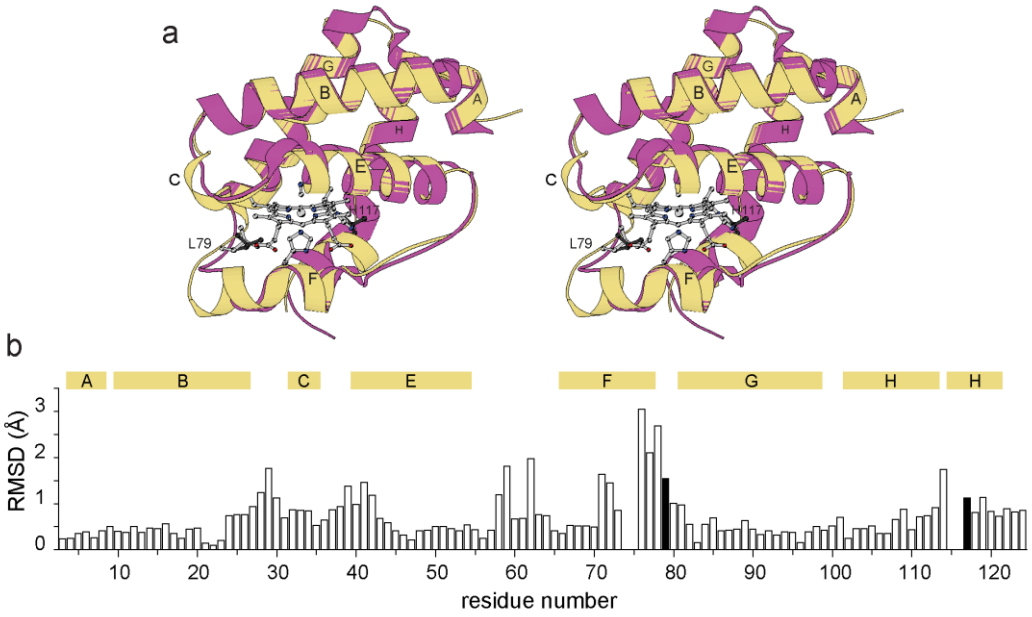


Figure 6.2. Structural comparison of cyanomet *C. eugametos* L1637 CtrHb (PDB ID: 1DLY) with cyanomet *Synechocystis* sp. PCC 6803 GlbN-A (PDB ID: 1S69). (a) Stereo view of the superimposition of the two structures. *Synechocystis* GlbN-A is in yellow with L79 and H117 shown in light gray sticks. Dark gray sticks mark L75 and T111 in CtrHb. (b) Cα rmsd for the superimposition shown in (a). The numbering is that of GlbN. The black vertical bars indicate the position of L79 and H117. The secondary structure of GlbN is indicated with horizontal bars above the rmsd values.

At neutral pH and in the ferrous state, CtrHb is primarily a 5-coordinate species [29]. Thus, CtrHb has distinct heme pocket composition, local backbone geometry, and iron coordination that render it a rigorous test case for the histidine–heme modification of the type depicted in Figure 6.1b.

In this chapter, we inspected the reactivity of CtrHb variants with a histidine facing either the heme 2- or 4-vinyl group on the proximal side of the heme. We demonstrate that the histidine–heme modification can be successfully implanted in CtrHb. Although spontaneous crosslinking does not occur upon reduction to the ferrous state, the addition of cyanide prior to reduction promotes the reaction. The results support that histidine alkylation may be a generalizable reaction for the production of proteins containing a non-dissociable heme.

Materials and Methods

Plasmid construction and protein production

The gene coding for WT CtrHb (UNIPROTKB Q08753, residues 44–164, "H19" in reference [32]) was synthesized with 5'-*Nde*I and 3'-*Bam*HI restriction sites and inserted in a pJexpress 404 vector by DNA 2.0 (Menlo Park, CA). Because expression levels were low with this plasmid, WT and mutated CtrHb genes were subsequently cloned into a pET3c vector (Novagen, Inc., Madison, WI). Site directed mutagenesis was performed to introduce the T111H or L75H

replacement using the QuikChange (Qiagen, Valencia, CA) protocol provided by the manufacturer and primers purchased from Integrated DNA Technologies (Coralville, IA). Restriction enzymes and ligase were from New England Biolabs (Ipswich, MA), and sequencing was performed by GENEWIZ, Inc. (South Plainfield, NJ).

E. coli BL21(DE3) cells (New England Biolabs) were used for overexpression of CtrHb. Cells were grown on M9 medium supplemented with ampicillin (Sigma-Aldrich, St. Louis, MO or Research Products International, Corp., Mount Prospect, IL). Induction was achieved by addition of 1 mM dioxane-free isopropyl β-D-1-thiogalactopyranoside (Santa Cruz Biotechnologies, Santa Cruz, CA). For uniformly labeled ^{15}N protein, $^{15}NH_4Cl$ (Sigma-Aldrich) was used as the sole nitrogen source.

CtrHb purification

WT and T111H CtrHbs were purified from inclusion bodies as described previously for *Synechocystis* GlbN [33]. The protocol calls for treatment of the harvested and washed inclusion bodies with 8 M urea, followed with apoprotein refolding by passage over a gel filtration column (Sephadex G-50 Fine, Sigma-Aldrich) equilibrated with 50 mM Tris, 50 μM EDTA pH 8.0 at 4 °C. Further purification was performed by passage through an anion exchange column (DEAE Sephacel, Sigma) either before or after addition of bovine hemin chloride (Fe(III)-protoporphyrin IX, Sigma/Alfa Aesar) to generate the ferric holoprotein. Elution was achieved with a 0–0.4 M NaCl gradient. Although WT and T111H CtrHbs could be lyophilized and resuspended without loss or damage, such was not the case for L75H CtrHb, and solutions of this protein were used

immediately upon purification. The yields were ~15 mg/L for L75H CtrHb, ~39 mg/L for T111H CtrHb, and ~50 mg/L for WT CtrHb when the cells were grown in minimal medium.

Apoprotein concentrations were estimated using a calculated extinction coefficient of 10 mM^{-1} cm^{-1} at 280 nm. The optical spectra of the WT CtrHb in the ferric, ferrous, cyanomet, and ferrous cyanide states were comparable to those reported in the original work of Guertin and colleagues [29, 32, 34] (Supporting Information Table S6.1). The T111H and L75H replacements affected the spectral properties to varying degrees, with minimal effect on the cyanomet state (Supporting Information Figure S6.1). Ferric holoprotein concentrations were determined on a per-heme basis with an extinction coefficient of 117 mM^{-1} cm^{-1} at 410 nm at neutral pH, back-calculated from the published WT cyanomet coefficient [34].

NMR sample preparation

Lyophilized protein was resuspended in either 100 or 250 mM phosphate buffer in H_2O or 100 mM phosphate in D_2O. 10% v:v D_2O was added to the H_2O samples. Protein concentration varied from ~0.6 mM to ~3 mM (final volume of ~300 μL). Sample pH ranged between 7.1 and 7.5. Iron reduction was performed by adding 3- to 13-fold excess sodium dithionite (DT, Sigma-Aldrich). Alternatively, freshly purified protein was concentrated, exchanged in the appropriate buffer, and used without lyophilization. Cyanomet CtrHb was generated by adding a 5-fold excess of potassium cyanide to ferric holoprotein solutions.

Heme modification

Stocks of ferric WT (as a negative control), T111H, and L75H CtrHbs were prepared either from lyophilized protein, or, in the case of L75H CtrHb, from freshly purified protein. A 5-fold excess of cyanide was added to each sample, followed by 20–30 min incubation at room temperature to ensure saturation of the heme site. An excess of DT (200- to 500-fold for optical samples or 3- to 13-fold for NMR samples) was then added to the solutions. Cyanomet CtrHb samples were incubated for ~3–8 h (WT and T111H) or ~0.5–2 h (L75H) and then oxidized by exposure to air or potassium ferricyanide. Aliquots of the total samples were saved at different stages of the process for optical characterization.

Hemochromogen and enhanced chemiluminescence assays

Reacted CtrHb samples were analyzed by the hemochromogen assay [35, 36] for rapid assessment of *b* heme modification. Aliquots were taken from samples used for NMR analysis, diluted to reach an absorbance between 0.3 and 0.6 in the α/β region (500–600 nm), and treated as previously described [20] with basic pyridine followed by DT reduction. Absorbance measurements were performed on an AVIV Spectrophotometer 14DS UV-Vis (AVIV Biomedical Inc., Lakewood, NJ) as described previously [22].

Covalent heme attachment was detected by enhanced chemiluminescence (ECL) [37, 38]. Samples of CtrHb were prepared for DT reduction in pair, one with cyanide present (referred to as "KCN+") and the other without cyanide (referred to as "KCN–"). CtrHb samples were also prepared with 10-fold excess

imidazole (Sigma) or 10 mM azide (Sigma) as alternatives to cyanide as the distal iron ligand. *Synechocystis* GlbN treated with excess DT to generate GlbN-A [39] was used as a covalently-linked positive control.

The ECL assay relies on the peroxidase activity of the heme group. This activity is inhibited by cyanide and high concentrations of either imidazole or azide. Prior to ECL detection, reduced samples containing these ligands were treated by passage over a ~1.3 cm^3 DEAE column. This procedure did not eliminate all exogenous ligand, and the ECL results for these samples are therefore only qualitative.

Electrophoresis was performed with Mini-PROTEAN® 16.5% Tris-Tricine Precast Gel (BIO-RAD, Hercules, CA). The protein samples (5 μM) were denatured by boiling for 10 min with SDS but in the absence of β-mercaptoethanol. After separation, the gels were thoroughly rinsed and exposed to the ECL reagent (Immobilon Western Chemiluminescent HRP Substrate, Millipore, Billerica, MA) prepared per manufacturer's instructions. The ECL reaction was allowed to proceed for 7 min before applying the film (Carestream® Kodak® BioMax® light film, Sigma-Aldrich, St. Louis, MO). Exposure times were 30 s for samples containing no exogenous ligand and 10 s, 30 s, or 7 min for samples partially stripped of exogenous ligand. After the ECL procedure, gels were stained with Coomassie Blue for protein detection.

Reaction kinetics monitored by absorbance spectroscopy

Samples of WT, T111H, and L75H CtrHb containing ~4 μM protein were prepared in 100 mM phosphate buffer pH 7.0 from concentrated stocks (lyophilized WT and T111H) or freshly prepared protein (L75H). Five-fold excess

KCN was added to the samples, and CN^- binding was monitored via electronic absorption spectroscopy. Approximately 500-fold excess DT was added to cyanide-saturated samples and reduction was monitored as above. Time course experiments were performed on a Varian Cary 50 Bio UV-Vis spectrophotometer with acquisition parameters similar to those reported previously [22].

NMR spectroscopy

NMR data were acquired on Bruker AVANCE or AVANCE-II spectrometers operating at a ^{1}H Larmor frequency of 600.13 or 600.53 MHz, respectively, and each equipped with a cryoprobe. Reduction and reaction progress of cyanomet CtrHbs were monitored by ^{1}H 1D NMR spectroscopy with water presaturation. Reactants and products were additionally studied using ^{1}H-^{1}H NOESY, DQF-COSY, TOCSY, and natural abundance ^{1}H-^{13}C HMQC spectra [40-42]. These data were sufficient to assign most heme resonances and locate many conserved heme pocket residues of each cyanomet CtrHb. Labile protons were identified by acquisition of the same data in 99.9% D_2O. ^{1}H-^{15}N HSQC and long-range (lr) histidine selective ^{1}H-^{15}N HMQC spectra [43] were used to confirm the assignments of the engineered and native histidines within WT, T111H and L75H cyanomet CtrHbs [21, 22]. ^{1}H chemical shifts were referenced to DSS indirectly through the water resonance (4.76 ppm at 298 K); ^{13}C and ^{15}N chemical shifts were referenced as described in [44]. NMR data were processed with TopSpin 2.1 or NMRPipe [45]. Spectral analysis was conducted with Sparky 3 [46].

Results

Choice and preparation of CtrHbs

The selection of sites for histidine introduction was guided by inspection of the CtrHb and GlbN structures (Figure 6.2). A significant proportion of CtrHb molecules contain the heme group in a "flipped" orientation that interchanges pyrroles A and B with pyrroles D and C, respectively [31]. These two forms of the protein are referred to as "major" and "minor" heme rotational isomers according to their relative population at equilibrium. The X-ray structure (PDB ID: 1DLY [31]) is that of the major isomer. The two heme conformations effectively double the number of vinyl geometries available for histidine addition. Minimal modeling (amino acid replacement in 1DLY and optimization with UCSF Chimera [47]) suggests that a histidine at position 111, the closest equivalent to H117, can reach the heme 2-vinyl group as positioned in the major heme rotational isomer. The favored rotameric state of H111 points the ring away from the heme toward solvent and generates few clashes. Another potentially reactive site, facing the 4-vinyl group of the major heme rotational isomer, was identified at position 75 as the equivalent of H79 in L79H GlbN. At this position, clashes are predicted regardless of histidine rotameric state. Sites 75 and 111 were chosen despite their suboptimal geometric properties for lack of better candidates.

The apoproteins of WT CtrHb and its L75H and T111H variants aggregated into inclusion bodies during *E. coli* cell growth. Purification included unfolding in urea, refolding over a size-exclusion column, and reconstitution with ferric

heme. For WT and T111H CtrHb, stoichiometric amounts of heme were added prior to a final purification step by anion exchange chromatography. L75H CtrHb had low ferric heme affinity, and passage through this last column resulted in substantial heme loss. To minimize this problem, heme was titrated to a slightly sub-stoichiometric level as the last step of holoprotein preparation, after apoprotein purification by anion exchange chromatography. It is likely that the presence of H75 disrupts the hydrophobic packing near the heme and interferes with heme binding. This was also the proposed explanation for the low heme affinity of the L79H variants of *Synechocystis* GlbN [22].

NMR characterization of cyanomet CtrHbs

Ferric CtrHbs to which KCN was added to generate the cyanomet state (S = 1/2) yielded high quality NMR spectra. Figure 6.3a, b, and c show the ^{1}H 1D spectrum of cyanomet WT CtrHb and the T111H and L75H variants. In all cases, two sets of resonances are detected per protein. These arise from the heme rotational isomers, which typically exchange slowly on the chemical shift time scale [48, 49]. Assignments of most heme resonances of both isomers in WT CtrHb and the T111H and L75H variants were obtained by established procedures [50]. Portions of spectra with representative through-bond and through-space connectivities are shown in Supporting Information Figure S6.2–S11. Heme and protein assignments confirm that in WT, T111H, and L75H cyanomet CtrHbs, the major isomer corresponds to approximately 60–70% of the sample and has the heme positioned as in the crystallographic WT structure (1DLY [31], Supporting Information Figure S6.12A–F). NMR data also support that key features of the WT CtrHb heme pocket, such as the presence

of Tyr20 (B10) as an H-bond donor to bound cyanide (Supporting Information Figure S6.5) and the participation of Gln41 (E7) and Gln45 (E11) in an extended H-bond network (Supporting Information Figure S6.6) [31], are largely unaltered by the T111H and L75H replacements (data not shown). Heme ^{1}H chemical shifts of WT and variant cyanomet CtrHbs are listed in Table 6.1.

Inspection of Figure 6.3 shows that at neutral pH the histidine replacement at position 111 perturbs the spectrum more profoundly than at position 75.

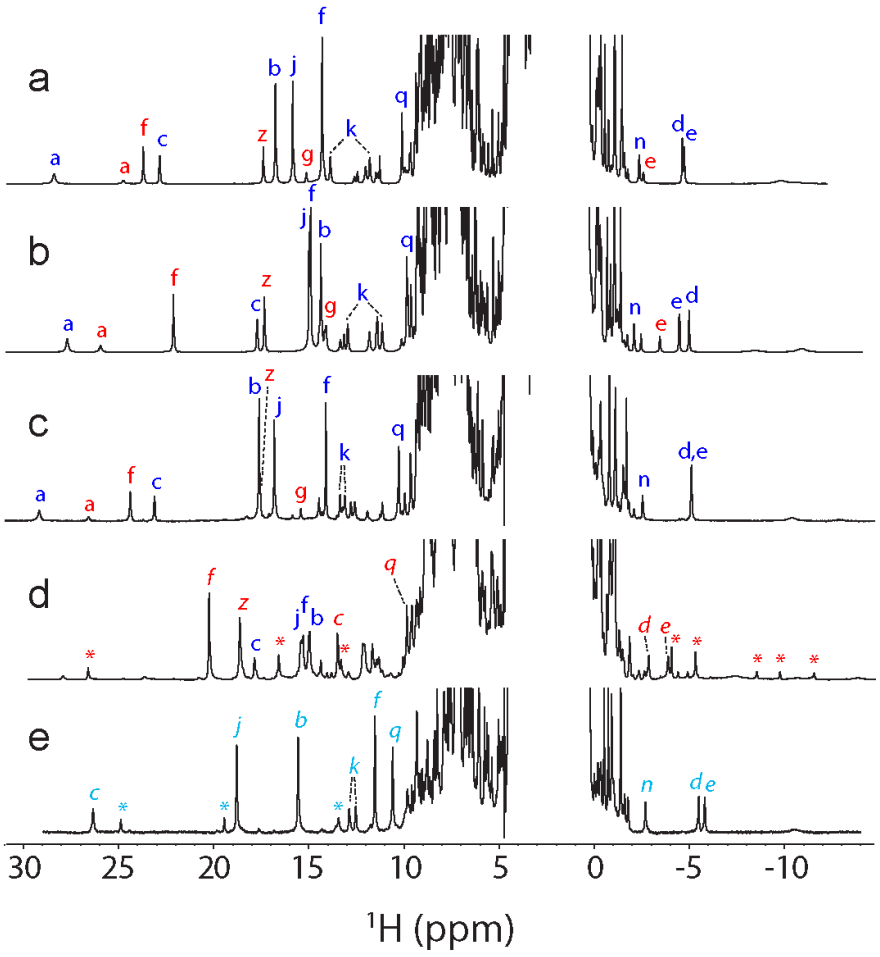


Figure 6.3. Comparison of WT, T111H, and L75H cyanomet CtrHb ^{1}H NMR spectra. (a–c). Spectra collected prior to DT treatment. The major and minor heme orientational isomers are indicated with blue and red letters, respectively. (a) Wild-type cyanomet CtrHb (72% major, 28% minor). (b) T111H cyanomet

CtrHb (67% major, 33% minor). (c) L75H cyanomet CtrHb (62% major, 38% minor). (d–e). Spectra collected after DT treatment and reoxidation. (d) T111H cyanomet CtrHb-A[4] product mixture. The predominant form (~60%, H111 reaction with the 4-vinyl group) is marked with red italics. The spectrum also contains 25% of the unreacted major isomer (starting material, b, blue letters), and a small amount of a third species (marked with red *). (e) L75H cyanomet CtrHb-B product (~90%, H75 reaction with the 4-vinyl group). A second product species (10%, likely H75 reaction with the 2-vinyl group) is marked with cyan *. Peak labeling: a, Tyr20 OηH; b, heme 1-CH_3; c, heme 2-vinyl Hα; d, e, heme 2-vinyl $H\beta_{cis}$, $H\beta_{trans}$; f, heme 3-CH_3; g, heme 4-vinyl Hα; j, heme 5-CH_3; k, heme 6-propionate Hα, Hα'; n, heme 7-propionate Hβ; q,Tyr20 CεHs; z, heme 8-CH_3.

In the major heme isomer of T111H CtrHb, the substituents of the B pyrrole undergo a relatively large chemical shift perturbation, apparently related to the reorientation of the 2-vinyl group from a "*cis*"-like conformation in WT CtrHb (i.e., with Cβ out of the heme plane and pointing toward the 3-CH_3, Supporting Information Figure S6.2) to a "*trans*"-like conformation (i.e., Cβ pointing toward the 1-CH_3, Supporting Information Figure S6.7). NOEs are detected between H111 Hδ2 and the heme 1-CH_3, whereas H111 Hε1 makes dipolar contact with F48 Hζ. The geometry is consistent with an outward position of the histidine Nε (g^- rotameric state) and forecasts a low probability for reaction with the 2-vinyl Cα. In the minor isomer, H111 Hδ2 contacts the 4-vinyl $H\beta_{cis}$ and exhibits the same NOE to F48, in support of a similar conformation for H111 in both major and minor isomers. ^1H-^{15}N lr-HMQC data collected at neutral pH do not reveal the H111 spin system (Supporting Information Figure S6.13), presumably because of broadening due to exchange between neutral and protonated states.

The heme chemical shifts of cyanomet L75H CtrHb form a pattern similar to that of WT CtrHb. The side chain of H75 in the major isomer was identified by

comparison with WT spectra; the unusual Hε1 and Hδ2 chemical shifts (9.30 ppm and 6.15 ppm, respectively) reflect paramagnetic and heme ring current contributions. NOEs to the heme and proximal histidine further define the position of the imidazole ring (not shown). Observed contacts are between H75 Hε1 and H68 (proximal histidine) Hβs, and between H75 Hδ2 and the heme 3-CH_3 and 4-vinyl. The 4-vinyl has a *trans*-like orientation (i.e., Cβ pointing toward the 3-CH_3, Supporting Information Figure S6.10, and analogous to the WT and T111H variant, Supporting Information Figure S6.2 and S7, respectively). The data are consistent with a g^- rotameric state, which orients the imidazole ring toward the heme and may favor reaction. Comparison of WT and L75H ^{1}H-^{15}N lr-HMQC spectra confirms the histidine assignments and further indicates that, at neutral pH, the engineered H75 side chain is in the Nε2H tautomeric state (Supporting Information Figure S6.13).

Heme modification in T111H and L75H CtrHb

The first attempt at producing CtrHbs with covalently attached heme was based on results obtained with *Synechococcus* and *Synechocystis* GlbNs. For these proteins to undergo the reaction, it suffices to reduce the heme iron [39]. Ferric T111H and L75H CtrHb were incubated with excess DT for variable periods of time, up to ~2 h, oxidized with potassium ferricyanide, and then exposed to an excess of KCN for NMR comparison with their respective untreated form. Absence of reaction is hinted at by unchanged optical spectra and made apparent by ECL assay of denaturing gels, which detects heme only in the dye front of DT-treated CtrHbs (Figure 6.4a, lanes 1–3). In this assay *Synechocystis* GlbN-A shows a strong signal at the protein's molecular weight

(Figure 6.4a, lane 4), as expected. The NMR spectra of the treated CtrHbs are not presented as they are identical to those of the non-reduced starting material.

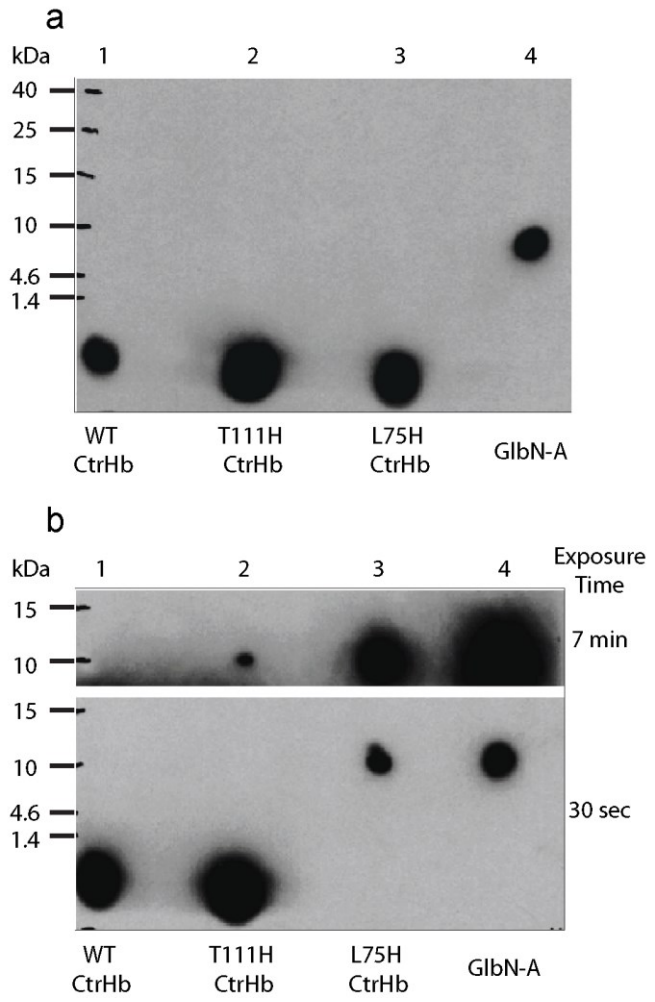


Figure 6.4. ECL SDS-PAGE detection of covalently bound heme in CtrHbs. (a) Ferric samples were reduced with 2 mM DT for ~30 min then re-oxidized with air before being subjected to electrophoresis. The image was obtained with a 30-s exposure. The heme of WT CtrHb and the T111H and L75H variants (lane 1-3) migrates at the dye front. WT GlbN-A (lane 4) is included as a positive control for covalent linkage. (b) Ferric WT and variant CtrHbs were first saturated with cyanide and then reduced for ~30 min. The figure shows images of film that was exposed for either 30 s or 7 min. The heme from WT CtrHb (lane 1) migrates with the dye front. The T111H CtrHb lane (lane 2) shows incomplete heme attachment. The longer exposure time was necessary to

observe the faint crosslinked T111H CtrHb band. The L75H CtrHb lane (lane 3) shows complete crosslinking as in the GlbN-A control (lane 4).

Failure to react can be due to various factors, including heme dissociation, incorrect histidine–vinyl geometry or unfavorable electron distribution of the porphyrin ring [22]. Binding of a suitable (permissive) ligand on the distal side may alleviate some of these problems. We therefore chose to initiate reduction from the stable and well-folded cyanomet state. The resulting ferrous–cyanide complex has a distinctive optical spectrum [34] (Supporting Information Figure S6.1A-D), and although it is not expected to be as stable as its ferric counterpart [51], it may persist long enough to assist in the reaction.

Reduction of cyanomet WT CtrHb with 500-fold excess DT at pH 7.0 was monitored optically (Fig. 5a). Over a period of 40 min, the Soret maximum shifts from 416 nm to 431 nm and sharpens considerably. The end spectrum displays resolved α and β bands and corresponds to that of the cyanide adduct of ferrous CtrHb reported by Milani and coworkers [34] (Supporting Information Table S6.1). The transition displays isosbestic points, which is an indication of an apparent two-state process. At longer incubation times, a slow but nearly uniform decay in absorbance is observed and attributed to DT-mediated heme bleaching. Reduction of the WT protein without exogenous ligand but under otherwise identical conditions is considerably faster (complete within 2 min) and leads to a broad Soret band at neutral pH (Supporting Information Figure S6.1B). The WT CtrHb behavior was used to gauge the reaction of the variants.

Cyanomet T111H CtrHb shows a response similar to that of cyanomet WT CtrHb, albeit on a slightly faster time scale for reduction (complete within 25 min under the same conditions as WT CtrHb). In addition, after 2 h of incubation with excess DT, the Soret maximum has moved from 430 nm to 428 nm (Fig. 5b, red to orange traces). Similar changes are observed for the α and β bands. These blue shifts suggest a decreased degree of conjugation of the macrocycle as caused by crosslink formation.

Reduction of cyanomet L75H CtrHb gives distinct results (Fig. 5c). Addition of DT leads to a rapid (< 2 min) sharpening of the Soret maximum with a shift to 426 nm. The intensity is at first lower than expected based on the WT extinction coefficient and increases over time while remaining at 426 nm. The final spectrum (Supporting Information Figure S6.1D) differs from that of the ferrous protein (Supporting Information Figure S6.1B) and is inconsistent with a simple loss of cyanide from the reduced state.

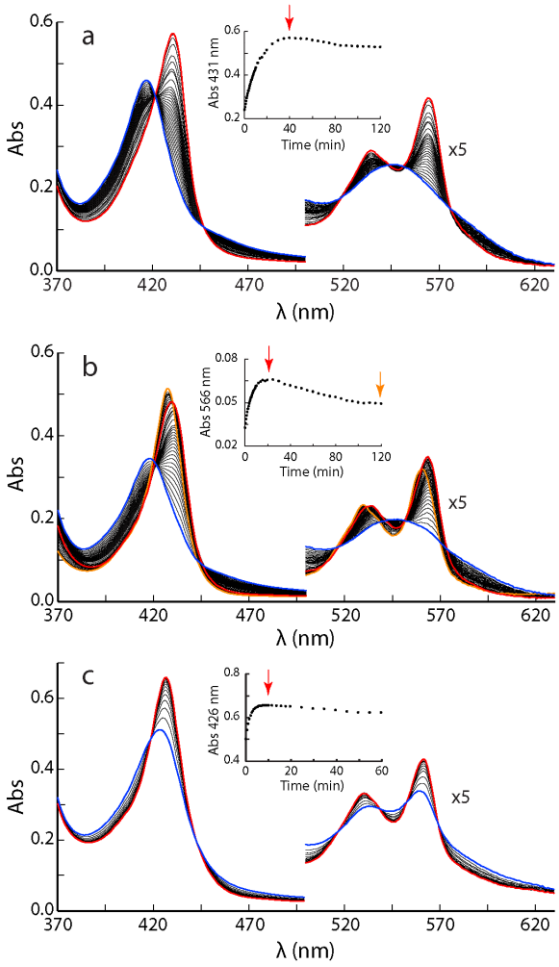


Figure 6.5. DT-mediated reduction of cyanomet CtrHbs monitored by UV-vis spectrophotometry. In each reaction, spectra were taken every 30 s for the first 10 min, every 2 min for the following 20 min, and every 5 min for the remaining 1.5 h. (a) Reduction of cyanomet WT CtrHb. The reduction progresses over 40 min from the blue spectrum to the red spectrum. Inset: kinetic trace at 431 nm; the decrease in signal after 40 min (red vertical arrow) is attributed to damage by DT. (b) Reduction of cyanomet T111H CtrHb. The initial response is similar to WT (blue to red spectrum) but occurs within 25 min. An additional blue shift of the spectrum is distinguished after this first phase (red to orange spectrum). Inset: kinetic trace at 566 nm to show the biphasic nature of the spectral evolution. The red and orange arrows mark the time of the red and orange spectra. (c) Reduction of cyanomet L75H CtrHb. The Soret band shifts to ~426 nm within the ~10-s manual mixing dead time. The reduction progresses over 10 min from the blue spectrum to the red spectrum, which is distinct from ferrous L75H CtrHb and corresponds to crosslinked L75H. Inset: kinetic trace at 426 nm, the decrease in signal after 10 min (red vertical arrow) is attributed to damage by DT. Note that only the first 60 min are shown.

The hemochromogen assay was performed to judge heme integrity. Both T111H and L75H CtrHbs treated with DT in the cyanomet state exhibit a blue shift of the hemochrome α and β maxima compared to WT CtrHb (data not shown) in support of a modified heme. Furthermore, Figure 6.4b illustrates that ECL staining of treated cyanomet L75H CtrHb detects heme only where the protein has migrated, similar to WT GlbN-A and in support of covalent attachment. ECL staining of treated cyanomet T111H CtrHb showed the presence of heme associated with the protein as well as free heme migrating with the dye front.

The changes observed in the optical spectra upon reduction of cyanomet T111H and L75H CtrHbs, along with the ECL results, indicated the presence of a covalent bond between the heme and protein in treated samples, but did not reveal the nature of the bond. Thus, these data warranted a detailed examination of the products by NMR spectroscopy. For both variants, the ferric cyanomet form was first generated by incubation with excess potassium cyanide and upon saturation, reacted with excess DT and monitored kinetically by ^{1}H NMR spectroscopy for at least 2 h (Supporting Information Figure S6.14–16). At the end of this period, the samples were exposed to potassium ferricyanide or air and completely reoxidized to generate the $S = 1/2$ cyanomet state. WT cyanomet CtrHb, when subjected to this treatment shows signs of modification. Reaction extent is small but reproducible, clearly detectable by NMR spectroscopy (Supporting Information Figure S6.14C), but does not yield a covalently attached heme (Figure 6.4b, lane 1). This illustrates the potential for

protein damage, likely due to the production of O_2^- or H_2O_2, when using DT to generate the crosslink. Comparison of the cyanomet ^{1}H-1D NMR spectra of T111H and L75H CtrHbs before and after treatment shows significant differences (Figure 6.3b–e).

NMR Identification of the heme modification in T111H CtrHb

The T111H CtrHb spectrum resulting from the treatment (Figure 6.3d) contains a multitude of hyperfine shifted lines with intensities suggesting a mixture of at least three species. The new predominant set of resonances corresponds to a major product accounting for ~60% of the sample. The inability for DT to reduce cyanomet T111H CtrHb completely under these NMR conditions can be seen from Supporting Information Figure S6.15 and explains the second set of peaks, which corresponds to the cyanomet major heme orientational isomer of the starting material (~25%). Finally, a third set corresponds to a minor product that may be due to a different modification of the heme or DT-damaged protein. Notably, after treatment, peaks corresponding to the minor cyanomet isomer of the T111H starting material have vanished. We therefore hypothesize that the primary product contains a covalent linkage between H111 and the heme 4-substituent in an orientation analogous to the minor isomer starting material. We refer to this product as T111H CtrHb-A^4 (Supporting Information Figure S6.20A).

Signals corresponding to heme methyl groups of T111H CtrHb-A^4 are assigned on the basis of their high intensity and characteristic paramagnetic downfield ^{1}H and upfield ^{13}C shifts. DQF-COSY spectra reveal one vinyl and both propionate spin systems. Intra-heme NOEs between the 1-CH_3 and 8-CH_3,

along with contacts between the heme 1- and 2-substituents indicate that the heme 2-vinyl group has remained intact (Supporting Information Figure S6.17). Cyanomet T111H CtrHb ^{1}H-^{15}N lr-HMQC data collected before and after DT treatment and compared with WT data indicate the presence of a peculiar imidazole system [52] in the T111H CtrHb-A^{4} product, exhibiting two downfield shifted ^{15}N signals *J*-coupled to a proton at 8.36 ppm, itself coupled to a proton at 6.48 ppm. These resonances are attributed to H111 (Table 6.2). Restraints defining the position of the side chain include: H111 Hε1 to heme 3-CH_3 and F48 Hζ and Hε protons, and H111 Hδ2 to the heme 3-CH_3 and V115 side chain methyls. Both H111 Hε1 and Hδ2 protons have shared NOEs (strong and weak, respectively) to a methyl group at 3.39 ppm, which is assigned to the modified heme 4-CβH_3 group. Accordingly, a scalar connectivity is observed between the 4-CβH_3 and a single proton at 6.15 ppm (Supporting Information Figure S6.17) corresponding to the modified heme 4-CαH. This proton, like the 4-CβH_3, is in dipolar contact with H111 Hε1. Although J_{HN}-coupling between H111 $^{15}N\varepsilon 2$ and the newly produced 4-CβH_3 or 4-CαH was not detected in the long-range data, the observed NOEs, heme methyl ^{1}H chemical shift dispersion pattern [53], propionate shifts, and unusual H111 ^{15}N shifts are all consistent with modification occurring at H111 and the heme 4-substituent (Supporting Information Figure S6.20A).

NMR Identification of the heme modification in L75H CtrHb

Unlike WT and T111H cyanomet CtrHbs, which are reduced slowly and with different efficiencies between the major and minor forms (Supporting Information Figure S6.14–15), DT reduction of L75H cyanomet CtrHb leads to

complete and rapid disappearance of all cyanomet resonances (Supporting Information Figure S6.16). Reoxidation results in a spectrum (Figure 6.3e) distinct from the starting material (Figure 6.3c) and consistent with high yield of one major product (> 90 %, henceforth L75H CtrHb-B). Only one heme vinyl group can be identified in DQF-COSY data of the product (Supporting Information Figure S6.18). The intact vinyl Hα exhibits an NOE to a heme methyl, itself having a weak intraheme connectivity to a second heme methyl. The vinyl and methyl chemical shifts (Table 6.2) and NOE pattern are analogous to the WT 2-vinyl ↔ 1-CH_3 ↔ 8-CH_3 sequence and suggest that the reaction occurs on pyrrole C. Also similar to the WT, characteristic 1-CH_3 ↔ F48 and 2-vinyl Hα ↔ F80 contacts are observed. Of the 2-vinyl β_{cis} and β_{trans} protons, the latter is oriented toward a third heme methyl, assigned as the heme 3-CH_3 (Supporting Information Figure S6.18C). This methyl shows distinctive dipolar contacts, including strong NOEs to the side chain methyls of V83 and the ring of F80. In addition, the heme 3-CH_3 displays an NOE to a methyl group, the latter (–0.49 ppm, Supporting Information Figure S6.18D) *J*-coupled to a proton resonating at 0.54 ppm. The relative intensity of the 0.54 ppm and –0.49 resonances is consistent with a 4-CαH-CβH_3 heme moiety. The 3:1:1:3 splitting pattern observed in the ^{13}C dimension of the ^{1}H-coupled natural abundance ^{13}C spectrum confirms that the signal at –0.49 ppm corresponds to a methyl group (Supporting Information Figure S6.19). The 4-CαH is oriented towards the heme 5-CH_3 and the 4-CβH_3 is proximal to the heme 3-CH_3, as in the starting material L75H major isomer.

Because WT CtrHb does not form a crosslink, H75 is the logical candidate for modification. ^{1}H-^{15}N lr-HMQC spectra collected on cyanomet WT and L75H CtrHbs before and after treatment demonstrate that, in L75H CtrHb-B, signals corresponding to intact H75 are missing and a new set of shifted cross peaks are detected. Both H75 imidazole nitrogens exhibit *J*-coupling to a proton at 7.12 ppm (Hε1). The ring ^{15}Nε2 is also *J*-coupled to a proton at 5.42 ppm (Hδ2) and to the heme 4-CβH$_3$ (–0.49 ppm, Fig. 6).

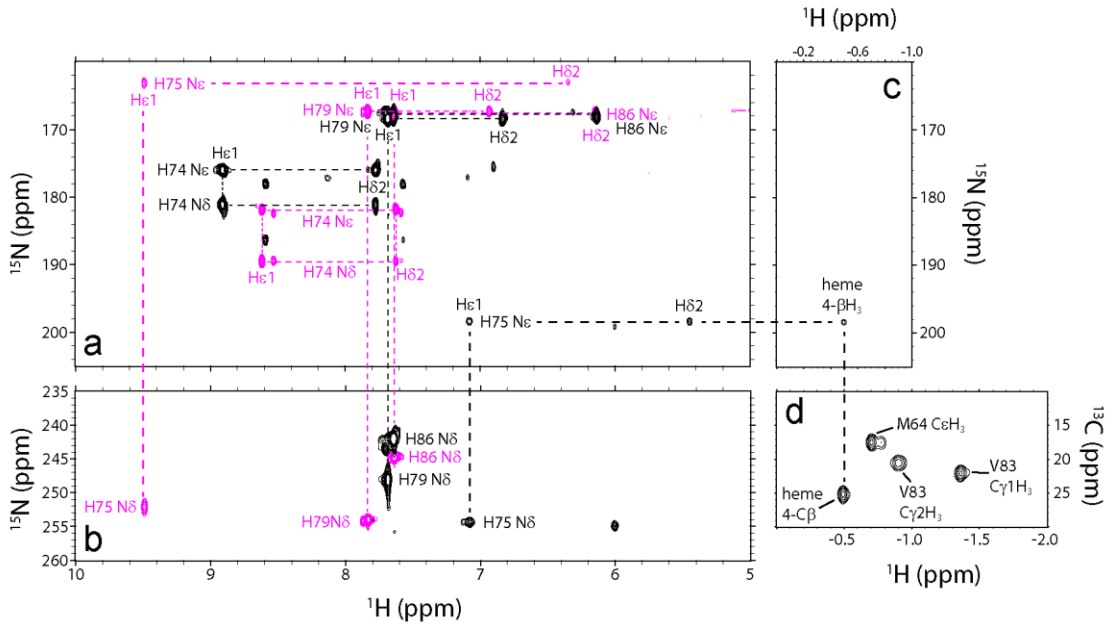


Figure 6.6. Histidine assignment in cyanomet L75H CtrHb and product CtrHb-B by NMR spectroscopy (pH 7.1–7.2, 298 K). (a) Overlay of ^{1}H-^{15}N lr-HMQC data (^{15}N upfield region). Signals from all histidine residues (except proximal H68) were detected. Reactant (magenta) and product (black) L75H histidine spin systems are connected with dashed lines. (b) Overlay of ^{1}H-^{15}N lr-HMQC data (^{15}N downfield region). Colors are as in (a). Note that the similarity of the H86, H74, and H79 imidazole signals (a, b) in CtrHb and CtrHb-B. Conversely, H75 Nε2, along with Hε1 and Hδ2, undergoes significant changes. (c) ^{1}H-^{15}N lr-HMQC data (^{1}H far-upfield region). In L75H cyanomet CtrHb-B, H75 Nε2 has a cross peak at –0.49 ppm, indicative of histidine N-alkylation. (d) Portion of a natural abundance ^{1}H-^{13}C HMQC spectrum collected on L75H cyanomet CtrHb-B. The coupled ^{1}H-^{13}C HSQC spectrum (Supporting Information Figure S6.19)

confirms that the resolved peak at a proton shift of –0.49 ppm is a methyl group.

NOE connectivities define the position of the modified histidine: H75 Hδ2 has strong dipolar interactions with both the heme 3- and 4-CβH_3 groups, whereas H75 Hε1 is oriented towards the 5-CH_3 group (data not shown). Together, these data support that the H75 ring undergoes a rotation from its original position and forms the anticipated crosslink, with R stereochemistry at the heme 4-Cα, as depicted in Supporting Information Figure S6.20B. Interestingly, for both T111H and L75H CtrHbs, the engineered histidine reorients in order to react with the heme.

Mechanistic considerations

The data accumulated with GlbN [20-22, 26] has provided a set of criteria for the addition of histidine to the ferrous heme. The first is proximity. The histidine must be within reach of a heme vinyl, and sp^3 geometry at the vinylic Cα atom must be sterically possible; however, the histidine need not be optimally oriented in the starting material as long as the barrier to adopt the productive state can be overcome thermally. The second is the ionization state of the histidine. In GlbN, data are consistent with the histidine providing the proton to the vinyl Cβ. Protonation is the rate determining step and the reaction should be carried out at a sufficiently low pH. A third criterion is the nature of the iron distal ligand. These are either permissive (histidine, cyanide, water/5-coordinate) [20] or inhibiting (O_2 and NO, data not shown; CO) [21, 26]. The inhibiting ligands have a slow dissociation rate constant, and their

electron-withdrawing property is presumably responsible for the inhibition. In GlbN, the permissive ligands are not tightly bound and may or may not be released prior to protonation and nucleophilic attack.

The behavior of the CtrHb variants brings new insights into the mechanism and limitations of the histidine addition. Ferrous WT CtrHb at neutral pH is reported to be a mixture of 4-, 5-, and 6-coordinate species, with equilibrium favoring the 5-coordinate high-spin complex [29]. In preliminary experiments, this protein was found to lose heme rapidly to apomyoglobin (half-life of ~8 min at pH 7.7, data not shown). Optical spectra suggest the same mixture of species in reduced T111H CtrHb and a slight increase in the population of a 6-coordinate species in L75H CtrHb (Supporting Information Figure S6.1B). The failure of ferrous T111H and L75H CtrHbs to react therefore appears related to absent or fleeting heme–protein contacts.

The modification occurs when reducing the well-folded cyanomet state. *A priori*, steric (i.e., correct seating of the heme), electrostatic, or heme electronic factors can all contribute to the reactivity. Attempts to crosslink in the presence of azide as an alternative anionic ligand were not successful (Supporting Information Figure S6.21) and imidazole coordination led only to partial crosslinking (Supporting Information Figure S6.22). The possibility that cyanide promotes the covalent modification by stabilizing the protonated vinyl seems unlikely given that binding of this anionic ligand is not expected to perturb significantly the electron distribution of the ferrous heme [54]. Thus, the data point to steric effects as the major determinants of reactivity in the absence of inhibiting ligands.

The protein conformation adopted when cyanide is bound may influence the reaction via subtle effects on the vinyl groups. We showed above that in the cyanomet complex, the major heme orientational isomer of WT CtrHb has the 2-vinyl group out of the heme plane and in a *cis*-like orientation, whereas T111H CtrHb undergoes a rearrangement to a *trans*-like orientation (Supporting Information Figure S6.11). The stereochemistry of the products indicates that the reactive vinyl is in a *trans*-like conformation. This conformation has a higher degree of conjugation with the porphyrin ring than *cis*-like conformations [55] and may undergo protonation more readily. Whether a *trans* vinyl configuration is in general a necessary condition for reaction is not clear. Regardless, the lack of reaction between H111 and the 2-vinyl (major isomer) in T111H CtrHb shows that a *trans* vinyl near a histidine is an insufficient feature for reaction. Facile heme reorientation associated with low heme and exogenous ligand affinity in the ferrous state presents H111 with the 4-vinyl group as a competitive alternative to the 2-vinyl group. Thus, the yield of T111H CtrHb-A^4 exceeds the minor isomer content in the cyanomet starting material. Likewise, the yield of L75H CtrHb-B exceeds the initial major isomer content. As anticipated, heme reorientation on a relatively rapid time scale emerges as a significant advantage for reaction.

In natural heme proteins, the vinyl groups are often surrounded by hydrophobic side chains. In prior work, we surveyed over 300 structures for the presence of leucines with Cα atom within 7 Å of a vinyl Cα atom and Cβ atom within 6 Å of the same vinyl Cα atom [15]. These criteria identified L79 in GlbN and L75 in CtrHb as candidates for histidine replacement and heme modification. Many other leucines were found. Examples include L106 in

nitrophorin 2 (1PEE), L211 in chlorite dismutase (2VXH), L37 in ascorbate peroxidase (1APX), and L71 in cytochrome b_5 (3NER), chosen to highlight proteins that have architectures different from that of globins. Many isoleucines, filtered by the same distance criteria, also occupy positions where a histidine may react. For in vitro applications, screening of pH conditions and permissive exogenous ligands that seat the heme properly can be systematically explored. Thus, it is reasonable to expect that many natural heme proteins can be modified to undergo the reaction, or artificial proteins can be designed to present a histidine with appropriate geometry to the heme.

Conclusion

The present work has several implications for the chemistry of heme proteins. (1) The existence of the PTM in GlbN [40] and its engineered presence in GlbN [22] and CtrHb cautions that the PTM may occur in additional heme proteins, not only wild-type but also unwittingly in His-tagged versions, (histidine) variants, or designed proteins. (2) In preparing the histidine Nε2–heme vinyl Cα crosslink, care must be exercised in the choice of a reducing agent to avoid heme and protein damage, as occasionally caused by the oxidative by-products of aerobic DT reduction. Anaerobic DT treatment or reduction with alternative agents such as a ferredoxin system [56] if sufficiently powerful may be preferable. (3) It is possible to engineer the crosslink in a *b* heme protein using relatively mild reaction conditions. Although detailed kinetic analysis was not performed on CtrHb variants, the products are

consistent with the mechanism described previously [21], in which case the pK_a of the reactive histidine, in addition to its ability to adopt the geometry of the product, is an important factor. (4) The necessity of using the cyanomet adduct as the species to be reduced rather than the ligand free protein (as in GlbNs) is likely related to the conformation of the starting material. In view of the CtrHb results, we expect that successful application to other proteins will depend principally on the position of the engineered histidine, the local flexibility of the supporting structure, the absence of inhibiting distal ligands, and possibly the orientation of the heme vinyl group.

The ability to prepare non-natively crosslinked versions of heme proteins extends the range of heme chemistry questions that can be addressed by biophysical methods. For example, to what extent does protonation of the alkylated histidine influence the heme reduction potential? How do the histidine-heme linkages differ from the cysteine–heme thioether linkages of *c*-type cytochromes? What is the electronic influence of the distal ligand? Practical usage of the crosslinking reaction includes the formulation of artificial enzymes or oxygen transporters with non-dissociable heme; such systems are expected to have an elongated lifetime, have increased resistance to proteolytic cleavage, and may be capable of functioning over a wide range of conditions (e.g. high temperature, low pH) for industrial or biomedical applications.

Acknowledgement

This study was supported by National Science Foundation grant MCB-1330488. Data were collected at the Johns Hopkins Biomolecular NMR Center with assistance from Dr. Ananya Majumdar.

References

[1] M.S. Hargrove, D. Barrick, J.S. Olson, Biochemistry 35 (1996) 11293-11299.

[2] D.A. Landfried, D.A. Vuletich, M.P. Pond, J.T.J. Lecomte, Gene 398 (2007) 12-28.

[3] A.B. Cowley, M.L. Kennedy, S. Silchenko, G.S. Lukat-Rodgers, K.R. Rodgers, D.R. Benson, Inorg. Chem. 45 (2006) 9985-10001.

[4] T.G. Traylor, V.S. Sharma, Biochemistry 31 (1992) 2847-2849.

[5] M. Gullotti, L. Santagostini, E. Monzani, L. Casella, Inorg. Chem. 46 (2007) 8971-8975.

[6] B.G. Fritz, X. Hu, J.L. Brailey, R.E. Berry, F.A. Walker, W.R. Montfort, Biochemistry 50 (2011) 5813-5815.

[7] A.M. Thompson, A.R. Reddi, X. Shi, R.A. Goldbeck, P. Moënne-Loccoz, B.R. Gibney, T.R. Holman, Biochemistry 46 (2007) 14629-14637.

[8] Y.M. Kim, H.A. Bergonia, C. Muller, B.R. Pitt, W.D. Watkins, J.R. Lancaster, Jr., J. Biol. Chem. 270 (1995) 5710-5713.

[9] S. Kundu, M.S. Hargrove, Proteins 50 (2003) 239-248.

[10] A.I. Alayash, Trends Biotechnol. 32 (2014) 177-185.

[11] B.R. Gibney, F. Rabanal, K.S. Reddy, P.L. Dutton, Biochemistry 37 (1998) 4635-4643.

[12] J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, Chem. Rev. 114 (2014) 4366-4469.

[13] R.G. Kranz, C. Richard-Fogal, J.S. Taylor, E.R. Frawley, Microbiol. Mol. Biol. Rev. 73 (2009) 510-528.

[14] S.E. Bowman, K.L. Bren, Nat. Prod. Rep. 25 (2008) 1118-1130.

[15] D.A. Mavridou, S.J. Ferguson, J.M. Stevens, IUBMB Life 65 (2013) 209-216.

[16] M. Braun, I.G. Rubio, L. Thöny-Meyer, Appl. Microbiol. Biotechnol. 67 (2005) 234-239.

[17] M. Braun, L. Thöny-Meyer, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 12830-12835.

[18] P.D. Barker, J.C. Ferrer, M. Mylrajan, T.M. Loehr, R. Feng, Y. Konishi, W.D. Funk, R.T. MacGillivray, A.G. Mauk, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 6542-6546.

[19] O. Daltrop, J.W. Allen, A.C. Willis, S.J. Ferguson, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 7872-7876.

[20] H.J. Nothnagel, N. Love, J.T.J. Lecomte, J. Inorg. Biochem. 103 (2009) 107-116.

[21] H.J. Nothnagel, M.R. Preimesberger, M.P. Pond, B.Y. Winer, E.M. Adney, J.T.J. Lecomte, J. Biol. Inorg. Chem. 16 (2011) 539-552.

[22] M.R. Preimesberger, B.B. Wenke, L. Gilevicius, M.P. Pond, J.T.J. Lecomte, Biochemistry 52 (2013) 3478–3488.

[23] T.K. Das, M. Couture, Y. Ouellet, M. Guertin, D.L. Rousseau, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 479-484.

[24] J.A. Hoy, B.J. Smagghe, P. Halder, M.S. Hargrove, Protein Sci. 16 (2007) 250-260.

[25] D.A. Vuletich, C.J. Falzone, J.T.J. Lecomte, Biochemistry 45 (2006) 14075-14084.

[26] M.R. Preimesberger, M.P. Pond, A. Majumdar, J.T.J. Lecomte, J. Biol. Inorg. Chem. 17 (2012) 599-609.

[27] M.P. Pond, A. Majumdar, J.T.J. Lecomte, Biochemistry 51 (2012) 5733-5747.

[28] M. Dellarole, C. Roumestand, C. Royer, J.T.J. Lecomte, Biochim. Biophys. Acta 1834 (2013) 1910-1922.

[29] M. Couture, T.K. Das, H.C. Lee, J. Peisach, D.L. Rousseau, B.A. Wittenberg, J.B. Wittenberg, M. Guertin, J. Biol. Chem. 274 (1999) 6898-6910.

[30] T.K. Das, M. Couture, H.C. Lee, J. Peisach, D.L. Rousseau, B.A. Wittenberg, J.B. Wittenberg, M. Guertin, Biochemistry 38 (1999) 15360-15368.

[31] A. Pesce, M. Couture, S. Dewilde, M. Guertin, K. Yamauchi, P. Ascenzi, L. Moens, M. Bolognesi, EMBO J. 19 (2000) 2424-2434.

[32] M. Couture, M. Guertin, Eur. J. Biochem. 242 (1996) 779-787.

[33] J.T.J. Lecomte, N.L. Scott, B.C. Vu, C.J. Falzone, Biochemistry 40 (2001) 6541-6552.

[34] M. Milani, Y. Ouellet, H. Ouellet, M. Guertin, A. Boffi, G. Antonini, A. Bocedi, M. Mattu, M. Bolognesi, P. Ascenzi, Biochemistry 43 (2004) 5213-5221.

[35] C. de Duve, Acta Chem. Scan. 2 (1948) 264-289.

[36] E.A. Berry, B.L. Trumpower, Anal. Biochem. 161 (1987) 1-15.

[37] D.W. Dorward, Anal. Biochem. 209 (1993) 219-223.

[38] C. Vargas, A.G. McEwan, J.A. Downie, Anal. Biochem. 209 (1993) 323-326.

[39] B.C. Vu, D.A. Vuletich, S.A. Kuriakose, C.J. Falzone, J.T.J. Lecomte, J. Biol. Inorg. Chem. 9 (2004) 183-194.

[40] B.C. Vu, A.D. Jones, J.T.J. Lecomte, J. Am. Chem. Soc. 124 (2002) 8544-8545.

[41] C.J. Falzone, J.T.J. Lecomte, J. Biomol. NMR 23 (2002) 71-72.

[42] M.P. Pond, D.A. Vuletich, C.J. Falzone, A. Majumdar, J.T.J. Lecomte, Biomol. NMR Assign. 3 (2009) 211-214.

[43] J.G. Pelton, D.A. Torchia, N.D. Meadow, S. Roseman, Protein Sci. 2 (1993) 543-558.

[44] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, J. Biomol. NMR 6 (1995) 135-140.

[45] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, J. Biomol. NMR 6 (1995) 277-293.

[46] T.D. Goddard, D.G. Kneller, University of California, San Francisco (2006).

[47] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, J. Comput. Chem. 25 (2004) 1605-1612.

[48] G.N. La Mar, N.L. Davis, D.W. Parish, K.M. Smith, J. Mol. Biol. 168 (1983) 887-896.

[49] K.B. Lee, G.N. La Mar, L.A. Kehres, E.M. Fujinari, K.M. Smith, T.C. Pochapsky, S.G. Sligar, Biochemistry 29 (1990) 9623-9631.

[50] G.N. La Mar, J.D. Satterlee, J.S. de Ropp, in: K.M. Smith, K. Kadish, R. Guilard (Eds.), The Porphyrin Handbook, vol. 5, Academic Press, Burlington, MA, 2000, pp. 185-298.

[51] A. Bolli, C. Ciaccio, M. Coletta, M. Nardini, M. Bolognesi, A. Pesce, M. Guertin, P. Visca, P. Ascenzi, FEBS J. 275 (2008) 633-645.

[52] J.A. Vila, J. Phys. Chem. B 116, American Chemical Society (2012) 6665-6669.

[53] I. Bertini, C. Luchinat, G. Parigi, F.A. Walker, J. Biol. Inorg. Chem. 4 (1999) 515-519.

[54] A. Boffi, E. Chiancone, S. Takahashi, D.L. Rousseau, Biochemistry 36 (1997) 4505-4509.

[55] M.P. Marzocchi, G. Smulevich, J. Raman Spect. 34 (2003) 725-736.

[56] A. Hayashi, T. Suzuki, M. Shin, Biochim. Biophys. Acta 310 (1973) 309-316.

Table 6.1. Heme *b* and engineered histidine chemical shifts of WT, T111H, and L75H cyanomet CtrHbs[a].

	WT	WT*	T111H	T111H*[b]	L75H	L75H*
1-CH_3	16.77	3.2	14.38	3.13?	17.64	
2-v α	22.87	12.46	17.75	13.20	23.14	12.85
2-v β_c, β_t	–4.55, –4.65	–1.69, –2.49	–4.84, –4.34	–2.34, –3.32	–5.10, –5.11	–1.11, –1.42
3-CH_3	14.35	23.71	14.97	22.08	14.17	24.42
4-v α	6.68	15.17	7.04	14.23	6.59	15.46
4-v β_c, β_t	0.29, 0.58	–1.53, –0.12	0.38, 0.72	–1.13, –0.05	0.33, 1.12	–1.81, –0.47
5-CH_3	15.87	5.65	15.02	6.02?	16.83	6.17?
6-p α, α'	13.88, 11.76	9.35, 2.75	12.96, 11.16	9.28, 2.77	13.35, 13.08	8.69, 3.60
6-p β, β'	0.32, –0.98	–0.84, –1.53	0.30, –0.99	–1.63, –0.89	0.09, –1.06	–1.27, –2.08
7-p α, α'	7.84, 3.68	14.21, 11.38	7.97, 4.57	14.46, 11.44	7.06, 3.66	12.79, 11.90
7-p β, β'	–1.34, –2.31	1.00, –0.22	–1.99, –1.08	0.59, –0.68	–1.56, –2.53	0.62, –0.42
8-CH_3	6.17	17.42	7.49	17.36	5.91	17.64
H111 Hε1, Hδ2	n/a	n/a	7.86, 7.28	7.91, 7.32	n/a	n/a
H75 Hε1, Hδ2	n/a	n/a	n/a	n/a	9.43, 6.29	
H75 Nδ1, Nε2	n/a	n/a	n/a	n/a	252.1, 163.0	

[a] pH 7.2, 298 K. [b] A * refers to the minor heme orientational isomer. A ? indicates a tentative assignment.

Table 6.2. Reacted heme and alkylated histidine chemical shifts of T111H-A[4] and L75H-B cyanomet CtrHbs[a]

	T111H-A[4]	L75H-B
1-CH_3	2.02	15.57 (–26.3)
2-v α	13.47	26.44
2-v β_c, β_t	–2.83, –3.84	–5.46, –5.78
3-CH_3	20.26	11.55 (–31.1)
4-α	6.13	0.54
4-β-CH_3	3.45	–0.49 (25.2)
5-CH_3	5.88	18.81 (–34.8)
6-p α, α’	8.50, 1.87	12.91, 12.56
6-p β, β’	–1.03, –1.84	–0.27, –1.57
7-p α, α’	15.46, 12.19	7.74, 3.89
7-p β, β’	0.53, –0.81	–1.75, –2.66
8-CH_3	18.65	5.66 (–12.7)
H111 Hε1, Hδ2	8.36, 6.48	n/a
H111 Nδ1, Nε2	250.6, 220.7	n/a
H75 Hε1, Hδ2	n/a	7.12, 5.42
H75 Nδ1, Nε2	n/a	254.6, 198.6

[a] pH 7.2, 298 K. Values in parentheses are ^{13}C shifts.

Supporting Information for

Introduction of a covalent histidine–heme linkage in a hemoglobin: A promising tool for heme protein engineering.

Table S6.1: Optical properties of WT, T111H, and L75H CtrHbs[a]

State/Protein[b]	WT CtrHb	T111H CtrHb	L75H CtrHb
ferric (this work)	409 (117)[c] 535 (8.7)	411 (117) 531 (9.6)	409 (117) 530 (12)
ferric[d]	410 (110) 535 (8.5)		
cyanomet (this work)	416 (97.6)[e] 546 (11)	417 (114) 545 (13)	416 (119) 544 (13)
cyanomet[e]	416 (97.6) 547 (11)		
ferrous (this work)	427 (98.3) 529 (sh), 557 (13)	429 (96.8) 530 (sh), 559 (13)	425 (111) 529 (10), 558 (17)
ferrous[d]	426 (108) 529 (sh), 557 (15)		
ferrous cyanide (this work)[f]	431 (142) 534 (14), 564 (21)	430 (152) 534 (15), 564 (22)	n/a
ferrous cyanide[e]	434 (149)		
ferrous cyanide (X-linked, this work)	n/a	428 (163) 530 (15), 561 (20)	426 (184) 530 (18), 561 (23)

[a] Data at pH 7.1, room temperature. Wavelength maxima are given in nm (extinction coefficients are in mM^{-1} cm^{-1}).

[b] For all three proteins, mass spectrometry data (Acquity / Xevo-G2 UPLC-MS, Waters) were consistent with the coded sequences from which the initial methionine had been cleaved. The experimental molecular masses were: WT CtrHb, 13297.7 Da (expected 13298.2 Da); L75H CtrHb, 13321.1 Da (expected 13322.2 Da); and T111H CtrHb (^{15}N labeled), 13491.7 Da (expected 13493.2 Da for 98% ^{15}N abundance). The WT CtrHb mass was unaffected by aerobic reduction with 2mM dithionite.

[c] Extinction coefficient calculated from the cyanomet value of 97.6 mM^{-1} cm^{-1} provided in [1]. The 117 mM^{-1} cm^{-1} value was applied to all three ferric CtrHbs. All other values calculated from the ferric state reference spectra.

[d] From [2], pH 7.5.

[e] From [1], pH 7.0, 20° C.

[f] The same spectrum with a Soret at 431 nm is obtained upon DT reduction of a cyanomet sample or addition of cyanide to a DT-reduced sample. This is in contrast to the observations published in [1].

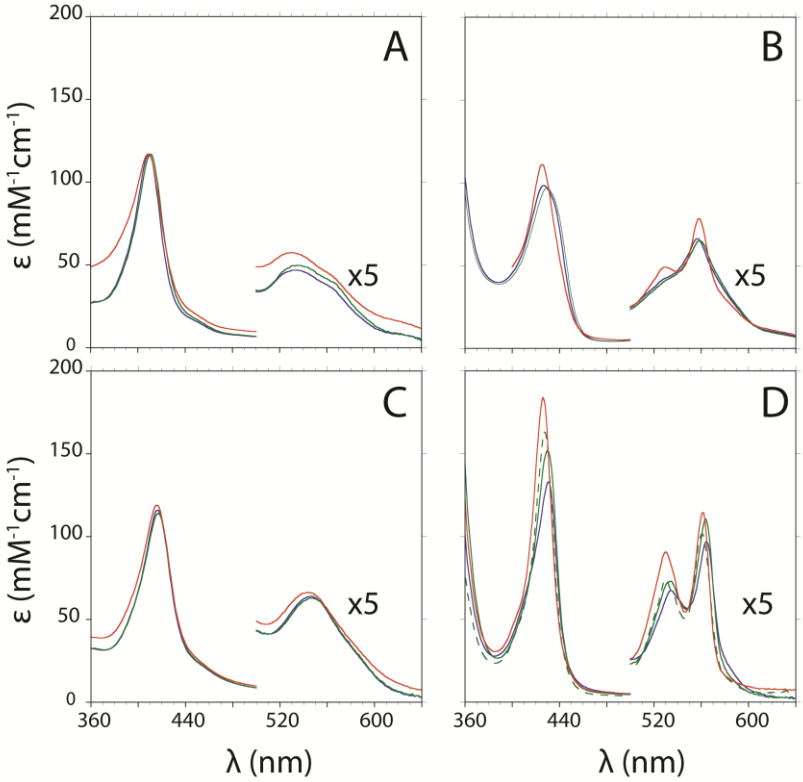


Figure S6.1. Reference optical spectra of WT (blue), T111H (green), and L75H (red) CtrHbs. All protein samples (4–10 μM) were prepared in 100 mM phosphate, pH ~7.1, and spectra were collected at room temperature. (A) Ferric state. The L75H variant exhibits some light scattering caused by aggregation. (B) Ferrous state. (C) Cyanomet state, prepared by addition of a 5-fold excess of KCN. (D) Ferrous cyanide-bound state prepared by DT reduction (~2 mM) of the cyanomet starting material. For T111H CtrHb, the solid line corresponds to the spectrum obtained initially after reduction; the dashed green line corresponds to the final product. Note the blue shift of the T111H and L75H variants relative to WT CtrHb.

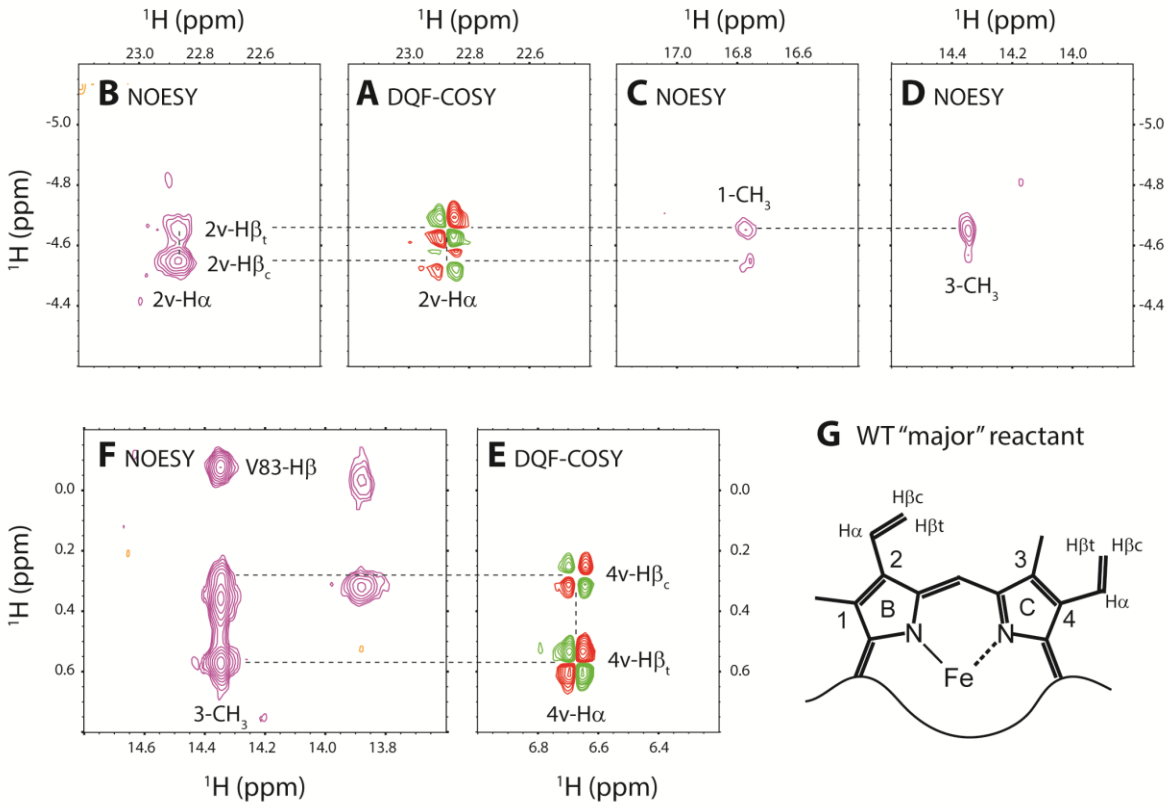


Figure S6.2. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in WT cyanomet CtrHb (major isomer) and vinyl orientation diagram. (A) *J*-coupled protons corresponding to the 2-vinyl group. (B) Intra-vinyl NOEs of the 2-vinyl group support 2-β_{cis} and 2-β_{trans} assignments. (C) The 2-vinyl β protons exhibit weak NOEs to the heme 1-CH_3. (D) The 2-vinyl β_{trans} proton is within dipolar contact with the heme 3-CH_3. Together, these connectivities support a *cis*-like orientation for the 2-vinyl group. (E) *J*-coupled protons corresponding to the 4-vinyl group. (F) The 4-vinyl β protons exhibit strong NOEs to the heme 3-CH_3 indicating it adopts a *trans*-like orientation. (G) Cartoon model depicting the orientations of vinyl groups within cyanomet WT CtrHb.

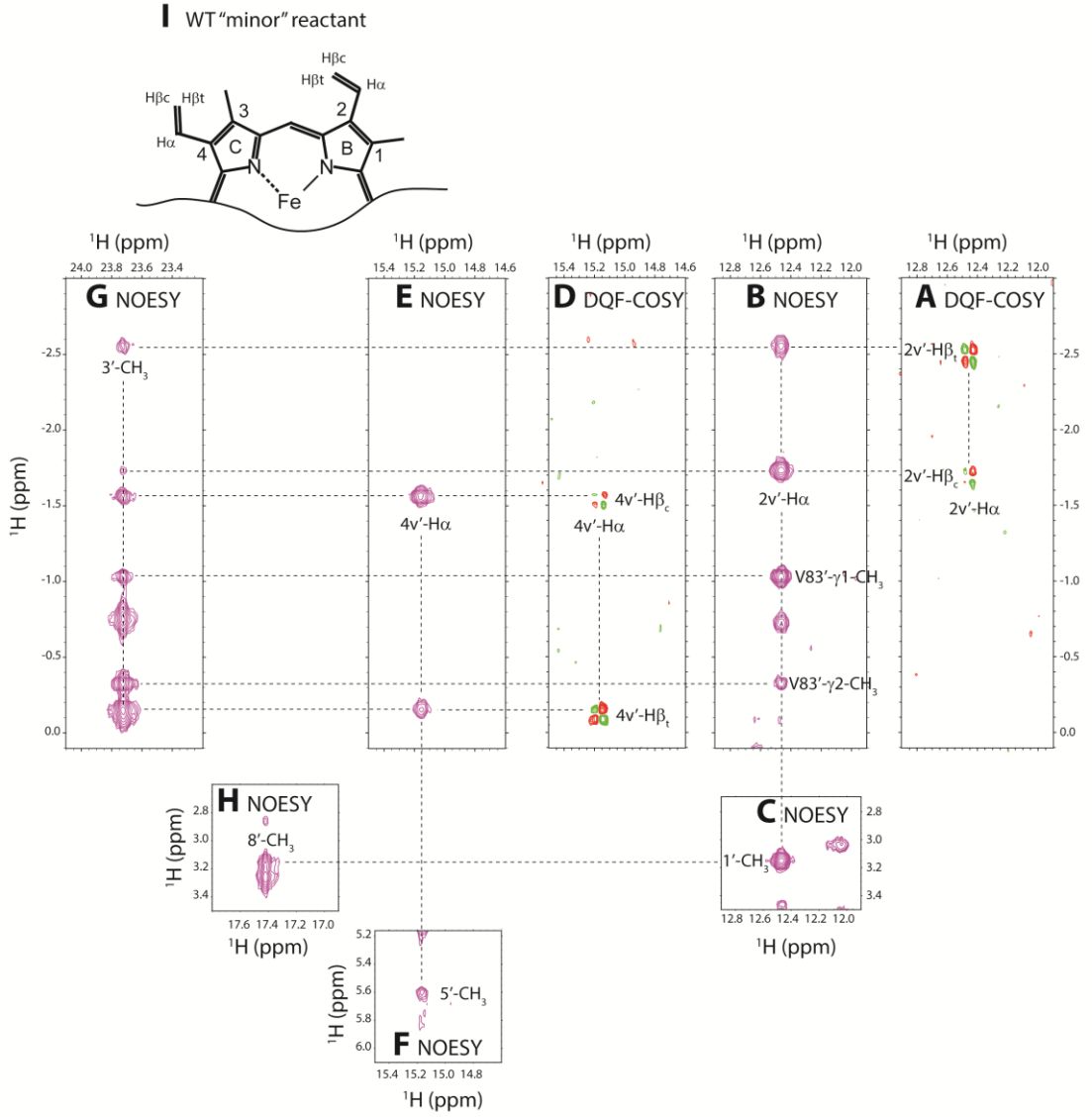


Figure S6.3. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in cyanomet WT CtrHb (minor isomer) and vinyl orientation diagram. (A) *J*-coupled protons assigned to the 2-vinyl group. (B) The 2-vinyl Hα has strong NOEs to the neighboring Hβs. (C) The 2-vinyl Hα also shows a strong contact to the heme 1-CH_3. (D) *J*-correlated signals assigned to the 4-vinyl group. (E) The 4-vinyl Hβs exhibit strong intra-vinyl NOE to the Hα. (F) The 4-vinyl Hα is in dipolar contact with the heme 5-CH_3. (G) The 2- and 4-vinyl Hβs are proximal to the heme 3-CH_3. (H) The heme 8-CH_3 shows an NOE to the 1-CH_3. The heme connectivity pattern is thus: 8-CH_3 ↔ 1-CH_3 ↔ 2-vinyl ↔ 3-CH_3 ↔ 4-vinyl ↔ 5-CH_3. (I) Cartoon model depicting vinyl orientations in cyanomet wild-type CtrHb.

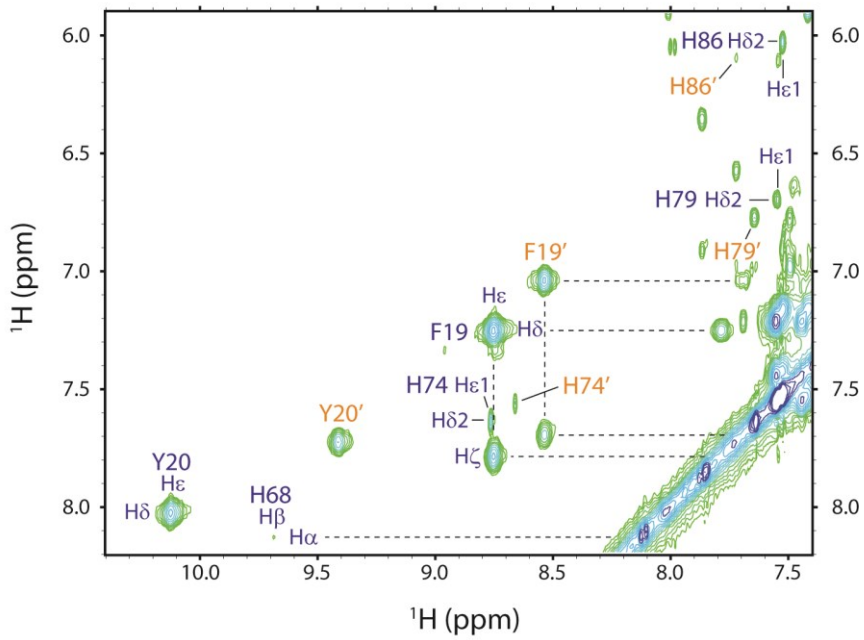


Figure S6.4. Downfield aromatic region of a ^{1}H-^{1}H TOCSY spectrum collected on cyanomet WT CtrHb. Assignments for some key heme pocket residues [H68 (His F8), Y20 (Tyr B10), and F19 (Phe B9)] and additional histidines are included. Peaks corresponding to the major heme orientational isomer are labeled in blue; peaks assigned to the minor heme orientational isomer are labeled in orange and denoted with prime.

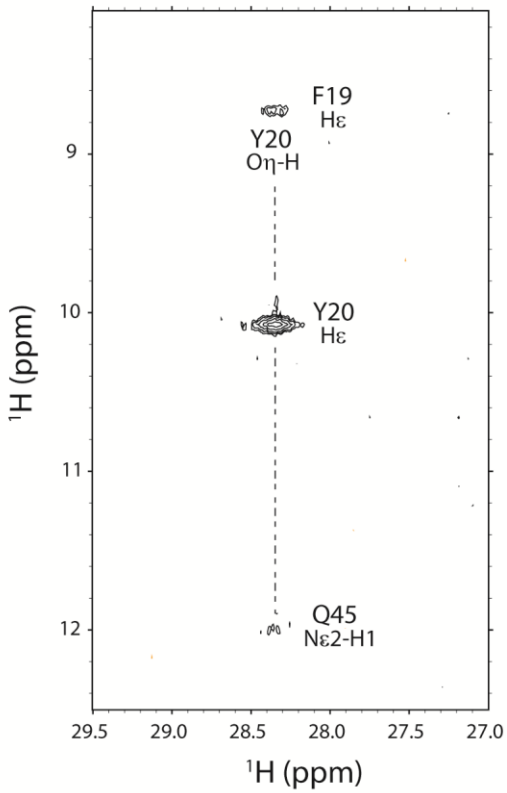


Figure S6.5. Key NOEs exhibited by the distal H-bonding Y20 OηH (Tyr B10) in cyanomet CtrHb. A downfield exchangeable proton (~28.4 ppm) is detected in WT NOESY data collected in H_2O. A strong NOE connects this signal to the Y20 CεHs. Weaker contacts to the distal residues F19 and Q45 are also observed. The NOEs, lability, and far downfield shift of the 28.4 ppm proton support its assignment to the distal Y20 OηH (Tyr B10), which stabilizes exogenous cyanide. Corresponding signals were observed for both cyanomet T111H and L75H CtrHb variants (see Figure 6.2, peaks labeled a), demonstrating that the variations did not appreciably perturb the distal cyanide H-bond network.

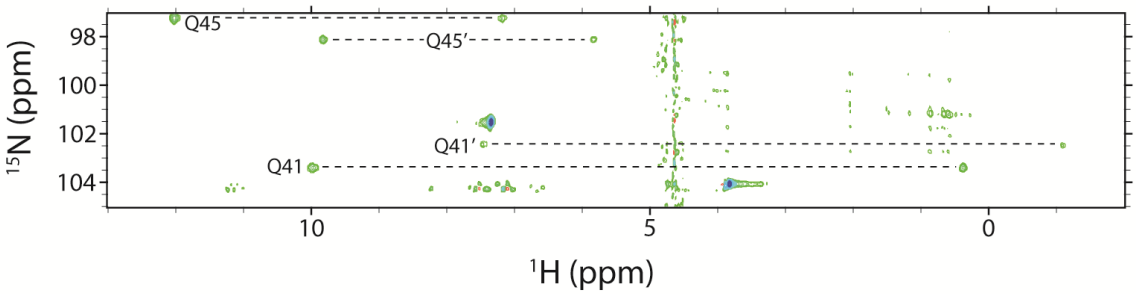


Figure S6.6. ^{1}H-^{15}N HSQC spectral region showing the highly shifted NH_2 systems of distal H-bonding Q41 and Q45 (Gln E7 and E11) residues. The highly shifted proton resonances of four NH_2 groups were tentatively assigned to the distal H-bonding residues Q45 and Q41 in the major and minor heme orientational isomers. The same signals are detected in both cyanomet T111H and L75H CtrHbs.

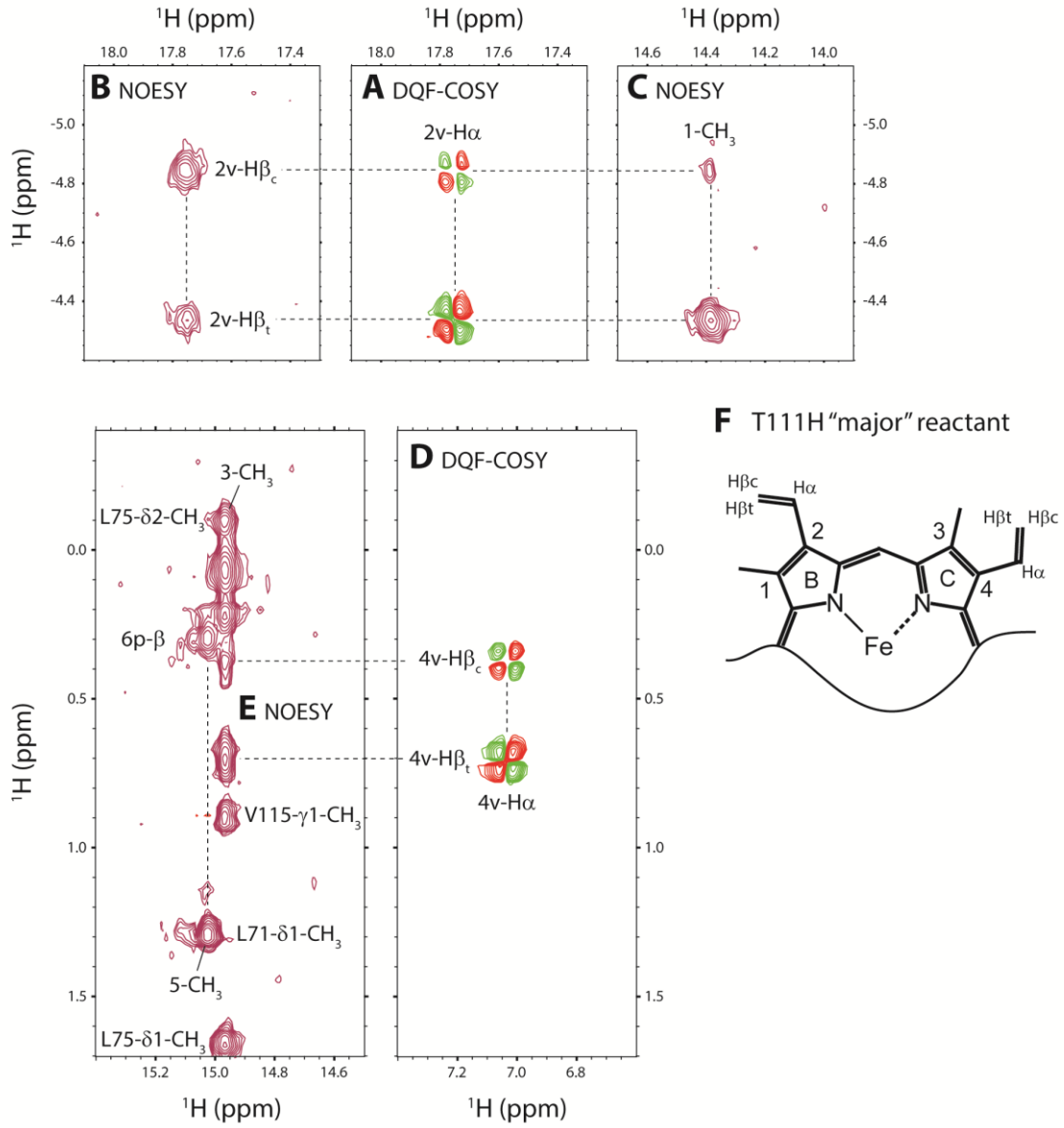


Figure S6.7. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in cyanomet T111H CtrHb (major isomer) and vinyl orientation diagram. (A) *J*-coupled AMX spin system assigned to the heme 2-vinyl group. (B) Intra-vinyl NOEs confirm the 2-β_{cis} and 2-β_{trans} proton assignments. (C) Strong NOE contacts between the 2-Hβs and the heme 1-CH_3 group orient the 2-vinyl in the *trans* conformation. These data support distinct 2-vinyl orientations between WT and T111H CtrHbs. (D) *J*-coupled signals corresponding to the heme 4-vinyl substituent. (E) NOEs between the 4-vinyl Hβs and the heme 3-CH_3 indicate that the former adopts a *trans* orientation, as observed for WT CtrHb. (F) Heme cartoon depicting the vinyl orientations consistent with NOE data.

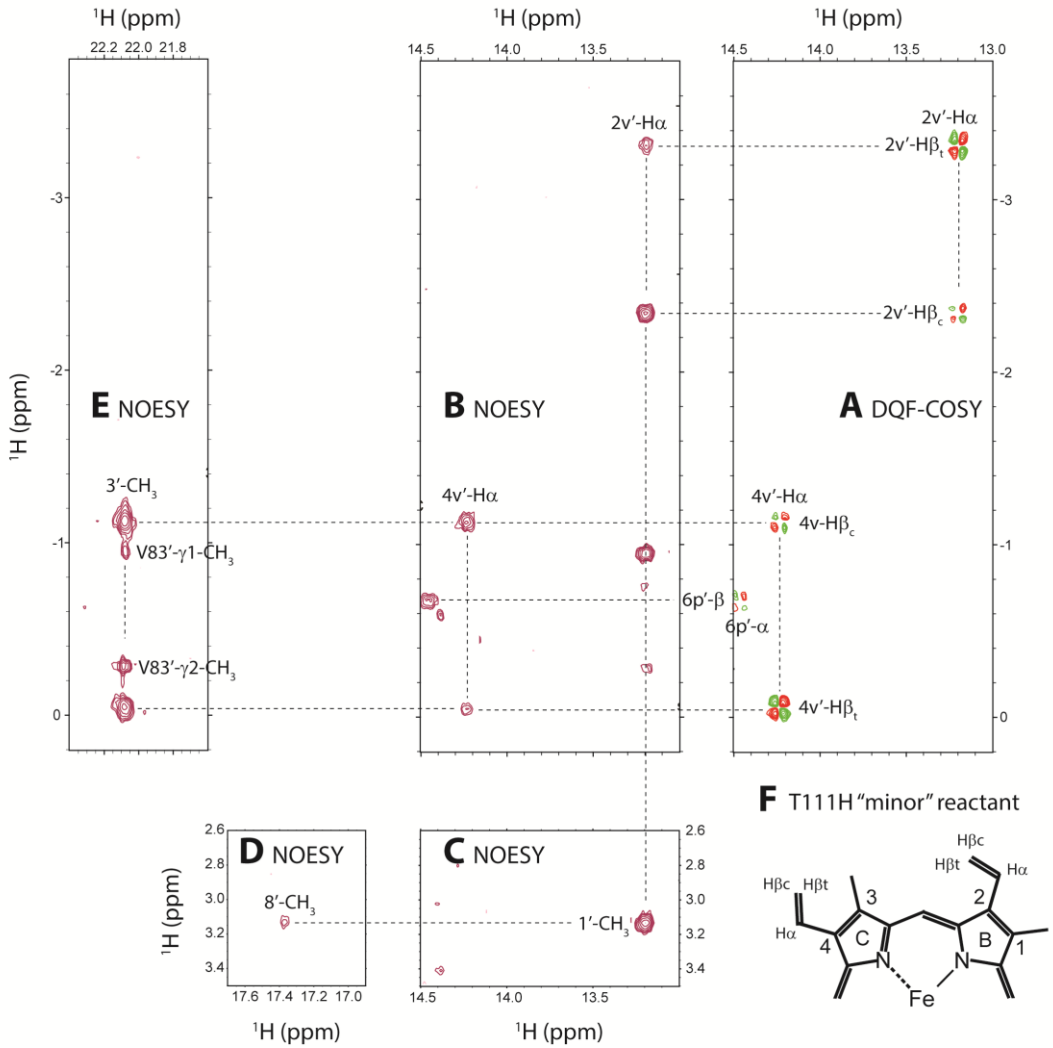


Figure S6.8. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in T111H cyanomet CtrHb (minor isomer) and vinyl orientation diagram. (A) *J*-correlated signals corresponding to the 2- and 4-vinyl groups. (B) Corresponding region displaying the intra-vinyl NOEs, confirming the assignment of β_{cis} and β_{trans} protons. (C) The 2-vinyl Hα displays a strong dipolar contact to the heme 1-CH_3 group supporting a 2-vinyl *cis* orientation. (D) A weak NOE was detected between the heme 1- and 8-CH_3. (E) The 4-vinyl β protons both show strong contacts to the heme 3-CH_3, indicative of the 4-vinyl group adopting a *trans* orientation. (F) Cartoon depiction of the heme vinyl group orientations.

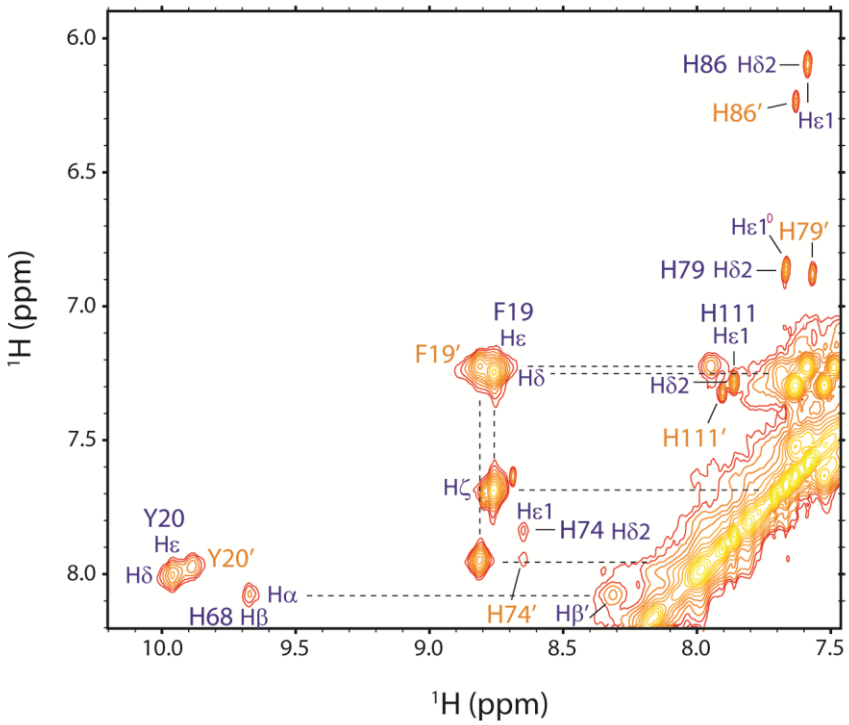


Figure S6.9. Downfield aromatic region of a ^{1}H-^{1}H TOCSY spectrum collected on cyanomet T111H CtrHb. Detectable histidine signals, along with distal heme pocket residues [Y20 (Tyr B10), F19 (Phe B9)] are labeled (major orientational isomer = blue, minor isomer = orange). Note the presence of the His111 Hε1 ↔ Hδ2 correlations. Interestingly, there is a closer chemical shift similarity between major and minor heme isomers in T111H CrHb compared to the WT CtrHb.

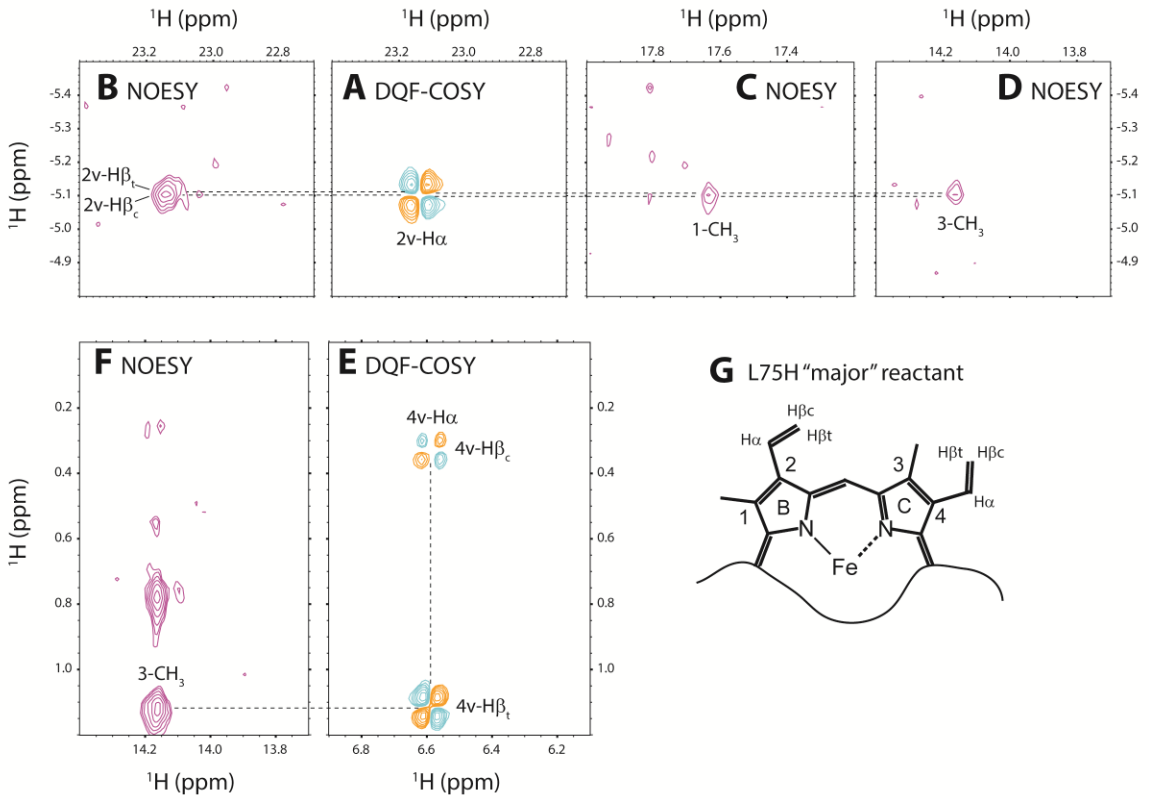


Figure S6.10. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in cyanomet L75H CtrHb (major isomer) and vinyl orientation diagram. (A) *J*-correlations exhibited between the 2-vinyl Hβs (overlapped) and the 2-vinyl Hα. (B) Corresponding region showing NOEs between the 2-vinyl Hβs and Hα. (C) The 2-vinyl Hβs are in weak dipolar contact with the heme 1-CH_3. (D) The 2-vinyl Hβs also show weak NOEs to the heme 3-CH_3 group, supporting a *cis* configuration. (E) DQF-COSY cross-peaks assigned to the heme 4-vinyl β_{cis}, β_{trans} and α protons. (F) The 4-vinyl Hβ_{trans} exhibits a strong NOE contact to the heme 3-CH_3 group, indicative of a 4-vinyl *trans* geometry. (G) Cartoon model showing the heme 2- and 4-vinyl orientations consistent with NMR data.

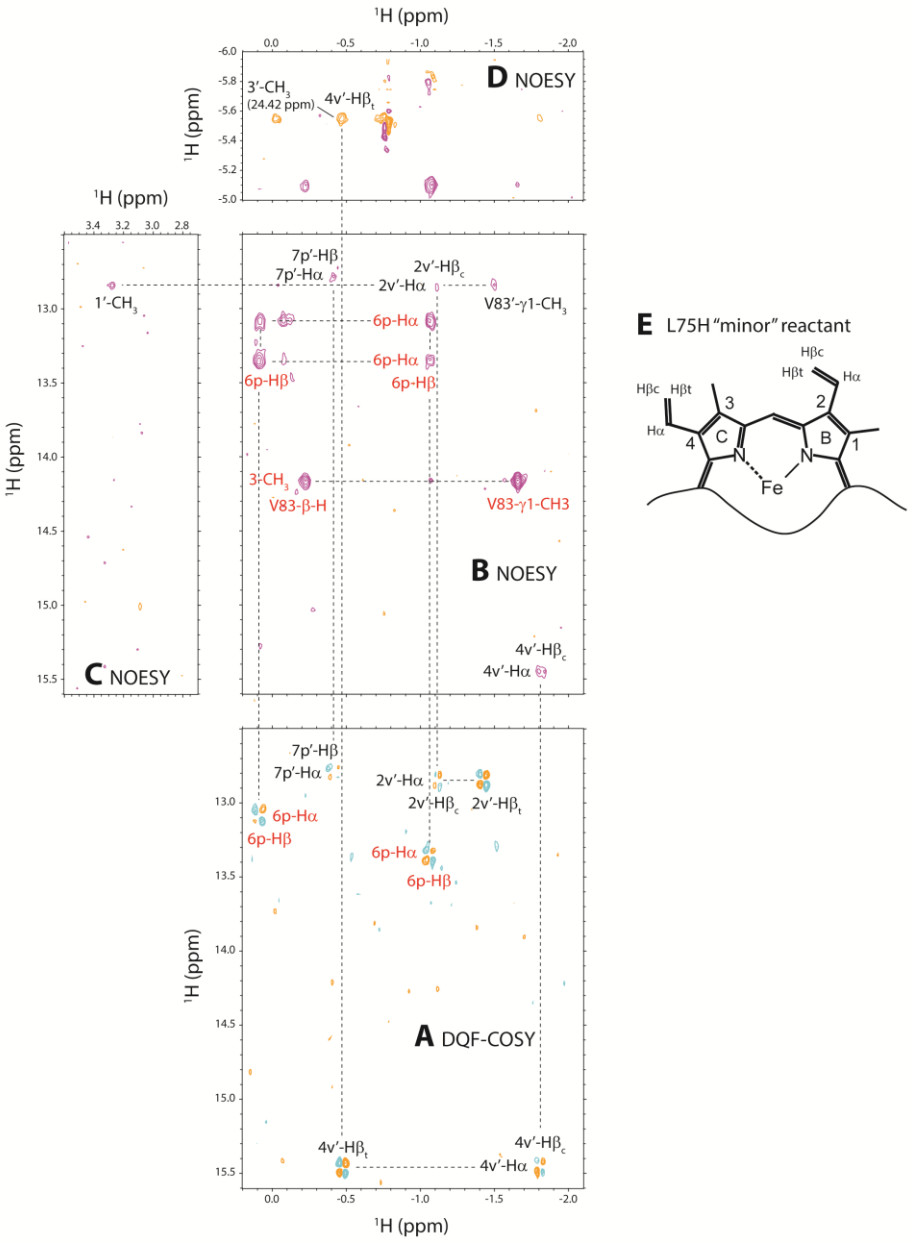


Figure S6.11. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme vinyl assignments in L75H cyanomet CtrHb (minor isomer) and vinyl orientation diagram. (A) *J*-correlations between the α and β protons of the heme 2- and 4-vinyl groups (black labels). Signals annotated with red labels derive from the L75H cyanomet CtrHb major heme isomer. (B) Corresponding region of the NOESY spectrum, highlighting intra-heme vinyl NOEs (labels as in A). (C) The 2-vinyl Hα exhibits a dipolar contact to the heme 1-methyl. Because the Hβs do not show such effects, the data support a 2-vinyl *cis* orientation. (D) The 4-vinyl β_{trans} proton shows a strong NOE to the heme 3-CH_3 (folded). This strong contact defines a 4-vinyl *trans* orientation. (E) Heme structural model depicting vinyl orientations consistent with NOE data.

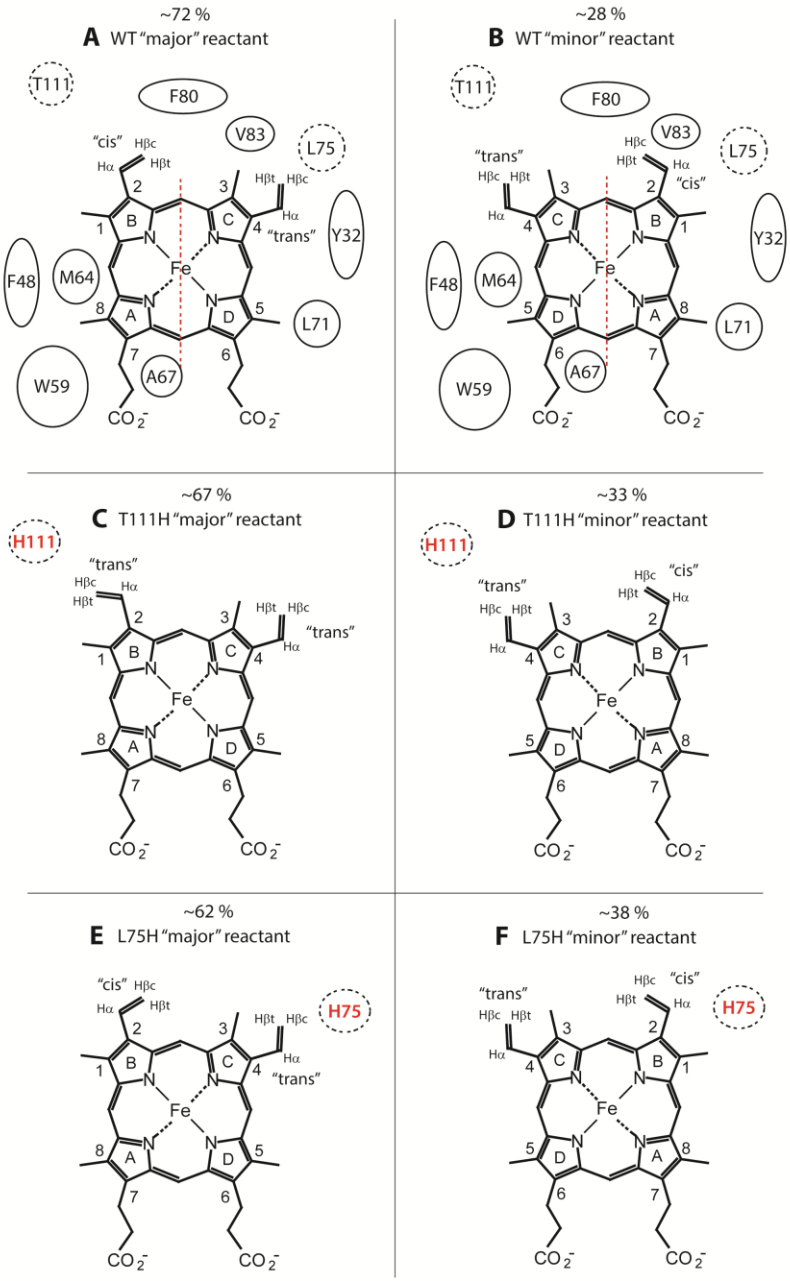


Figure S6.12. Structural summary of heme orientational isomers in wild-type (A, B), T111H (C, D), and L75H (E, F) cyanomet CtrHbs. The relative populations of each isomer, as estimated from integration of resolved ^{1}H 1D peak, are given in percent. These populations did not change appreciably over the course of data collection indicating extremely slow exchange. For the WT CtrHb major and minor isomers (top row), several residues in proximity to the heme are depicted as ellipses. Within the left column (major heme isomers), the T111H variant displays a distinct heme 2-vinyl orientation with respect to WT and L75H CtrHbs. The minor heme isomers (right column) all displayed similar intra-heme NOEs, which supports a common vinyl configuration. In each depiction, the proximal histidine is above the page toward the reader.

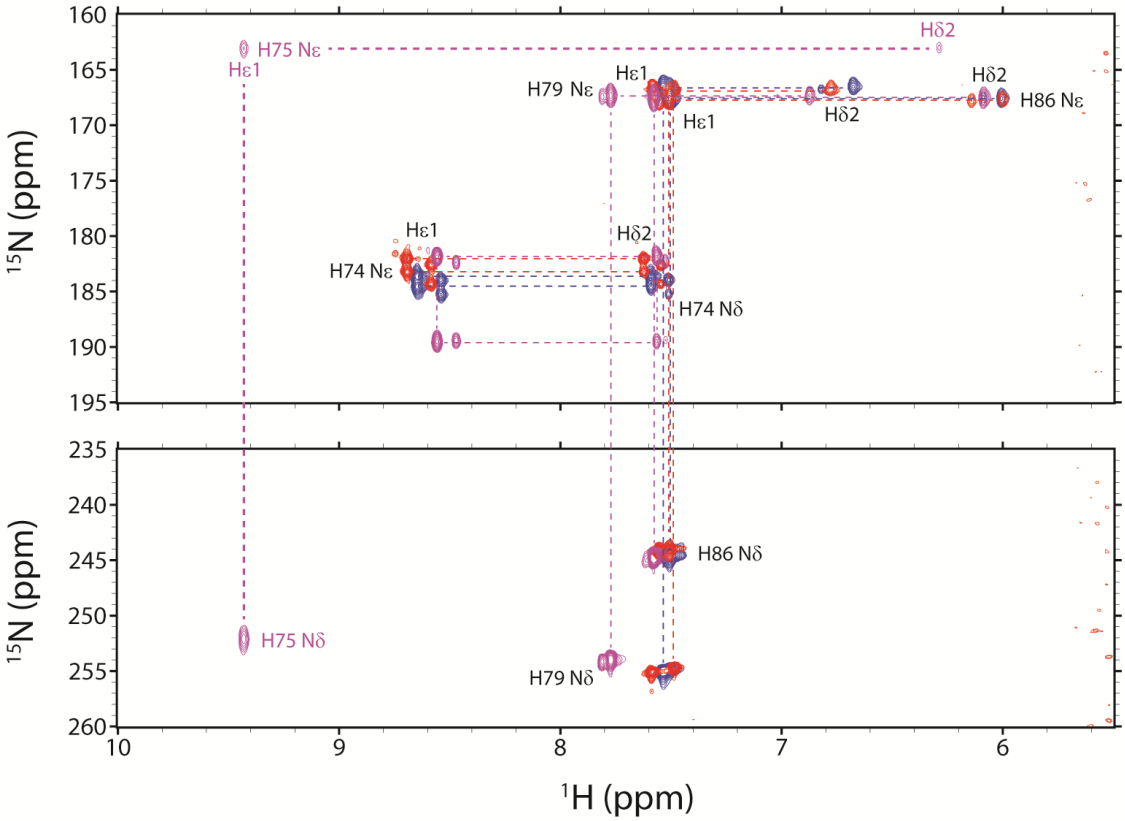


Figure S6.13. ^{1}H-^{15}N long-range HMQC ($1/2J_{NH}$ ~ 22 ms) spectral overlay of WT, T111H, and L75H cyanomet CtrHbs. All non-proximal histidine residues (H74, H79, and H86) are assignable in cyanomet WT CtrHb (blue peaks, black labels). For the T111H variant (red peaks), resonances corresponding to H74, H79, and H86 overlay well with the wild-type, but additional signals corresponding to H111 are not observed, likely because of exchange broadening. On the other hand, the L75H variant (magenta peaks) exhibits a new, highly shifted set of resonances attributable to H75 (magenta labels), in addition to minimally perturbed H74, H79, and H86 signals. The unusual chemical shifts observed for H75 protons are consistent with its close proximity to the paramagnetic (S = 1/2) heme.

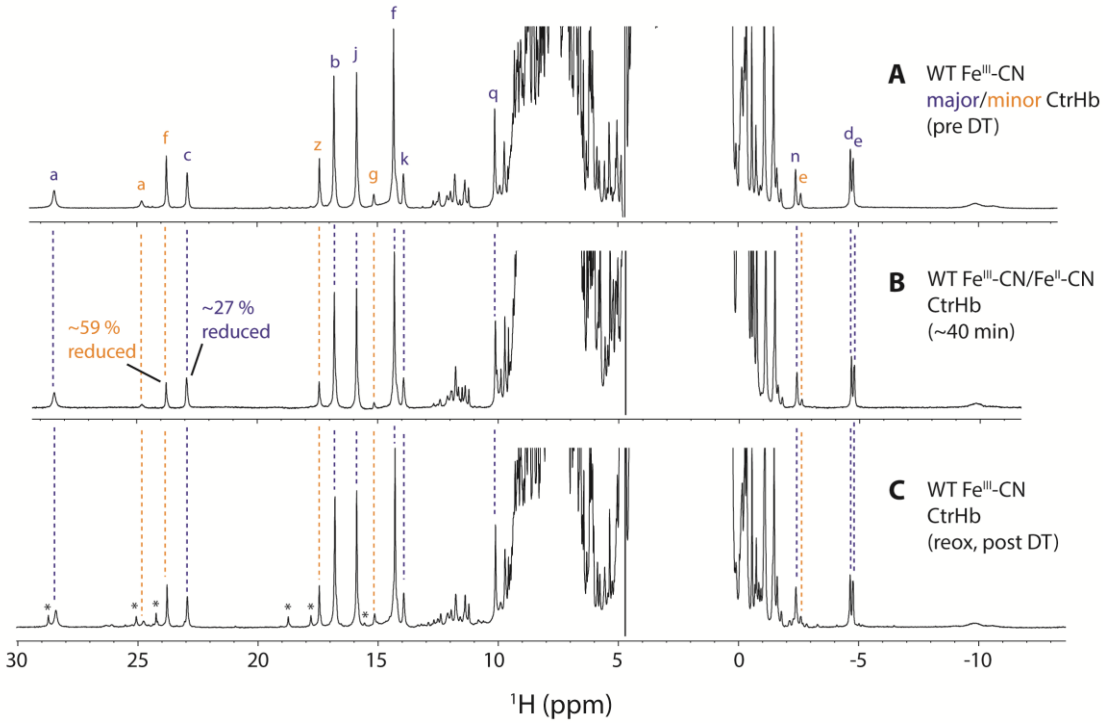


Figure S6.14. ^{1}H NMR spectra monitoring DT reduction of cyanide bound WT CtrHb (^{15}N amide decoupled). (A) Cyanomet (Fe^{III}-CN) ^{15}N-labeled protein spectrum prior to DT reduction. Assignments for select heme and protein ^{1}H signals are included. Peaks are labeled as in Figure 6.2 of the main text: a, Tyr20 OηH; b, heme 1-CH_3; c, heme 2-vinyl Hα; d, e, heme 2-vinyl $H\beta_{cis}$, $H\beta_{trans}$; f, heme 3-CH_3; g, heme 4-vinyl Hα; j, heme 5-CH_3; k, heme 6-propionate Hα, Hα'; n, heme 7-propionate Hβ; q, Tyr20 CεHs; z, heme 8-CH_3. The major heme orientational isomer (blue labels) occurs at a ~2.6:1 ratio over the minor isomer (orange labels). (B) Mixture of cyanomet (Fe^{III}-CN, paramagnetic) and ferrous cyanide (Fe^{II}-CN, diamagnetic) proteins ~40 min following reduction with 6 mM DT. The dashed lines connect remaining cyanomet signals. Assuming no damage occurred to the protein, the decrease in cyanomet peak intensities indicates that approximately 27% of the major isomer was reduced. Interestingly, the minor isomer displays increased susceptibility to reduction (~59% reduced), suggesting that the major and minor heme orientational isomers differ in their ability to stabilize exogenous cyanide in the ferrous state. (C) Following ~6 h DT reduction, the sample was completely reoxidized. The predominant forms detected were identical to those observed prior to reduction indicating minimal reaction (dashed lines). However, an additional minor species is detected (asterisks) in an indication of heme side reaction. A–C. Sample conditions: pH ~7.1, 10 % D_2O, 298 K.

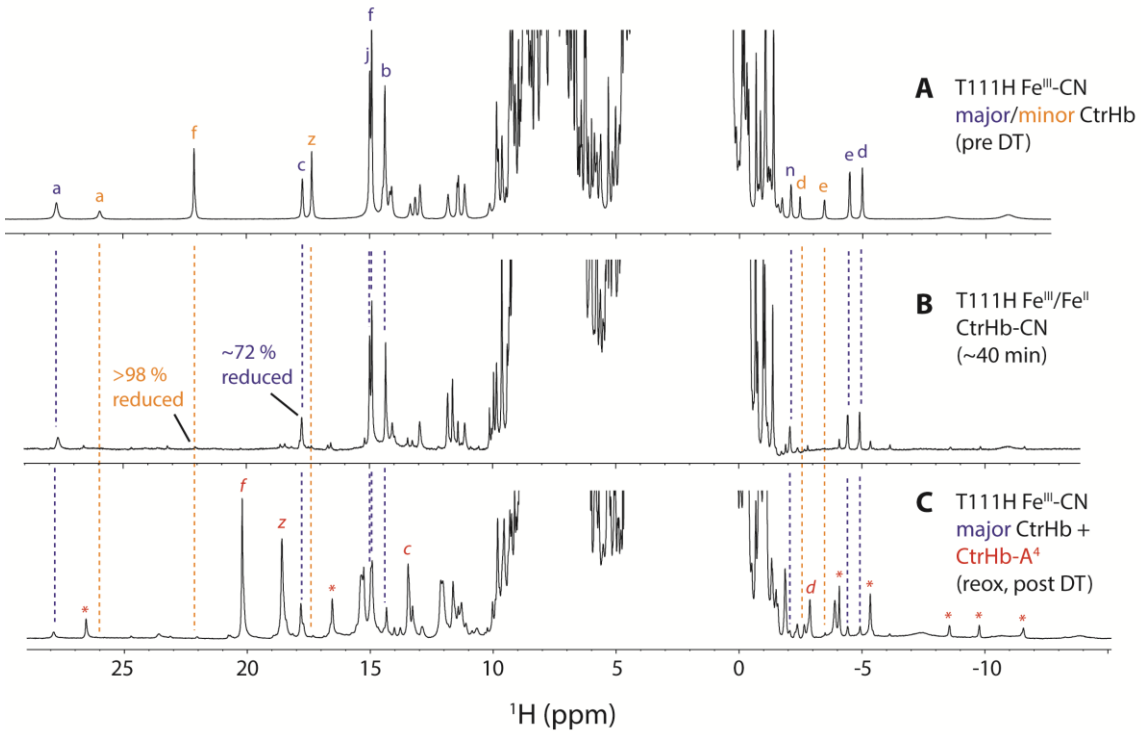


Figure S6.15. ^{1}H NMR spectra following DT reduction of cyanide bound T111H CtrHb. (A) Cyanomet (Fe^{III}-CN) protein spectrum before DT reduction. Assignments for select heme and protein ^{1}H signals are included. Peaks are labeled as in Figure 6.2 of the main text: a, Tyr20 OηH; b, heme 1-CH_3; c, heme 2-vinyl Hα; d, e, heme 2-vinyl Hβ_{cis}, Hβ_{trans}; f, heme 3-CH_3; j, heme 5-CH_3; n, heme 7-propionate Hβ; z, heme 8-CH_3. The major heme orientational isomer (blue labels) occurs at a ~2:1 ratio over the minor isomer (orange labels). (B) Mixture of cyanomet (Fe^{III}-CN, paramagnetic) and ferrous cyanide (Fe^{II}-CN, diamagnetic) T111H proteins ~40 min following reduction with 6 mM DT. The dashed lines connect remaining cyanomet signals. Under the assumption of no protein damage, the decrease in cyanomet peak intensities indicates that approximately 72% of the major isomer was reduced. In contrast, virtually none of the cyanomet minor isomer remains. Therefore, as in WT CtrHb, the T111H minor heme orientational isomer displays an enhanced susceptibility to DT reduction when compared to the major isomer. However, both isomers of T111H CtrHb are reduced more readily than WT CtrHb. (C) Following ~7 h DT reduction, the sample was completely reoxidized. The predominant form (~60%, red labels) corresponds to the major reaction product (cyanomet T111H CtrHb-A^4). Additionally, the unreacted major isomer is observed (blue dashed lines). At least one other form (red asterisks) is noted. This form was not studied further, but we note that the shifts do not suggest formation of T111H CtrHb-A^2. Sample conditions: pH ~7.1, 10 % D_2O, 298 K.

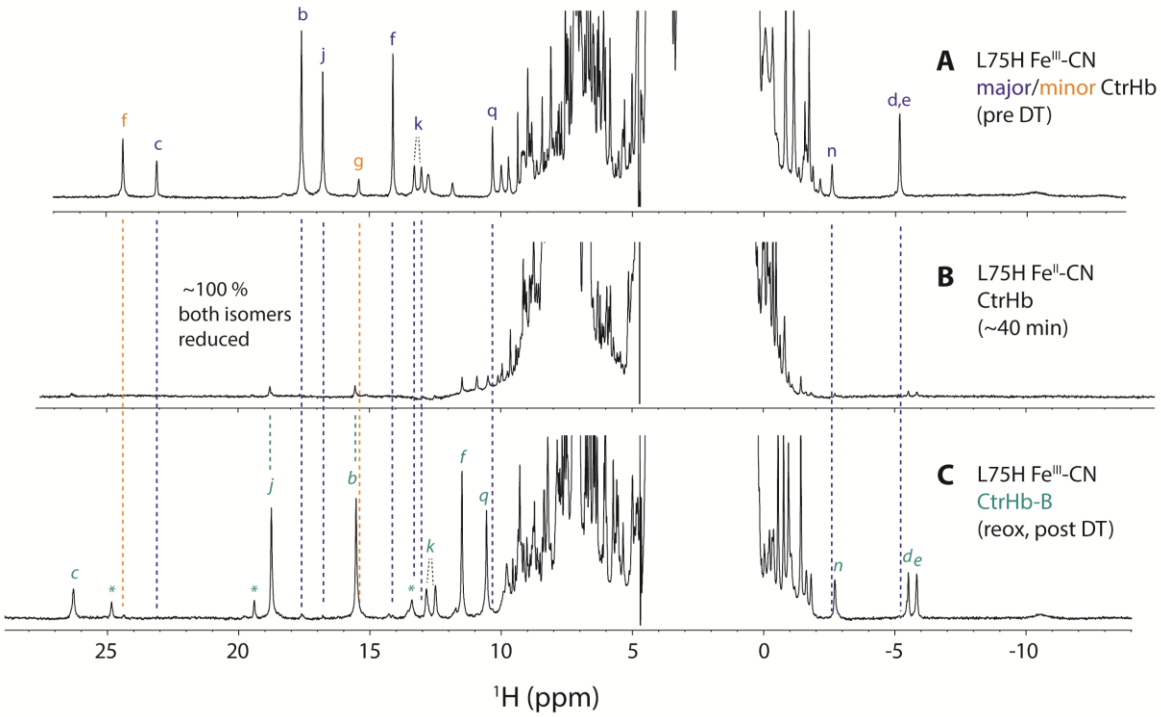


Figure S6.16. ^{1}H NMR spectra following DT reduction of cyanide bound L75H CtrHb. (A) Cyanomet (Fe^{III}-CN) protein spectrum prior to DT reduction (100% D_2O, pH* = 7.4). Assignments for select heme and protein ^{1}H signals are included. Peaks are labeled as in Figure 6.2 of the main text: a, Tyr20 OηH; b, heme 1-CH_3; c, heme 2-vinyl Hα; d, e, heme 2-vinyl $H\beta_{cis}$, $H\beta_{trans}$; f, heme 3-CH_3; g, heme 4-vinyl Hα; j, heme 5-CH_3; k, heme 6-propionate Hα, Hα'; n, heme 7-propionate Hβ; q, Tyr20 CεHs; z, heme 8-CH_3. The major heme orientational isomer (blue labels) occurs at a ~3:2 ratio over the minor isomer (orange labels). (B) The sample was exchanged into H_2O (pH 7.3) and reduced with ~8 mM dithionite. After ~40 min, the sample contained no detectable resonances attributable to either the major or minor cyanomet reactants. Thus, unlike both T111H and WT cyanomet CtrHbs, the L75H variant is completely reduced to the ferrous cyanide bound form. (C) Following ~2 h DT reduction, the sample was completely reoxidized and exchanged back into D_2O (pH* 7.2). Both major and minor isomer reactant resonances had essentially vanished (orange and blue dashed lines) but a new set of resonances, corresponding to > 90% of the sample was observed and attributed to the primary product: L75H cyanomet CtrHb-B (cyan labels). A minor product (< 10%, cyan asterisks) was also detected but not examined in this work. Preliminary evidence (lr-HMQC) suggests the minor product in L75H is H75 Nε2–2-Cα linkage. All data were collected at 298 K.

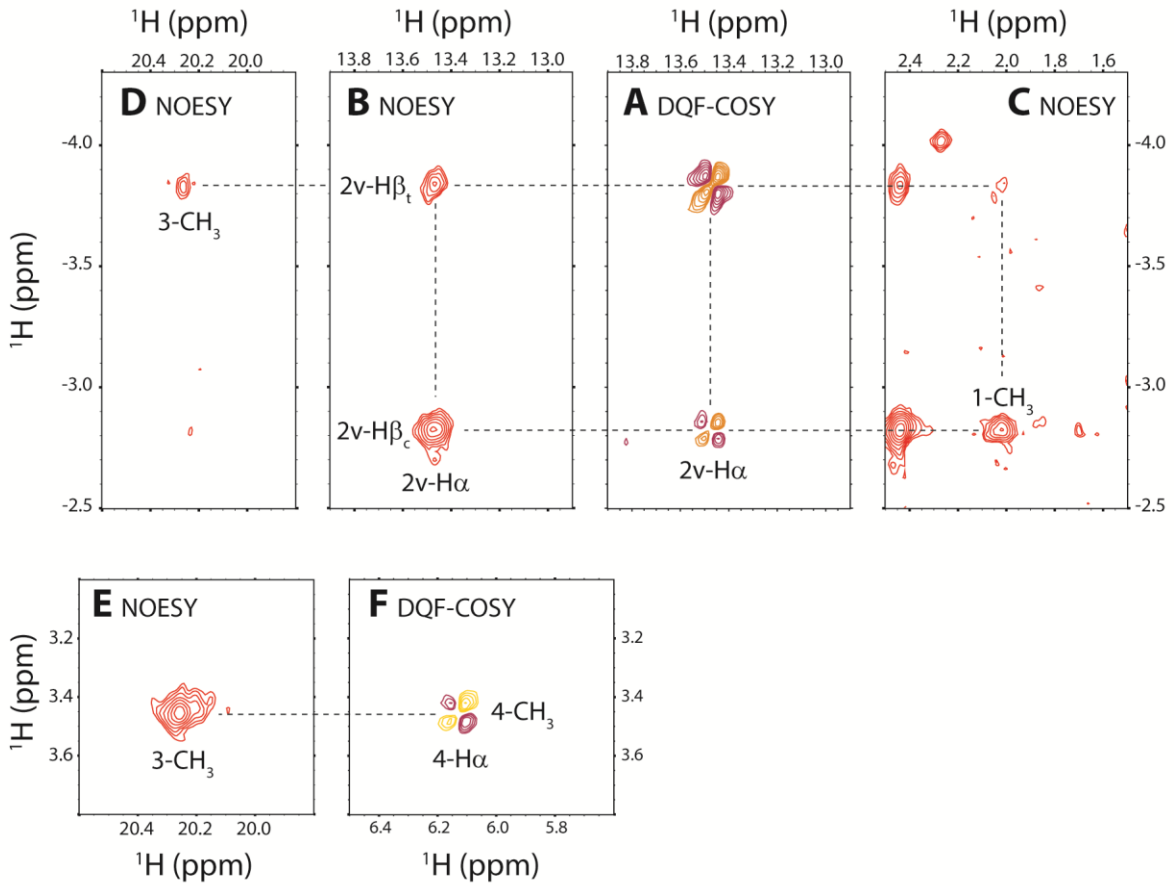


Figure S6.17. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme assignments in cyanomet T111H CtrHb-A^{4} (covalent adduct). (A) *J*-correlated resonances assigned to the 2-vinyl Hα and Hβs. (B) Corresponding region showing NOEs in support of the 2-vinyl Hβ_{cis} and Hβ_{trans} assignments. (C) The 2-vinyl Hβ_{cis} is in dipolar contact with the heme 1-CH_3. (D) The 2-vinyl Hβ_{trans} displays an NOE to the heme 3-CH_3. Together, these data support a *cis* or "twist" configuration for the 2-vinyl group. No additional vinyl group was detected. (E) The 3-CH_3, which typically displays NOEs to the heme 4-vinyl, instead showed a strong contact with a methyl group (~3.4 ppm), in agreement with a modification at the heme 4-substituent. (F) The tentative 4-heme methyl group displayed a *J*-correlation to a signal at (~6.1 ppm), suggesting the modification of the vinyl CαH=CβH_2 group into a CαH-CβH_3 substituent.

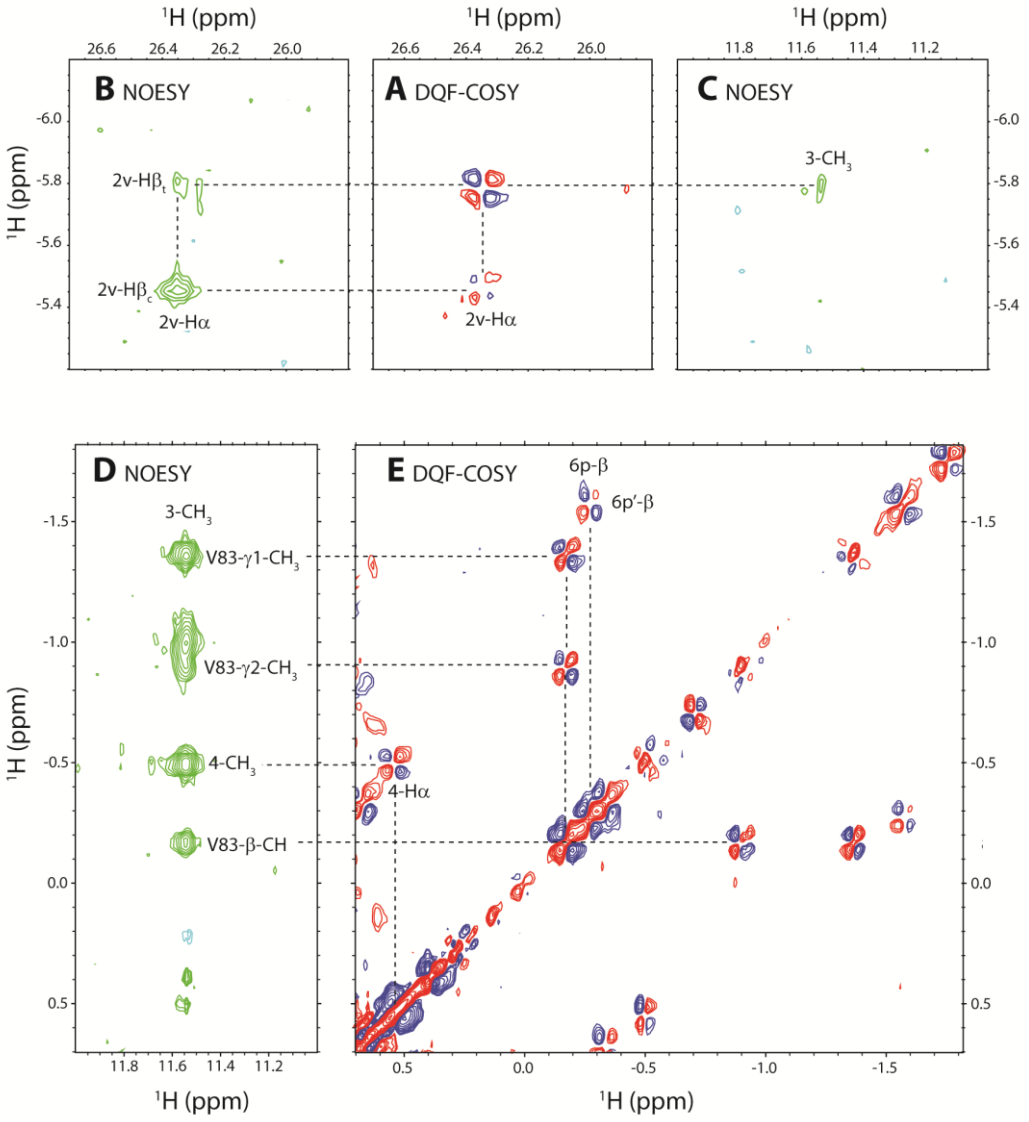


Figure S6.18. Portions of ^{1}H-^{1}H NOESY/DQF-COSY spectra for heme assignments in L75H cyanomet CtrHb-B (covalent product). (A) *J*-correlations exhibited between 2-vinyl protons. (B) Corresponding region showing NOEs in support of 2-vinyl Hβ_{cis} and Hβ_{trans} assignments. (C) The 2-vinyl Hβ_{trans} displays a weak NOE to the heme 3-CH_3. The intra-heme connectivities are consistent with the 2-vinyl group adopting a *cis* or "twist" orientation. (D) The upfield methyl protons (–0.5 ppm) *J*-correlated to His75 in the ^{15}N histidine-selective lr-HMQC spectrum (Figure 6.5, main text) displays a strong NOE to the heme 3-CH_3. This contact, along with the absence of a second vinyl system in DQF-COSY data, indicates that the 4-substituent has undergone modification. (E) The heme 4-CβH_3 displays a single *J*-correlation assigned to the heme 4-CαH.

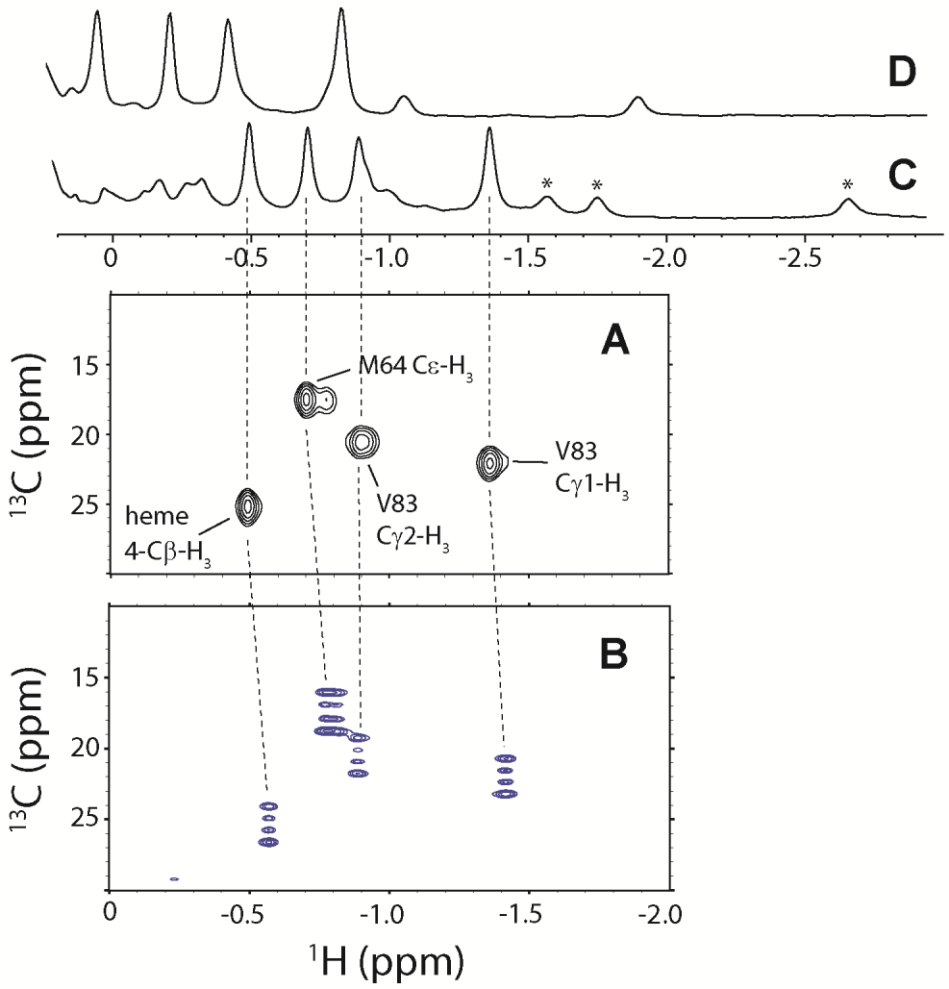


Figure S6.19. Identification of the resolved peak at a ^{1}H shift of –0.49 ppm as a methyl group. (A) The ^{1}H-^{13}C HMQC data shown in Fig. 6. (B) The corresponding coupled ^{1}H-^{13}C HSQC data clearly showing a 3:1:1:3 quartet at –0.49 ppm. (C) The ^{1}H 1D data at 298 K. (D) The ^{1}H 1D data at 313 K. The region shown contains the methyl groups of Val83 and Met64. All four methyl groups are highly sensitive to temperature (Curie effect) as illustrated in (D). (A) and (B) were collected on different samples and different spectrometers. Slight changes in temperature and pH explain the small shifts in the proton dimension. For comparison, the peaks marked with asterisks correspond to heme propionate signals with single proton intensity.

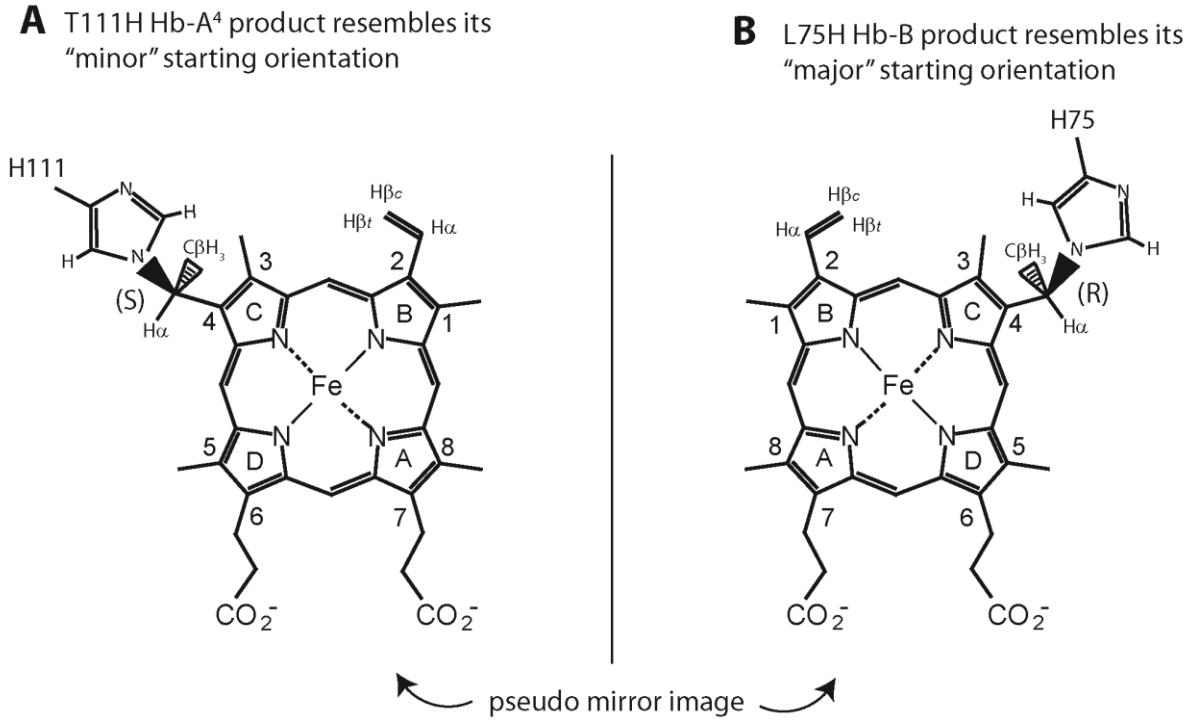


Figure S6.20. Proposed structures of the histidine-heme modifications in (A) T111H CtrHb-A^4 and (B) L75H CtrHb-B. The T111H product is tentative as no long-range coupling was observed between the modified His111 Nε and the heme. However, NOEs between the His111 Hδ2 and Hε1 are consistent with the orientation depicted. With regards to the L75H product, the long-range ^{15}Nε2-CβH$_3$ *J*-correlation establishes that modification occurred at His75 Nε2. Since a methyl group was detected at the 4-substituent, a Nε2-Cα– CβH$_3$ His-heme linkage can be inferred. As shown, the model is consistent with NOEs observed between the modified H75 side chain and covalently bound heme.

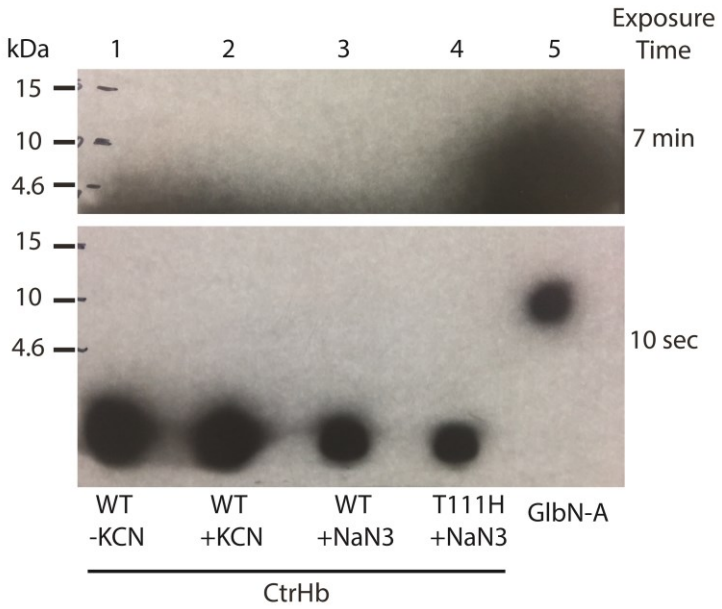


Figure S6.21. ECL SDS-PAGE showing no heme attachment in T111H CtrHb when azide is used in place of cyanide. Five-fold excess KCN (+KCN) or 10 mM NaN_3 (+NaN_3) was added to the ferric form of the designated protein and incubated for 20 min. Subsequently, 2 mM DT was added to each sample and incubated for 3 h at RT, followed by passage over a ~1.3 cm^3 DEAE column. The figure shows film that was exposed for either 10 s or 7 min. As expected, all of the WT CrtHb samples (lanes 1–3) showed the heme migrating with the dye front. The T111H CtrHb +NaN_3 sample (lane 4) also showed heme only migrating with the dye front. No crosslinked heme was detected in this sample even after the 7 min exposure. To account for possible inhibition of peroxidase activity by azide, a sample of the more reactive L75H CtrHb was treated with azide/DT, then purified by DEAE chromatography. The protein lost heme on the column. This demonstrates that, unlike cyanide, azide does not facilitate crosslink formation in T111H or L75H CtrHb. Crosslinked GlbN (GlbN-A, lane 5) was used as a control.

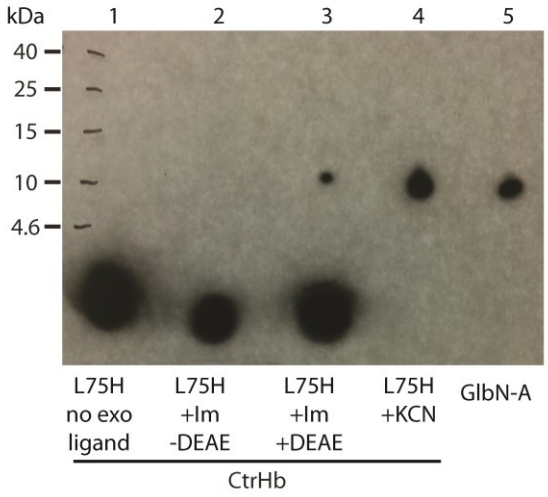


Figure S6.22. ECL SDS-PAGE demonstrating partial modification of L75H CtrHb when imidazole is used instead of cyanide. Ten-fold excess imidazole (+Im) or 5-fold excess KCN (+KCN) was added to L75H and incubated for at least 4 h or 20 min with imidazole or cyanide, respectively. All L75H CtrHb samples, except the L75H CtrHb sample to which no exogenous ligand was added (lane 1), were subsequently reduced with 2 mM DT for 3 h. The L75H CtrHb +Im/-DEAE sample in lane 2 does not show heme migrating with the protein. However, this sample was not passed over a DEAE column to strip some of the imidazole from the heme. Thus, when compared to the L75H CtrHb +Im/+DEAE sample (lane 3) that was cleaned by passage through a short DEAE column, imidazole is shown to inhibit the ECL reaction. In contrast to L75H CtrHb that was bound with cyanide prior to reduction (lane 4), the imidazole sample that was cleaned by passage through DEAE (L75H +Im/+DEAE, lane 3) only shows partial crosslinking. *Synechocystis* 6803 GlbN-A (lane 5) and L75H CtrHb with no exogenous ligand added were used as positive and negative controls, respectively. The film was exposed for 7 min.

Supporting Information References

[1] M. Milani, Y. Ouellet, H. Ouellet, M. Guertin, A. Boffi, G. Antonini, A. Bocedi, M. Mattu, M. Bolognesi, P. Ascenzi, Biochemistry 43 (2004) 5213-5221.

[2] M. Couture, M. Guertin, Eur. J. Biochem. 242 (1996) 779-787.

Chapter 7.

The NMR hydrogen bond scalar coupling as a sensitive probe for helix capping N-H···N hydrogen bonding in heme proteins

Matthew R. Preimesberger, Ananya Majumdar, Selena L. Rice, Lauren Que, and Juliette T.J. Lecomte

Author contributions:

MRP: designed research, assisted with protein preparation, coded NMR pulse sequences, performed NMR data acquisition and analysis, wrote the chapter

AM: advised on NMR data acquisition and interpretation

SLR: assisted with protein preparation

LQ: prepared cytochrome b_5 sample, assisted with NMR data acquisition

JTJL: designed research, analyzed data, wrote the chapter

Abbreviations: trHb, truncated hemoglobin; $^{n}J_{xy}$, scalar coupling between nuclei x and y separated by n bonds; HBC, hydrogen bond scalar coupling; $^{2h}J_{NN}$, two-bond N-N HBC; DT, sodium dithionite; PTM, posttranslational modification; GlbN, *Synechocystis* or *Synechococcus* hemoglobins without heme PTM; GlbN-A, GlbN with PTM between His117 Nε2 and heme 2-Cα; GlbN-B, GlbN with PTM between His79 Nε2 and heme 4-Cα; GlbN-AB, GlbN with PTM between His117 Nε2 and heme 2-Cα and a bond between His79 Nε2 and heme 4-Cα; CtrHb, heme domain of *C. eugametos* L1637 hemoglobin; THB1, hemoglobin 1 from *C. reinhardtii*; lr, long-range; H-bond, hydrogen bond; Hb, hemoglobin; τ, mixing time

Abstract

Although ubiquitous and absolutely essential to the structural and catalytic properties of biomacromolecules, hydrogen bonds (H-bonds) are often inferred through indirect methods. At medium resolution (2 Å), hydrogens within a X-ray diffraction structure are unobservable, thus the length and orientation of an H-bond is defined primarily using theoretical ideal geometry. NMR detection of hydrogen bond scalar couplings (HBCs) enables identification of donor, acceptor, and shared proton within an H-bond, at a site-resolved basis. These HBCs are extremely sensitive to changes in H-bond geometry, and are therefore a useful probe for perturbation of H-bonds in solution.

Here, we report the direct detection of amide N-H···Nδ histidine helix capping H-bonds in a series of model heme proteins by measurement of $^{2h}J_{NN}$ HBCs using HNN-correlation and spin-echo difference experiments. The set of proteins includes well-studied monomeric truncated hemoglobins (trHbs) from the eukaryotic algae *C. eugametos* (CtrHb), *C. reinhardtii* (THB1), the cyanobacteria *Synechococcus* sp. PCC 7002, *Synechocystis* sp. PCC 6803 (GlbNs), and the classic electron transfer protein rat microsomal cytochrome b_5. We unequivocally demonstrate that the wild-type trHbs contain a conserved H-bond capping the G-helix on the proximal side of the heme pocket. Cytochrome b_5 contains a distinct 3_{10}-helix capping N-H···Nδ interaction between D82 andH80. Accurate measurement of the $^{2h}J_{NN}$ coupling for N-H···Nδ H-bonds in wild-type proteins yielded a relatively constant magnitude ~5 Hz, independent of sequence context, secondary structure, or ligation of state of the heme iron (*bis*-histidine vs. histidine-cyanide).

GlbNs have the unusual ability among hemoglobins to bind covalently their heme group. The wild-type linkage connects H117 Nε2-heme 2-Cα and does not affect the magnitude of $^{2h}J_{NN}$ in neither cyanide bound nor *bis*-histidine GlbN forms. The His-vinyl modification can be engineered within different positions of GlbN (e.g. L79H variant) and has also been transplanted into CtrHb (L75H variant). In these instances, the engineered crosslink is adjacent to the helix capping N-H donor. HNN-COSY experiments show that in each case the N-H···Nδ H-bond was maintained. However, measurement of $^{2h}J_{NN}$ couplings in variants with engineered crosslink showed distinct effects on the N-H···N H-bond (*strain* vs. *optimization*), dependent on the background protein (cyanide bound L79H/H117A GlbN vs. L75H CtrHb) and heme iron ligation status (L79H/H117A GlbN *bis*-histidine and cyanide bound forms).

Overall, the model heme protein data display a positive correlation between the magnitude of $^{2h}J_{NN}$ (Hz) and ^{1}H chemical shift (ppm) of the amide donor. Correlations between amide ^{1}H and ^{15}Nδ1 histidine chemical shifts and $^{2h}J_{NN}$ and histidine ^{15}Nε2 are explained using a tautomeric bias argument: stronger N-H···N H-bonds (with larger $^{2h}J_{NN}$ and downfield ^{1}H shift) can only form with an Nδ acceptor and therefore select for the neutral Nε-H histidine tautomer of histidine leading to a large ^{15}Nδ1 downfield shift (and upfield ^{15}Nε2 shift). Comparison of the measured heme protein $^{2h}J_{NN}$ values with those of N-H···N H-bonds in DNA ($^{2h}J_{NN}$ ~ 6-10 Hz) suggests that the former interactions are on average slightly longer, and possibly weaker. Both amide N-H···Nδ1 H-bonds in proteins and imino N-H···N H-bonds in nucleic acids show a similar slope between $^{2h}J_{NN}$ and ^{1}H chemical shift (average ~1.1 Hz/ppm), in support of the

classic ^{1}H deshielding metric for the presence and length of H-bonds. Overall, the results indicate the NMR HBCs can be used to study the subtle deformation and relaxation of individual H-bonds in proteins.

Introduction

The folding and structural organization of biomacromolecules is defined in large part by the three-dimensional arrangement of hydrogen bonds (H-bonds). Binding interactions and conformational changes that alter the H-bonding patterns within proteins and nucleic acids are observable by many experimental techniques. Often neglected from consideration are structural elements where no apparent alteration in H-bonding occurs. In these instances, the strain or relaxation of individual hydrogen bonds may be energetically and functionally important, for example in mediating allosteric effects. However, perturbations such as the minute lengthening or bending of an H-bond are difficult to establish directly.

Nuclear magnetic resonance (NMR) spectroscopy enables the site-resolved identification of NMR-active nuclei involved in H-bonds via hydrogen bond scalar couplings (HBCs).[1–3] Using HBCs, it is possible to correlate both donor *and* acceptor heavy atoms, through their shared hydrogen.[4] The relatively large two-bond coupling, $^{2h}J_{NN}$, has been especially useful for the unambiguous assignment of Watson-Crick and non-canonical nucleic acid base pair N-H···N H-bonds.[5] In proteins, the predominant H-bond is of the backbone N-H···O=C' type, and the three-bond coupling, $^{3h}J_{NC'}$, has been similarly useful but with rare exception[6], are generally limited to small proteins.[7,3,8–10] HBCs detected from H-bonds involving ^{15}N-^{1}H···O=^{13}C ($^{3h}J_{NC'}$, and $^{2h}J_{HC'}$[11]) and ^{15}N-^{1}H···^{15}N ($^{2h}J_{NN}$, $^{1h}J_{HN}$) nuclei report on subtle changes in H-bonding geometry; both $^{3h}J_{NC'}$ and $^{2h}J_{NN}$ couplings are especially sensitive as they are proposed to be exponentially related to H-bond distance.[12,13] Thus, in addition to direct identification,

measurement of the magnitude of $^{2h}J_{NN}$ (or $^{3h}J_{NC'}$) provides a convenient proxy for the length, and possibly, the relative strength of individual H-bonds in solution.[14] In nucleic acids, interesting differences in N-H···N H-bonding can be detected between different types of base pairs. For example, comparison of ultra-high-resolution crystallographic structures of model A-U and G-C base pairs supports that the former have *slightly* shorter N-N H-bonding distances (~0.05 Å).[15,16] This observation appears to be corroborated when viewed from an individual hydrogen bond in solution, namely the central imino N3-H···N1 H-bond within an RNA (DNA) A–U (A-T) base pair typically displays $^{2h}J_{NN}$ values between 6–8 Hz, on average larger than the imino N1-H···N3 H-bond found in G–C pairs (5–7 Hz).[1,2,17] Similarly, cooperative networks of amino N2-H···N7 H-bonds found in G-quadruplex DNA display $^{2h}J_{NN}$ values of 6-8 Hz.[18] Hoogsteen base C-G pairs involving protonated cytidine ($N3^+$-H···N7) display the largest $^{2h}J_{NN}$ values yet measured in a biomacromolecule (10–11 Hz).[19] On the other hand, A-A mismatches, which have amino N6-H···N7 type H-bonds, tend to display very low $^{2h}J_{NN}$ values (2–3 Hz), in support of their non-ideal geometry.[20] The aforementioned examples highlight the extraordinary sensitivity of $^{2h}J_{NN}$ couplings to the local H-bonding environment. In proteins, N-H···N H-bonds are much less common and relatively few measurements of $^{2h}J_{NN}$ couplings have been reported.[21,22]

Owing to the N-H (i) ···O=C (i-4) pattern of H-bonds within a protein α-helix, N-terminal amide donors within the initiating turn do not have an available α-helical carbonyl acceptor.[23,24] Therefore, one situation in which N-H···N protein H-bonds do occur is through main chain N-H···Nδ1 histidine side chain helix capping interactions.[25] These H-bonds typically act to cap the N-termini of α-

helices and involve the amide of residue i and the histidine at position i+3. In general, N-capping interactions play an important role in determining the precise boundaries of helical structures.[26]

Recently, we reported the direct detection of threonine amide N-H···Nδ1 histidine α-helix capping H-bonds within the consensus ankyrin repeat (AR) TXXH motif.[27] In the study, we examined constructs containing three or four ARs and found that the magnitude of $^{2h}J_{NN}$ varied from repeat to repeat. Specifically, N-H···Nδ1 H-bonds within buried internal repeats yielded larger couplings ($^{2h}J_{NN}$ ~ 4 Hz) relative to the solvent exposed N-H···Nδ1 interactions ($^{2h}J_{NN}$ ~ 2 Hz). We also observed a corresponding trend in histidine pK_a from repeat to repeat: namely, H-bonds with larger $^{2h}J_{NN}$ values were associated with histidines having strongly depressed pK_a (< 3). Since formation of a cationic histidine requires cleavage of the N-H···Nδ1 His interaction, the observations led us to conclude that $^{2h}J_{NN}$ may serve as a reasonable proxy for relative N-H···N H-bond strength.[27] In the current study, we extend the methods used for N-H···N H-bond detection in AR proteins to a series of heme proteins, and systematically examine the strength of a conserved N-H···N H-bond within different structural contexts.

Previously, we have pointed out the presence of a conserved helix capping N-H···N hydrogen bond within several members of the Group I truncated hemoglobin (TrHb) lineage. The proposed interaction involves a donor amide N-H (residue i) and acceptor histidine Nδ1 (i+3), which initiates the G-helix on the proximal side of the heme pocket near the heme 4-vinyl group (Figure 7.1A). The conserved main chain-side chain N-cap is present in TrHbs from the

unicellular eukaryotes *Chlamydomonas eugametos* (CtrHb, S76-H79)[28] and *C. reinhardtii* (THB1, N87-H90)[29], and the cyanobacteria *Synechocystis* sp. PCC 6803[30–32] and *Synechococcus* sp. PCC 7002[33,34] (GlbNs, N80-H83, and T80-H83 respectively), and predicted in many others. These homologous TrHbs (~40–50% identity) display the same fold but offer a variety of sequence contexts for the expected amide N-H···N histidine H-bond. Additionally, TrHbs undergo a large conformational transition upon exogenous ligand binding to the distal iron site.[35–37,34,29] The regions most affected by this structural rearrangement (the B- and E-helices) are remote from the helix capping H-bond (Figure 7.1A), and therefore offer a favorable situation to assess subtle long-range coupling between the distal and proximal sides of the heme.

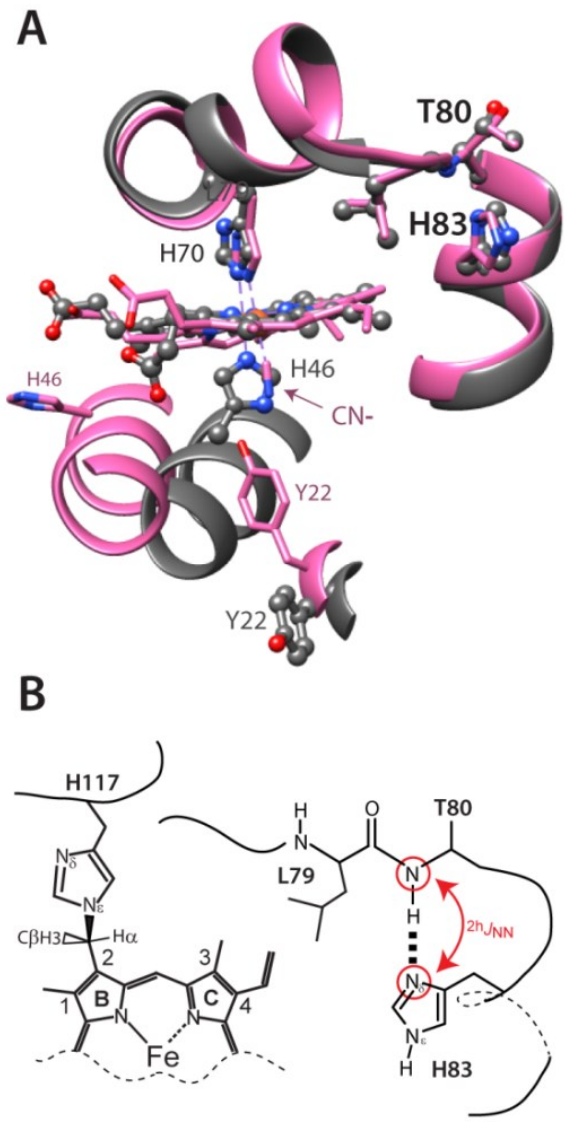


Figure 7.1. A. Alignment of *bis*-histidine (4MAX: gray) and cyanomet *Synechococcus* GlbN-A (4L2M: purple) heme and proximal histidine showing the large distal pocket rearrangement upon cyanide binding. The exogenous ligand is stabilized by a hydrogen bond network comprised of Tyr22 (B10), along with Gln43 (E7) and Gln47 (E10, not shown), whereas His46 (E11) has rotated out of the distal pocket. B. Cartoon diagram depicting the short FG loop and Thr80 N-H···Nδ1 His83 H-bond which caps the N-terminus of the G-helix. The red arrow depicts the $^{2h}J_{NN}$ coupling between amide and histidine ^{15}N nuclei. In GlbNs, the site of native heme PTM occurs on the H-helix near the C-terminus of the protein to form the H117 Nε2-2-Cα linkage.

Among all members of the hemoglobin superfamily, GlbNs have the distinctive ability to modify the heme group and form a covalent adduct. The irreversible post-translational modification (PTM) involves a non-coordinating histidine (H117) located on the H-helix near the heme 2-vinyl group.[38] Heme reduction in the absence of oxygen spontaneously leads to the H117 Nε2-2-Cα heme linkage (Figure 7.1B).[39,40] Recent work has demonstrated that analogous His-heme modifications can be engineered within different positions of GlbN (e.g. L79H, or L79H/H117A variants)[41], and implanted into CtrHb (e.g. T111H or L75H variants)[28], therefore enabling detailed examination of how native and engineered covalent crosslinks alter helix N-capping geometry (Supporting Information Figure S7.1). The wide range of perturbations available makes these TrHbs an excellent model set for the detailed interrogation of a conserved amide N-H···Nδ histidine H-bond.

Distinct from the previously discussed α-helix N-caps, histidine can also cap 3_{10}- type helices. In this arrangement, the amide N-H donor occurs at position i+2, relative to the histidine Nδ1 acceptor (i). This second type of N-H···N H-bond is exemplified in the electron transfer protein cytochrome b_5 and involves His80 Nδ1 and Asp82 N-H.[42] The 3_{10} N-cap occurs near the C-terminus of the protein, within the apoprotein folding core, greater than 20 Å from the heme binding pocket.[43] The 3_{10} capping N-H···Nδ1 H-bond in cytochrome b_5 occurs in a distinct motif and allows convenient comparison to the TrHb interactions.

Here, we report direct NMR detection and quantitative comparison of amide N-H···Nδ1 histidine H-bonds of several TrHbs (α-helix cap) and cytochrome b_5 (3_{10}-helix cap). NMR H-bond correlation experiments demonstrate that the

protein N-H···N H-bonds are readily assignable. Accurate measurement of $^{2h}J_{NN}$ HBCs within this set of proteins enables quantitative comparison of how the N-H···N H-bonds vary within different sequence and structural contexts. Furthermore, we highlight instances in which an engineered His-heme covalent modification induces helix capping N-H···N H-bond strain or relaxation.

Materials and Methods

Protein expression and purification

Overexpression and purification of recombinant hemoglobins was achieved as described previously.[28,29,39,41,44] In brief, plasmid DNA encoding either *C. eugametos* CtrHb, *C. reinhardtii* THB1, *Synechococcus* sp. PCC 7002 GlbN or *Synechocystis* sp. PCC 6803 GlbN (and variants thereof) was used to transform *E. coli* BL21(DE3) cells, which were then grown at 37 °C for ~24 h on minimal medium plates. Transformant colonies were used to inoculate minimal medium (M9) starter cultures (~60–100 mL) and grown for an additional ~24 h at 37 °C. At this point, the starter culture was divided into fourths and used to seed 4 × 500 mL M9 expression cultures. To produce uniformly ^{15}N-labeled protein, $^{15}NH_4Cl$ was used as the sole nitrogen source. For *C. eugametos* CtrHb, *C. reinhardtii* THB1 and *Synechocystis* GlbN, expression cultures were allowed to continue growing until an O.D. (550 nm) ~ 0.8 – 1.0 was achieved. Overexpression was induced by addition of 0.5 mM IPTG (final concentration), and was allowed to continue for ~5–8 h prior to cell harvesting by centrifugation. In the case of *Synechococcus* GlbN, leaky overexpression occurs

and the IPTG addition step is not required. Harvested cells were frozen at –20 °C until further use.

The recombinant hemoglobins expressed in this manner partitioned primarily into inclusion bodies. Cells were lysed by sonication; soluble and insoluble fractions were separated by centrifugation, and the lysate was discarded. Inclusion bodies were solubilized in ~8 M urea (50 mM Tris, 1 mM EDTA, pH 8) and passed over a G-50 size exclusion column to achieve apoprotein refolding and purification. Fractions deemed pure by SDS-PAGE were pooled and reconstituted with hemin to generate ferric holoprotein. Passage of the resulting solution over an anion exchange (DEAE) column removed any excess heme and resulted in > 95 % purity in the pooled protein samples. Buffer exchange (into ~1 mM phosphate, pH ~7) was achieved using an Amicon ultrafiltration unit and Millipore YM3 membrane (molecular weight cut off = 3 kDa). The pooled samples were concentrated to ~20 mL and protein concentration was assessed by optical spectroscopy as reported previously. Typical yields of recombinant hemoglobins were ~25–50 mg/L culture. Samples not used immediately for NMR spectroscopic studies were lyophilized and stored at –80 °C until needed.

A similar protocol was used for the expression and purification of ^{15}N-labeled rat microsomal cytochrome b_5. Following overexpression in *E. coli* BL21(DE3) cells, recombinant cytochrome b_5 primarily partitioned to the soluble fraction. An excess of hemin was added to the raw lysate in order to produce ferric holoprotein. Following the heme binding step, ferric cytochrome b_5 purification was achieved via DEAE anion exchange chromatography followed by G-50 size-

exclusion chromatography. Apo cytochrome b_5 was prepared from ferric holoprotein using an acid-butanone heme extraction procedure (1:1 H_2O/butanone, pH 2.5, 4 °C) followed by extensive dialysis and buffer exchange to remove residual butanone. Optical absorbance measurements on the resulting apocytochrome b_5 sample confirmed that > 99% of the heme had been removed.

NMR sample preparation

Lyophilized recombinant protein was re-dissolved in NMR buffer containing 100–200 mM sodium/potassium phosphate, pH ~7.5, 10% D_2O, and protein concentration ranged from approximately 500 μM to 5 mM. Under these conditions, all the heme proteins studied behaved as monomers. Cyanide bound hemoglobins were prepared by addition of 2–5 fold molar excess KCN. Ferrous cytochrome b_5 was produced by dithionite (DT) reduction (5-fold molar excess DT:protein) of the ferric protein. To prevent evaporation or oxidation, protein samples were transferred to NMR Shigemi tubes and sealed with Parafilm prior to NMR data collection.

NMR spectroscopy data acquisition:

NMR spectroscopy was conducted using 600 MHz Bruker Avance or Avance-II spectrometers, each equipped with a cryogenic probe. ^{1}H-^{15}N HSQC, histidine-selective ^{1}H-^{15}N long-range (LR) HMQC[45], Soft ^{1}H-^{15}N HNNδ-COSY[20], and quantitative $^{2h}J_{NN}$ constant-time spin-echo (CTSE) difference 1-D/2-D HSQC[20] spectra were acquired as detailed elsewhere.[27,39] See the Supporting Information Methods for pulse sequence scheme and experimental description.

A typical $^{2h}J_{NN}$-CTSE series consisted of 12 to 20 1-D (or 2-D) experiments using different τ mixing periods (with 2 to 3 duplicate τ periods collected at the beginning and end of the series). Unless otherwise noted, protein NMR samples were studied at 313 K. ^{1}H chemical shifts were referenced with respect to the water line (4.58 ppm at 313 K, 4.76 ppm at 298 K); ^{15}N chemical shifts were referenced indirectly using the Ξ ratio.[46]

NMR data processing, analysis and curve fitting:

NMR data were processed using NMRPipe[47] or Topspin 3.1 (Bruker BioSpin). Spectra were analyzed using the program Sparky 3.[48] For $^{2h}J_{NN}$ modulation data, 1D peak intensities were obtained using the Topspin 2.1 deconvolution (mixed Lorentzian/Gaussian) routine. 2D peak volumes were calculated by peak integration (Sparky 3). Peak intensities were tabulated and plotted as a function of the $^{2h}J_{NN}$ transfer time, τ, and fitted (Kaleidagraph) to the equation: $I(\tau) = A\cos(\pi J\tau)$[49] to extract the initial amplitude (A) and J-coupling ($^{2h}J_{NN}$). In most instances, the fitting error for $^{2h}J_{NN}$ was well below 0.1 Hz. The errors reported for $^{2h}J_{NN}$ combine the error of the fit with an estimate of the error based on identical (τ) experiments collected at the beginning and end of an experimental series (typically <5 %). The $^{2h}J_{NN}$ modulation curves plotted in Supporting Information Figure S7.5 have been normalized by their individually fitted amplitudes to facilitate comparison of peaks with difference intensities on the same y-axis.

Results

In the diamagnetic states of *Synechocystis* and *Synechococcus* GlbNs, the G-helix begins with a capping box. The N-H at the Ncap position (N80 and T80 respectively) is among the furthest downfield-shifted amides (^{1}H > 11 ppm) observed in ^{1}H-^{15}N HSQC spectra.[50,51] Subsequent identification of N87 within *C. reinhardtii* THB1[29] and S76 within *C. eugametos* CtrHb[28] (Figure 7.2A) confirms the highly deshielded environment observed for each Ncap NH. These signals remain at similar downfield shifts both ferric and cyanomet $S = 1/2$ states (not shown)[29,52,53]. Thus, the systematic downfield ^{1}H chemical shift cannot be accounted for by paramagnetism and is instead consistent with N-H···Nδ H-bonding.

Using uniformly ^{15}N-labeled CtrHb bound with cyanide (CtrHb-CN), we first collected ^{1}H-^{15}N HSQC spectra to high ^{15}N resolution (^{15}N acquisition time ~ 300 ms). CtrHb, like other heme proteins, binds heme in two distinct orientations with nearly equal affinity.[28,35,54] The "major" and "minor" isomers are related by a flip about the heme α–γ meso pseudosymmetry axis. The presence of two heme-bound conformations in slow exchange on the chemical shift timescale results in peak doubling for nuclei near the heme group. Supporting information Figure S7.2A illustrates the doubling of the S76 NH signals in CtrHb-CN. The high-resolution HSQC also reveals that both amide ^{15}N signals of S76 (7:3 major:minor isomers) split into partially overlapped doublets (splitting ~ 5 Hz, Supporting Information Figure S7.2A). No other N-H signals displayed this behavior. In addition, the S76 amide doublets collapse into singlets (Supporting Information Figure S7.2B) only when ^{1}H-^{15}N HSQC (^{15}N

acquisition time ~ 300 ms) spectra were acquired with histidine-selective ^{15}N decoupling (250 ppm) during the evolution period. This result supports interaction between amide and histidine ^{15}N nuclei via the two-bond HBC, $^{2h}J_{NN}$. In agreement, histidine-selective ^{1}H-^{15}N LR-HMQC spectra acquired on CtrHb-CN show that H79 is neutral and adopts the Nε2H tautomer with Nδ1 deprotonated and acting as an H-bond acceptor (Figure 7.2B,C).[45] Interestingly, high sensitivity data also demonstrate the presence of a weak cross peak between H79 Nδ1 (^{15}N = 255.1 ppm) and S76 ^{1}H (11.17 ppm), the latter split by $^{1}J_{NH}$ into a doublet (Supporting Information Figure S7.2C). Repeating the high-sensitivity LR HMQC experiment with application of ^{15}N broadband decoupling switched to the amide region (~120 ppm) during direct ^{1}H acquisition leads to complete refocusing of $^{1}J_{NH}$ and collapse of the S76 ^{1}H-^{15}Nδ1 H79 cross-peak (Supporting Information Figure S7.2D). Since these *J*-correlated nuclei are separated by fourteen covalent bonds, the observed signal must arise via the one-bond trans HBC, $^{1h}J_{HN}$.

HNN-COSY detection of N-H···N H-bonding and quantitation of $^{2h}J_{NN}$ in wild-type cyanomet CtrHb

To confirm the presence of the S76 N-H···Nδ1 H79 helix capping H-bond, we utilized a soft ^{1}H-(N)-^{15}N COSY experiment (see Supporting Information Methods).[20] Unlike the LR HMQC, the HNN-COSY pulse sequence utilizes the two-bond homonuclear HBC $^{2h}J_{NN}$, for N-H···N H-bond identification. This approach is advantageous because the $^{2h}J_{NN}$ coupling is often considerably larger than $^{1h}J_{HN}$[2,17] and therefore leads to a gain in sensitivity. Figure 7.2D illustrates the results of such an experiment as conducted on CtrHb-CN.

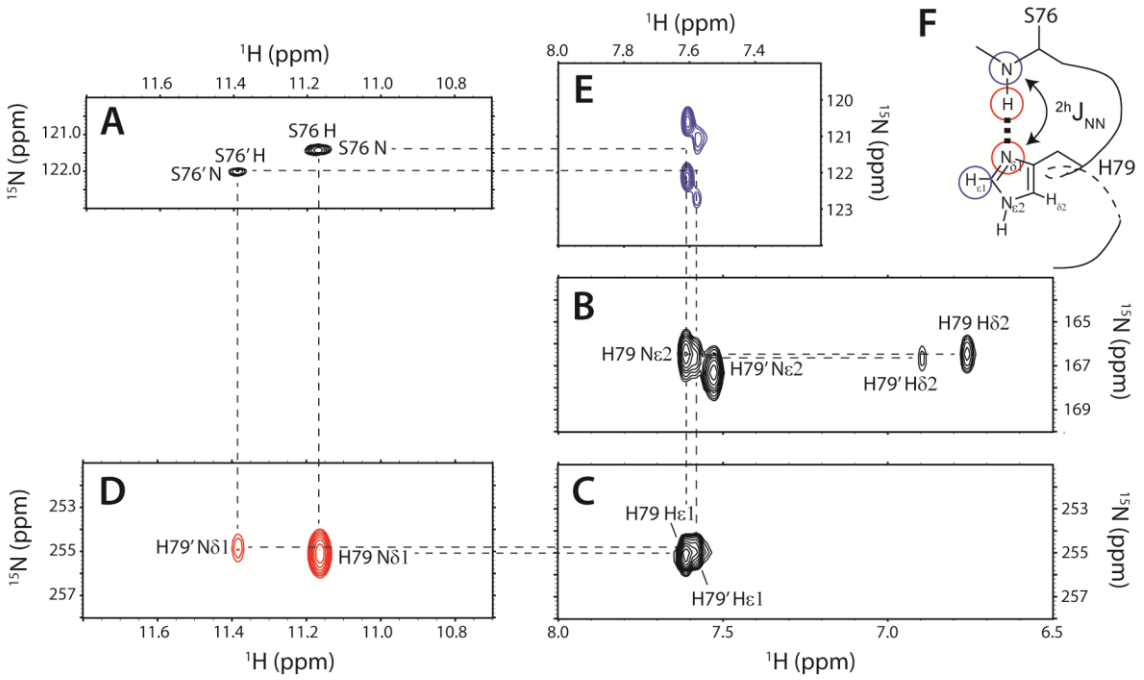


Figure 7.2. A. ^{1}H-^{15}N HSQC (downfield amide ^{1}H region): The large downfield shift of the Ser76 NH amide proton is indicative of N-H···N hydrogen bonding. Peak doubling (Ser76 NH', minor form) is due to slow orientational isomerization of the heme group. The major:minor population ratio is ~7:3. B. ^{1}H-^{15}N LR-HMQC (upfield ^{15}N region): Relatively strong His79 $^{2}J_{NH}$ Nε2–Hε1 and Nε2–Hδ2 intra-imidazole correlations are labeled. C. ^{1}H-^{15}N LR-HMQC (downfield ^{15}N region): The $^{2}J_{NH}$ Nδ1–Hε1 signal completes a Γ cross-peak pattern and indicates that His79 is neutral and adopts the Nε2-H tautomer. D. ^{1}H-(N)-^{15}N-COSY: Amide donor ^{1}H and histidine acceptor ^{15}Nδ1 nuclei are correlated via the homonuclear two-bond hydrogen bond scalar coupling, $^{2h}J_{NN}$. Two signals, corresponding to His79 Nδ1···H–(N) Ser76 nuclei were observed (red spectrum). No other peaks are detected. E. ^{1}H-(Nδ)-^{15}N-(H) LR COSY spectra confirm the N-H···N helix capping H-bond. In this experiment, the amide ^{15}N and histidine Hε1 are correlated via $^{2h}J_{NN}$ within the Ser76 N-H···Nδ1-Hε1 His79 H-bond. As expected from the NMR data shown in A–D, two clear doublets are observed, corresponding to His79 Hε1 - Ser76 N-(H) major and minor isomer signals (blue spectrum). The doublets arise due to the absence of ^{1}H decoupling during the ^{15}N evolution period, and active ^{15}N-^{1}H $^{1}J_{NH}$ coupling (~90 Hz). These signals independently verify the assignment of the S76 N-H···Nδ1 H79 helix capping hydrogen bond. F. Ser76 N-H···Nδ His79 helix capping scheme in *C. eugametos* cyanomet hemoglobin. Correlated nuclei determined by the NMR data shown in E. and F. are highlighted with red and blue circles respectively.

Two clear peaks are present in the spectrum and identify S76 amide ^{1}H and H79 Nδ1 within the N-H···N H-bonds (major and minor). In agreement, the LR HNN-COSY[55] spectrum shown in Figure 7.2E also displays two peaks and correlates H79 Hε1 with S76 N, the latter split by its directly attached proton ($^{1}J_{NH}$ ~ –90 Hz). The NMR connectivity pattern shown in Figure 7.2 is unequivocal evidence for the S76 N-H···Nδ H79 helix capping interaction in CtrHb-CN.

The $^{2h}J_{NN}$ coupling constant is a sensitive reporter for H-bond geometry and possibly strength.[14] We next measured the magnitude of the $^{2h}J_{NN}$ HBC using a high-precision quantitative constant-time spin-echo difference HSQC experiment.[20] With this approach, amide NH groups within N-H···Nδ1 H-bonds undergo $^{2h}J_{NN}$ modulation according to the controlled timing (τ) of a histidine-selective ^{15}Nδ1 inversion pulse. Repeating the experiment for different values of τ yields a modulation curve from which the magnitude of $^{2h}J_{NN}$ can be accurately extracted. Figure 7.3A presents the downfield region of 1-D data collected on CtrHb-CN. The intensities for the resolved S76 NH protons were plotted as a function of the $^{2h}J_{NN}$ modulation time τ, and fit according to the relationship $I(\tau) = A\cos(\pi J\tau)$.

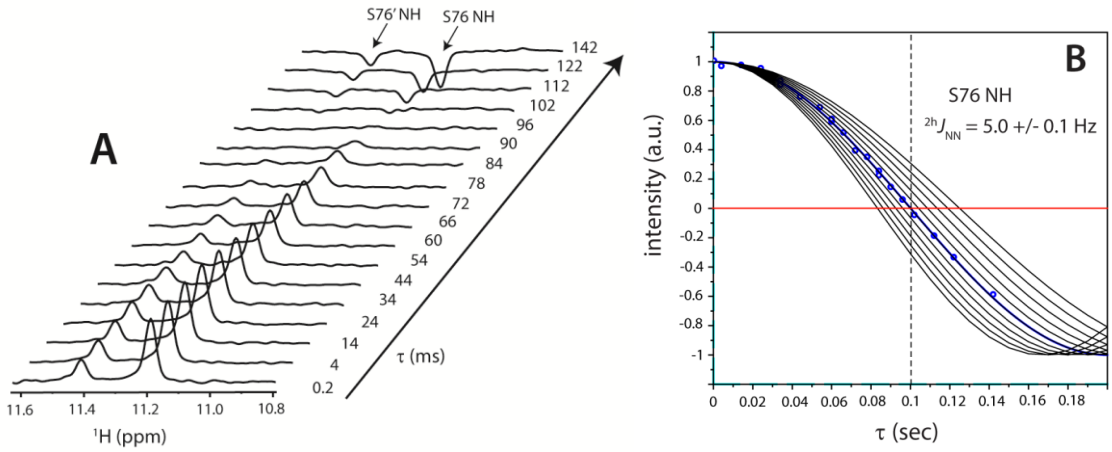


Figure 7.3. A. 1D spectral stack showing the resolved amide proton resonances of S76 (major and minor) as a function of the $^{2h}J_{NN}$-modulation time (τ). All other proton signals are independent of the $^{2h}J_{NN}$ modulation time and superimpose well. B. Peak intensities (blue circles) were obtained from the baseline-corrected data shown in a. and plotted as a function of the $^{2h}J_{NN}$-transfer time. The data were collected in random order and are normalized by the fitted maximum signal amplitude ($^{2h}J_{NN}$ mixing time = 0). Duplicate data points collected at the beginning and end of the experiment were in excellent agreement (< 5% error). A nonlinear least squares fit of the Ser76 NH data yields an accurate determination of the N-H···N H-bond $^{2h}J_{NN}$ = 5.0 +/- 0.1 Hz. The horizontal red line indicates zero intensity, and the black dashed vertical line corresponds to the null time for $^{2h}J_{NN}$ = 5 Hz (100 ms). Simulated $^{2h}J_{NN}$ curves (black lines) between 4 and 6 Hz in 0.2 Hz increments are included for comparison.

As can be seen from Figure 7.3B, the data are of exceptional quality and yield a well-defined value ($^{2h}J_{NN}$ = 5.0 ± 0.1 Hz, major isomer) for the S76 N-H···Nδ1 H79 CtrHb-CN interaction, in agreement with the directly observed splitting (Supporting Information Figure S7.2A,B). The S76-H79 H-bond of the minor heme isomer yielded a similar coupling constant ($^{2h}J_{NN}$ = 5.2 +/- 0.2 Hz, Supporting Information Figure S7.3B). Decreasing the temperature to 283 K shows that the S76 N-H···Nδ1 H79 H-bond is relatively insensitive to temperature changes ($^{2h}J_{NN}$ = 5.3 ± 0.2 Hz, major isomer, Supporting

Information Figure S7.3C). Overall, the magnitude of the CtrHb-CN $^{2h}J_{NN}$ couplings is within the range expected from the few previously reported instances of protein N-H···N H-bonds (2–11 Hz)[21,22,27] and provides a benchmark for comparison with the interactions in THB1, GlbN, and cytochrome b_5.

N-H···N H-bonding in cyanomet THB1 and cyanomet *Synechocystis* GlbN

We collected HSQC, LR HMQC, HNN-COSY and $^{2h}J_{NN}$ modulation data to assign and characterize the N-H···N H-bonds within THB1-CN (N87-H90) and *Synechocystis* GlbN-CN (N80-H83). These proteins yielded HNN-COSY signals (Figure 7.4) and $^{2h}J_{NN}$ values (Supporting Information Figure S7.3I,J,K,L) similar (4.6–5.0 Hz) to those of CtrHb-CN, supporting that sequence context and the identity of the amide donor (Ser in CtrHb, Asn in GlbN and THB1) has undetectable effects on the properties of the helix capping H-bond. Cyanomet GlbN with heme PTM (GlbN-A-CN) yielded a $^{2h}J_{NN}$ value (4.8 Hz, Supporting Information Figure S7.3J) similar to that in its unmodified form, an indication that the native His117-2-Cα heme crosslink has little influence on the N80 N-H···N H83 H-bond.

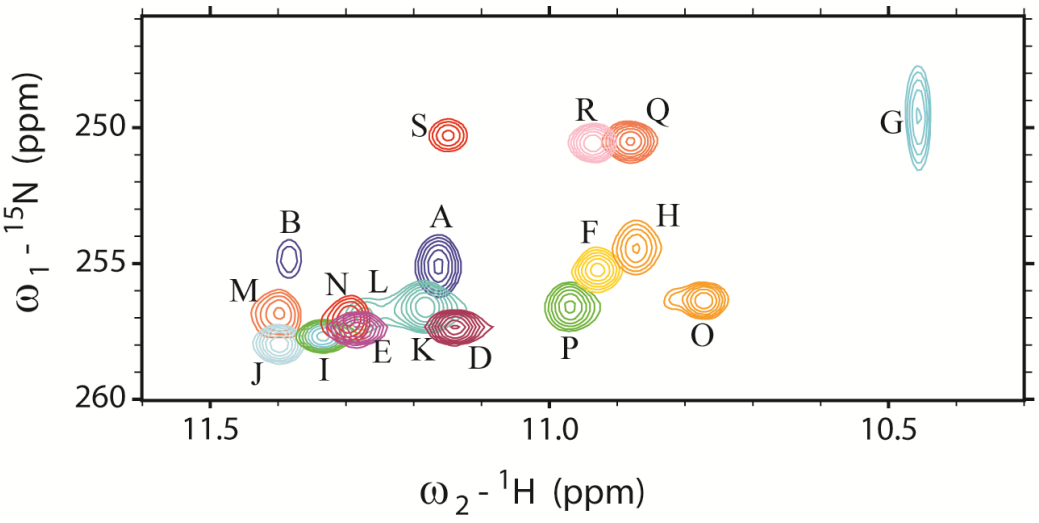


Figure 7.4. Overlay of HNN-COSY spectra collected on hemoglobin and cytochrome b_5 proteins. Each cross-peak in the spectrum corresponds to the amide ^{1}H and His ^{15}Nδ1 of an N-H···Nδ1 H-bond, correlated by the $^{2h}J_{NN}$ hydrogen bond scalar coupling. The peak labels are defined in the first column of Table 7.1.

N-H···N H-bonding in *bis*-histidine *Synechocystis* and *Synechococcus* GlbNs

To address how cyanide binding to the distal heme site alters the N-H···N N-cap, we prepared *Synechococcus* GlbN and *Synechocystis* GlbN with and without PTM in their ferric *bis*-histidine resting states. All four species yielded observable HNN-COSY cross-peaks (Figure 7.4) and $^{2h}J_{NN}$ values of 4.7–4.9 Hz (Supporting Information Figure S7.3D,E,O,P). These results indicate that the large conformational change that occurs upon distal CN binding (Figure 7.1A) does not appreciably affect the N-capping N-H···N H-bonds within the wild-type trHbs, regardless of PTM status. Thus, as may be expected for a highly conserved N-H···N H-bond, the wild-type proteins from *C. eugametos*, *C. reinhardtii*, *Synechococcus* and *Synechocystis* all display highly similar $^{2h}J_{NN}$

values in an indication of constant helix capping H-bond geometry and strength.

Influence of engineered His-heme crosslinking on the helix capping N-H···N H-bond

We next investigated the effects of an engineered histidine–heme crosslink on the helix capping N-H···N H-bonds. As discussed above, a nonnative PTM can be produced in *Synechocystis* GlbN to yield a single (L79H/H117A GlbN-B, H79 Nε-4-Cα linkage) or doubly crosslinked (L79H GlbN-AB, H117 Nε-2-Cα and H79 Nε-4-Cα linkages) protein (Supporting Information Figure S7.1C,D).[41] Additionally, the analogous nonnative His-heme covalent modification can be implanted into *C.eugametos* CtrHb (L75H CtrHb-B, H75 Nε-4-Cα linkage).[28] L79H/H117A and L79H *Synechocystis* GlbNs both exhibit endogenous *bis*-histidine coordination as in the wild-type. ^{1}H-^{15}N HSQC and ^{1}H-^{15}N LR HMQC data demonstrate systematic upfield shifts for N80 amide ^{1}H and H83 ^{15}Nδ1 acceptor nuclei relative to the wild-type reference. As shown in Figure 7.4, the direct correlation of N80 ^{1}H and H83 ^{15}Nδ1 signals demonstrates that the helix capping N-H···N H-bonds are maintained in GlbN variants with nonnative crosslink. However, quantitation of the $^{2h}J_{NN}$ couplings (4.5 Hz in L79H/H117A, 4.3 Hz in L79H, Supporting Information Figure S7.3F,G) suggests that relative to the *bis*-histidine wild-type, lengthening of the N80 N-H···Nδ1 H83 H-bond occurs upon formation of the engineered crosslink.

HSQC and LR HMQC spectra reveal that the N80 ^{1}H and H83 ^{15}Nδ1 resonances of L79H/H117A GlbN-B-CN and L79H GlbN-AB-CN have large downfield shifts relative to their respective *bis*-histidine forms. Subsequent

HNN-COSY experiments demonstrate the intact helix capping H-bonds (Figure 7.4). However, a ~20% increase in the measured $^{2h}J_{NN}$ coupling constants (Supporting Information Figure S7.3M,N) shows that upon distal cyanide binding, appreciable N80 N-H···Nδ1 H83 H-bond relaxation occurs in both L79H/H117A ($^{2h}J_{NN}$ = 5.4 Hz) and L79H GlbNs ($^{2h}J_{NN}$ = 5.0 Hz) relative to their *bis*-histidine complexes (4.5 and 4.3 Hz respectively). Thus, formation of the engineered H79 Nε-4-Cα linkage couples distal cyanide binding to the proximal (N-H···N helix cap) side of the heme pocket in a manner not observed in the wild-type.

Based on the relative $^{2h}J_{NN}$ values of N-H···N H-bonds in cyanomet wild-type GlbNs (4.8–5.0 Hz) and their non-natively crosslinked analogues (5.0–5.4 Hz), the S76 N-H···N H79 helix cap in cyanomet L75H CtrHb-B may be expected to show a similar or slightly enhanced $^{2h}J_{NN}$ value relative to wild-type CtrHb-CN. However, subsequent HNN-COSY (Figure 7.4) and $^{2h}J_{NN}$ modulation experiments (Supporting Information Figure S7.3G) demonstrate that the L75H CtrHb-B S76 N-H···Nδ1 H79 H-bond, although present, is considerably weakened ($^{2h}J_{NN}$ = 4.0 Hz) relative to its wild-type reference ($^{2h}J_{NN}$ = 5.0 Hz, see Table 7.1). These results suggest that upon forming the engineered H75 Nε-4-Cα crosslink in CtrHb-CN, detectable strain in the adjacent helix capping H-bond occurs. The ~20 % *decrease* in $^{2h}J_{NN}$ observed for the L75H CtrHb-B-CN variant relative to wild-type CtrHb-CN (5.0 Hz) is in the opposite direction expected by comparison with wild-type GlbN-(A)-CN (4.8–5.0 Hz) and L79H/H117A GlbN-B-CN (*increases* to 5.4 Hz). Thus, the engineered His-heme crosslink clearly affects the helix capping N-H···N H-bond differently (*strain* in cyanomet CtrHb, *relaxation* in cyanomet GlbN) depending on its structural context.

N-H···N H-bonding in the HPX motif of cytochrome b_5

The *bis*-histidine heme protein cytochrome b_5 contains a backbone N-H···Nδ1 histidine H-bond which caps the short 3_{10} helix near its C-terminus.[42] This motif involves H80 and D82 as H-bonding partners and is remote from the heme binding site.[43] We prepared the ferric, ferrous, and apo states of cytochrome b_5 to detect possible perturbations in H-bonding geometry upon iron redox changes and heme binding. As in the hemoglobins, the N-H donor proton resonates downfield from its random shift (ferric form shown in Supporting Information Figure S7.4A). Also in agreement with previous work, histidine-selective LR HMQC spectra confirm that H80 is neutral and that its Nδ1 is deprotonated (Supporting Information Figure S7.4B). HNN-COSY correlation spectra recorded on the ferric protein (Supporting Information Figure S7.4C) and subsequently the ferrous and apo forms (Figure 7.4), demonstrates the presence of D82 N-H···N H80 H-bond (Supporting Information Figure S7.4D) in all three forms of cytochrome b_5. Measurement of the $^{2h}J_{NN}$ coupling constants returned values of 5.0, 5.0, and 5.3 Hz for ferric (313 K), ferrous (313 K), and apo (298 K) forms (Supporting Information Figure S7.3Q-S). These results indicate that the HPD N-H···N H-bond of cytochrome b_5 depends little on the iron redox state or the presence of the heme. Additionally, the cytochrome b_5 H-bond appears to be of similar strength to those detected within the hemoglobin T/S/NxxH motifs. Thus, in the set of native heme proteins, it appears that the local secondary structure has minimal influence on amide N-H···Nδ histidine bonding geometry.

Discussion

$^{2h}J_{NN}$ and chemical shift correlations for amide N-H···Nδ His H-bonds

A remarkable study of N-H···N H-bonds in DNA by Dingley *et al.* demonstrates a positive correlation between the extent of imino donor ^{1}H deshielding and the magnitude of $^{2h}J_{NN}$, independent of base pair type.[19] Similar observations have been made between the amide ^{1}H shift and $^{3h}J_{NC'}$ within N-H···O=C type protein H-bonds.[7] The H-bonds detected in the current work provide an opportunity to analyze trends between $^{2h}J_{NN}$ and the chemical shifts of nuclei comprising N-H···N type protein H-bonds and facilitate comparison with the previous work on nucleic acids.

As shown in Figure 7.5A, the amide ^{1}H chemical shifts (ppm) and $^{2h}J_{NN}$ couplings (Hz) within heme protein N-H···N H-bonds are also positively correlated (i.e. greater deshielding, greater coupling). The slope of the correlation line (~0.9 Hz/ppm) is smaller than that observed for Hoogsteen and Watson-Crick base pairs in triplex DNA (~1.3 Hz/ppm)[19] but overall in good agreement. Owing to their proximity of the paramagnetic heme, the hemoglobin Thr, Ser, and Asn amide protons are expected to have small (~ –0.2 ppm) pseudocontact shifts. This effect, along with the different amino acid sequence contexts and types of helix capping N-H···N H-bonds (T/S/NxxH in TrHbs and HPD in cytochrome b_5), may account for some of the variability in the protein ^{1}H chemical shifts.

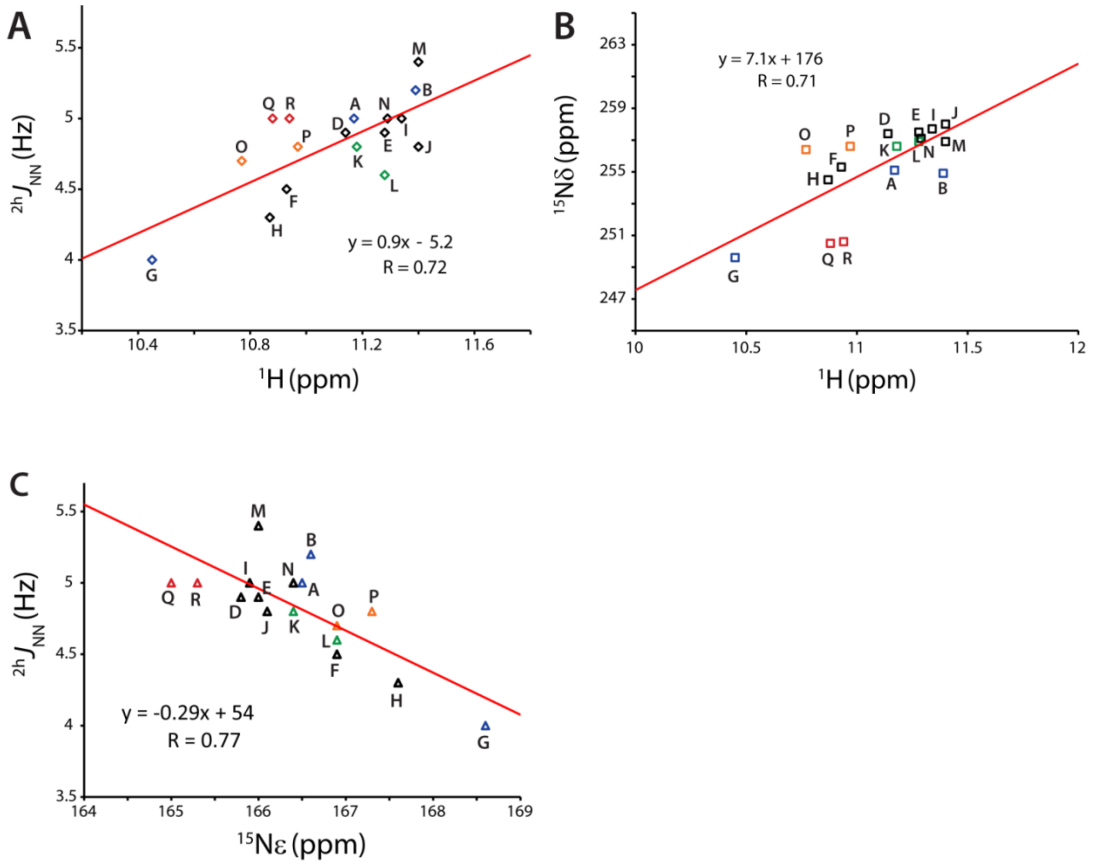


Figure 7.5. A. ^{1}H vs $^{2h}J_{NN}$: Plot of the amide ^{1}H chemical shift (ppm) and $^{2h}J_{NN}$ (Hz) obtained from the hemoglobin and cytochrome b_5 helix capping data (313 K) shown in Table 7.1. Linear regression indicates a positive correlation (slope ~ 0.9 Hz/ppm, R = 0.72) similar to that previously reported in nucleic acid N-H···N H-bonds (see Figure 7.6). Data points are labeled according to the first column of Table 7.1 and colored as follows: CtrHbs (blue), THB1 (green), *Synechococcus* GlbNs (orange), *Synechocystis* GlbNs (black), and cytochrome b_5 (red). B. ^{1}H vs ^{15}Nδ1: Plot of the amide ^{1}H donor and histidine acceptor ^{15}Nδ1 chemical shifts (ppm) observed in hemoglobin and cytochrome b_5 helix capping H-bonds (313 K). Interestingly, the ^{1}H and ^{15}Nδ1 chemical shifts are positively correlated (slope ~ 7.1 ^{15}N ppm/^{1}H ppm, R = 0.71), an indication that both nuclei become more deshielded within stronger N-H···N H-bonds. Data point labeling and colors are as in 7.5A. C. ^{15}Nε2 vs $^{2h}J_{NN}$: Correlation plot of the histidine ^{15}Nε2 chemical shift (ppm) and $^{2h}J_{NN}$ (Hz) obtained from hemoglobin and cytochrome b_5 data (313 K) shown in Table 7.1. The data show a negative linear correlation (slope ~ -0.3 Hz/ppm, R ~0.77) between the His ^{15}Nε2 chemical shift and $^{2h}J_{NN}$ coupling. Data point labeling and colors are as in 7.5A and 7.5B.

The presently-determined heme protein $^{2h}J_{NN}$ couplings and ^{1}H chemical shifts were combined with data collected on ankyrin repeat proteins[27] and the Watson-Crick/Hoogsteen DNA N-H···N H-bonds.[19] Remarkably, the combined data fall on a single line (slope ~1.1 Hz/ppm) spanning 9 to 15 ppm (^{1}H) and $^{2h}J_{NN}$ of 2 to 11 Hz (Figure 7.6). Despite this trend, the different absolute magnitude of protein and nucleic acid $^{2h}J_{NN}$ values supports that in solution, a typical N-H···N H-bond in the DNA triplex is shorter (and possibly stronger) than the helix capping N-H···N H-bond in proteins. It should be noted, however, that a protein N-H···N H-bond *can* yield $^{2h}J_{NN}$ values ~ 10 Hz, as exemplified by the His24-His119 side-chain-side-chain interaction observed in apomyoglobin.[21]

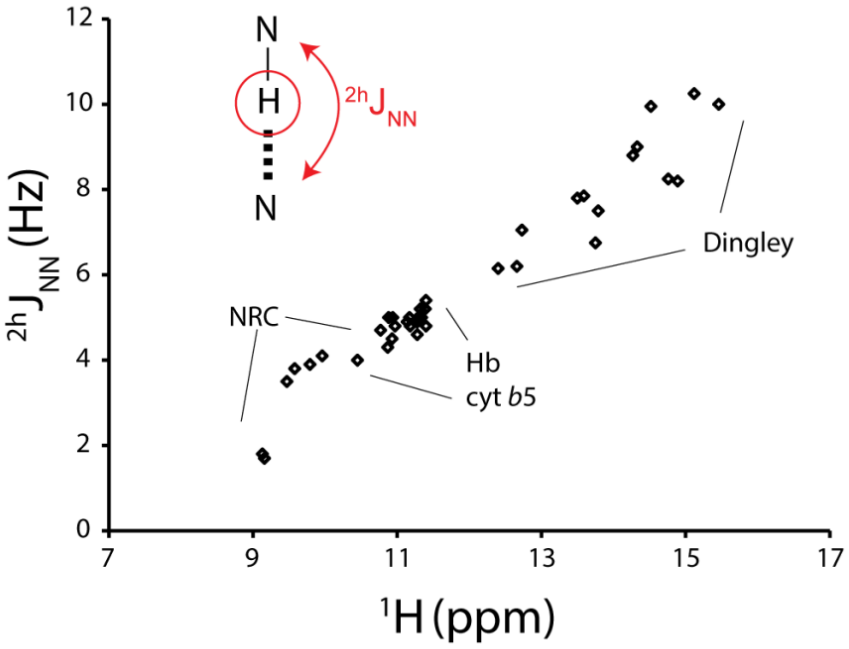


Figure 7.6. The amide N-H···Nδ1 His H-bonds in proteins (hemoglobins, cytochrome b_5, and ankyrin consensus repeat proteins (NRCs)) show a strong positive correlation between donor ^{1}H chemical shift and $^{2h}J_{NN}$. This trend is strikingly similar to that observed previously for imino N-H···N Watson-Crick and Hoogsteen base pairs in triplex DNA (data from Dingley *et al.*), despite the fact that the intrinsic ^{1}H shifts of protein amide and nucleic acid imino protons are different. A linear regression of the combined data returns a slope of ~1.1 Hz/ppm. Overall, the relative magnitudes of the $^{2h}J_{NN}$ values suggest that nucleic acid N-H···N H-bonds may be shorter (~2.8-2.9 Å) than those in proteins (~3.0-3.1 Å). I hypothesize that nucleic acid N-H···N H-bonds between base pairs are typically shorter, and likely stronger, due to their tendency to form a highly cooperative networks.

Direct comparison of protein amide and nucleic acid imino proton chemical shifts is hindered owing to several contributions unrelated to the presence or strength of an H-bond. For example, imino and amide protons have distinct intrinsic chemical shifts. Base stacking and the aromatic ring current effect on imino protons further complicates the comparison and adds to the noise. Ring

current contributions from the histidine and the porphyrin (~0.3–0.5 ppm) are also at work in the T/S/NxxH and HPD heme protein motifs and influence the amide proton chemical shift. Correction for these factors may improve the trend shown in Figure 7.6. Nevertheless, these combined protein and DNA data[19] predict a ~1.1 Hz/ppm slope between $^{2h}J_{NN}$ and ^{1}H chemical shift. Linear extrapolation of the plot shown in Figure 7.6 predicts a ^{1}H chemical shift of ~6.9 ppm for a non-H-bonded amide proton ($^{2h}J_{NN}$ = 0). This result is in excellent agreement with that observed in the plot of ^{1}H chemical shift against $^{3h}J_{NC'}$ (^{1}H ~6.8 ppm for $^{3h}J_{NC'}$ = 0).[7] Importantly, these combined protein and nucleic acid $^{2h}J_{NN}$ data strongly support the classical (but indirect) ^{1}H deshielding metric for the presence and length of H-bonds.

Additional information can be drawn from the chemical shifts of N-H···Nδ1 H-bonding nuclei. Classical ^{14}N NMR work on azoles[56] and extension to ^{15}N NMR of free histidine in solution[57], model imidazoles, and histidine within the catalytic triad of α-lytic protease[58], define the ring nitrogens as one of two types: the pyrrole α-type, which has an upfield ^{15}N shift (167.5 ppm, 176.5 ppm in cationic imidazole[45]), and the deprotonated pyridine β-type, which resonates downfield (^{15}N = 249.5, recently revised to 261.5–266.5 ppm[45,59]). Under fast-exchange on the chemical shift timescale, the ^{15}N chemical shifts of neutral histidine reflect the population weighted-average of Nε2-H and Nδ1-H tautomers.[57] If both tautomers are equally populated, the histidine side-chain ^{15}N signals will resonate at the average chemical shift of α and β type nitrogens (208.5 ppm). Hydrogen bond formation has several, competing effects on the ^{15}N shifts. 1. The tautomeric equilibrium of the imidazole side chain will be

biased from its intrinsic value (~4:1 Nε-H:Nδ-H) if an H-bonding interaction can occur with one nitrogen but not the other. This effect will lead to *increased* separation of the ^{15}N shifts.[57] The N-H···Nδ1 H-bonds reported here require a deprotonated Nδ1, therefore resulting in strong bias for the Nε2-H tautomer. 2. H-bonding also affects directly the imidazole ^{15}N shifts: a pure α-type nitrogen (proton donor) engaged in H-bonding becomes deshielded (shifts ~10 ppm downfield), whereas a pure β-type nitrogen (proton acceptor) becomes shielded (shifts ~10 ppm upfield).[60] Thus, a neutral histidine with both α and β type nitrogens engaged in H-bonds will lead to their ^{15}N chemical shifts moving closer to one another relative to the ^{15}N shifts of an H-bonded histidine with a single H-bond.[60] Figure 7.5B presents a correlation plot of the amide donor ^{1}H and histidine acceptor (β-type nitrogen) ^{15}Nδ1 chemical shifts for the N-H···N H-bonds reported herein. The heme protein N-H···Nδ1 H-bond data suggest that the extents of ^{1}H deshielding and ^{15}Nδ1 deshielding are *positively* correlated. I hypothesize that the first effect (histidine tautomeric bias) described above primarily determines the heme protein acceptor ^{15}Nδ1 chemical shift behavior. For example, a relatively optimized N-H···Nδ1 H-bond (e.g. CtrHb-CN S76-H79, $^{2h}J_{NN}$ ~ 5 Hz, ^{1}H = 11.17 ppm) leads to large depression of the imidazole ring pK_a (< 3.5, data not shown) and fully promotes the Nε2-H tautomeric form (at neutral pH) resulting in a downfield-shifted Nδ1 (255.1 ppm). In contrast, in a weaker N-H···Nδ1 H-bond with smaller $^{2h}J_{NN}$ coupling and ^{1}H shift (e.g. L75H CtrHb-B-CN S76-H79, $^{2h}J_{NN}$ ~ 4 Hz, ^{1}H = 10.45 ppm), the acceptor Nδ1 may increasingly sample a non H-bonded state. In the process, the Nδ1 will occasionally become formally protonated to form the neutral Nδ1-H tautomer

(leading to upfield-shift of $^{15}N\delta1$ = 249.6 ppm). Additionally, for very weak N-H···Nδ1 H-bonds, the histidine pK_a will be only moderately depressed relative to the standard value of 6.5. Such a histidine will routinely sample its charged form near neutral pH, also causing an apparent upfield Nδ1 shift. Thus, I argue that a moderately strong amide N-H···Nδ1 H-bond will promote the sole population of the neutral His Nε2-H tautomeric form, and result in a deprotonated Nδ1 nucleus with downfield chemical shift. The same argument can be extended to several well-studied proteins. Triose phosphate isomerase-His95, myoglobin-His24, xylanase-His149, and subtilisin Novo-His67 are all examples in which a known H-bonded histidine ^{15}N acceptor has chemical shift near or *downfield* of the classical limiting value.[61-64]

Histidine tautomeric bias caused by the N-H···Nδ1 H-bond can be applied to rationalize the correlation between the histidine $^{15}N\varepsilon2$, which is not directly involved in heme protein N-H···Nδ1 interaction, and the $^{2h}J_{NN}$ couplings within the heme protein set (Figure 7.5C). Specifically, stronger H-bonds with larger $^{2h}J_{NN}$ values promote the neutral His tautomer with deprotonated Nδ1. Accordingly, the histidine Nε2 adopts pure pyrrolic α-type character and shifts far-upfield (e.g. CtrHb-CN S76-H79, $^{2h}J_{NN}$ = 5.0 Hz, $^{15}N\varepsilon$ = 166.5 ppm). Longer N-H···Nδ1 H-bonds (e.g. L75H CtrHb-B-CN S76-H79, $^{2h}J_{NN}$ = 4.0 Hz) will lead to increased population of the charged or Nδ1-H histidine tautomer (with deprotonated Nε2), and in both instances, lead to a downfield shift of the Nε2 nucleus ($^{15}N\varepsilon2$ = 168.6 ppm).

Conclusion

In this work, we demonstrate that helix capping N-H···N H-bonds can be routinely detected in proteins by hydrogen bond scalar coupling experiments. Direct assignment of H-bonding nuclei was achieved by tailoring HNN-COSY experiments for protein amide N-H and His Nδ1 nuclei. Quantitation of the $^{2h}J_{NN}$ coupling constant provided a convenient metric for the length, and potentially strength, of H-bonds. The conserved helix capping hydrogen bond in the *wild-type* hemoglobin T/S/NxxH motifs yielded remarkably similar $^{2h}J_{NN}$ couplings, despite the proteins having less than 50% sequence identity, distinct iron coordination modes, and the presence or absence of the native H117-2-Cα heme crosslink (wild-type avg. ~ 4.9 Hz). Likewise, the N-H···N H-bond detected in the 3_{10} capping HPD motif of cytochrome b_5 also yielded $^{2h}J_{NN}$ values of 5 Hz.

In contrast, an engineered His-heme crosslink adjacent to the helix capping N-H···N H-bond had distinct effects depending on the background protein and iron ligation state. In *bis*-histidine GlbNs, formation of the nonnative H79-4-Cα heme crosslink lead to detectable *strain* in the helix capping N-H···N H-bond ($^{2h}J_{NN}$ = 4.3-4.5 Hz) relative to the wild-type (4.8-5.0 Hz). However, cyanide binding to the L79H GlbN variants causes significant *relaxation* of the N-H···N H-bond as evidenced by the ~20% increase in $^{2h}J_{NN}$ ($^{2h}J_{NN}$ = 5-5.4 Hz). Thus, unlike wild-type GlbNs, formation of the nonnative linkage couples distal ligand binding to the proximal side of the heme pocket. In contrast, the H75-4-Cα heme crosslink engineered in cyanide bound CtrHb-B lead to N-H···N H-bond deformation ($^{2h}J_{NN}$ = 4 Hz) relative to its wild-type reference ($^{2h}J_{NN}$ = 5 Hz).

The heme protein $^{2h}J_{NN}$ couplings reported in this work are linearly correlated with donor ^{1}H chemical shift, in a manner reminiscent of N-H···N H-bonds in Watson-Crick and Hoogsteen base pairs.[19] However, the absolute magnitude of the protein couplings tends to be smaller than those in nucleic acids, and hints that in solution, most DNA/RNA N-H···N H-bond lengths are shorter than those in the studied proteins. In sum, we have shown that strain or optimization of N-H···N H-bonds is observable using $^{2h}J_{NN}$ couplings. Future work will establish to what extent the $^{2h}J_{NN}$ couplings are related to local H-bond energetics.

Table 7.1. $^{2h}J_{NN}$ values, amide donor ^{1}H, histidine $^{15}N\delta1$ and $^{15}N\varepsilon2$ chemical shifts observed in heme protein N-H···Nδ1 helix capping H-bonds (313K)

Fig4	Protein	Fe	H-bond	$^{2h}J_{NN}$ (Hz)	^{1}H (ppm)	$^{15}N\varepsilon$ (ppm)	$^{15}N\delta$ (ppm)
	CtrHb, major isomer[a]	CN	S76 NH - Nδ H79	5.3 +/- 0.2	11.41	nd[c]	254.9
B	CtrHb, minor isomer	CN	S76 NH - Nδ H79	5.2 +/- 0.2	11.39	166.6	254.9
A	CtrHb, major isomer	CN	S76 NH - Nδ H79	5.0 +/- 0.1	11.17	166.5	255.1
G	L75H CtrHb-B	CN	S76 NH - Nδ H79	4.0 +/- 0.3	10.45	168.6	249.6
K	THB1, major isomer	CN	N87 NH - Nδ H90	4.8 +/- 0.1	11.18	166.4	256.6
L	THB1, minor isomer	CN	N87 NH - Nδ H90	4.6 +/- 0.3	11.28	166.9	256.9
I	*Syn*6803 GlbN	CN	N80 NH - Nδ H83	5.0 +/- 0.2	11.34	165.9	257.7
J	*Syn*6803 GlbN-A	CN	N80 NH - Nδ H83	4.8 +/- 0.2	11.40	166.1	258.0
D	*Syn*6803 GlbN	His	N80 NH - Nδ H83	4.9 +/- 0.2	11.14	165.8	257.4
E	*Syn*6803 GlbN-A	His	N80 NH - Nδ H83	4.9 +/- 0.2	11.28	166.0	257.5
O	*Syn*7002 GlbN	His	T80 NH - Nδ H83	4.7 +/- 0.2	10.77	166.9	256.4
P	*Syn*7002 GlbN-A	His	T80 NH - Nδ H83	4.8 +/- 0.2	10.97	167.3	256.6
F	L79H/H117A *Syn*6803 GlbN-B	His	N80 NH - Nδ H83	4.5 +/- 0.2	10.93	166.9	255.3
H	L79H *Syn*6803 GlbN-AB	His	N80 NH - Nδ H83	4.3 +/- 0.2	10.87	167.6	254.5
M	L79H/H117A *Syn*6803 GlbN-B	CN	N80 NH - Nδ H83	5.4 +/- 0.1	11.40	166.0	256.9
N	L79H *Syn*6803 GlbN-AB	CN	N80 NH - Nδ H83	5.0 +/- 0.1	11.29	166.4	257.1
Q	Rat cytochrome ferric b_5	His	D82 NH - Nδ H80	5.0 +/-0.1	10.88	165.0	250.5
R	Rat cytochrome ferrous b_5	His	D82 NH - Nδ H80	5.0 +/-0.1	10.94	165.3	250.6
S	Rat apocytochrome b_5[b]	n/a	D82 NH - Nδ H80	5.3 +/-0.1	11.15	165.0	250.3

[a]Data collected at 283 K, 10 % D_2O, [b]Data collected at 298 K, 10 % D_2O

[c]Not determined

References

1. Dingley, A. J. & Grzesiek, S. Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide 2JNN Couplings. *J. Am. Chem. Soc.* **120,** 8293–8297 (1998).
2. Pervushin, K. *et al.* NMR scalar couplings across Watson–Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc. Natl. Acad. Sci.* **95,** 14147–14151 (1998).
3. Cornilescu, G., Hu, J.-S. & Bax, A. Identification of the Hydrogen Bonding Network in a Protein by Scalar Couplings. *J. Am. Chem. Soc.* **121,** 2949–2950 (1999).
4. Dingley, A. J., Cordier, F. & Grzesiek, S. An introduction to hydrogen bond scalar couplings. *Concepts Magn. Reson.* **13,** 103–127 (2001).
5. Majumdar, A. & Patel, D. J. Identifying hydrogen bond alignments in multistranded DNA architectures by NMR. *Acc. Chem. Res.* **35,** 1–11 (2002).
6. Wang, Y. X. *et al.* Measurement of 3hJNC' connectivities across hydrogen bonds in a 30 kDa protein. *J. Biomol. NMR* **14,** 181–184 (1999).
7. Cordier, F. & Grzesiek, S. Direct Observation of Hydrogen Bonds in Proteins by Interresidue 3hJNC' Scalar Couplings. *J. Am. Chem. Soc.* **121,** 1601–1602 (1999).
8. Cordier, F. & Grzesiek, S. Quantitative comparison of the hydrogen bond network of A-state and native ubiquitin by hydrogen bond scalar couplings. *Biochemistry* **43,** 11295–11301 (2004).
9. Jaravine, V. A., Alexandrescu, A. T. & Grzesiek, S. Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci.* **10,** 943–950 (2001).

10. Liu, A., Hu, W., Majumdar, A., Rosen, M. K. & Patel, D. J. NMR detection of side chain–side chain hydrogen bonding interactions in 13C/15N-labeled proteins. *J. Biomol. NMR* **17,** 305–310 (2000).
11. Cordier, F., Rogowski, M., Grzesiek, S. & Bax, A. Observation of Through-Hydrogen-Bond 2hJHC′ in a Perdeuterated Protein. *J. Magn. Reson.* **140,** 510–512 (1999).
12. Cornilescu, G. *et al.* Correlation between 3hJNC‘ and Hydrogen Bond Length in Proteins. *J. Am. Chem. Soc.* **121,** 6275–6279 (1999).
13. Benedict, H. *et al.* Nuclear Scalar Spin−Spin Couplings and Geometries of Hydrogen Bonds. *J. Am. Chem. Soc.* **122,** 1979–1988 (2000).
14. Grzesiek, S., Cordier, F., Jaravine, V. & Barfield, M. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.* **45,** 275–300 (2004).
15. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. RNA double-helical fragments at atomic resolution: I. The crystal and molecular structure of sodium adenylyl-3′,5′-uridine hexahydrate. *J. Mol. Biol.* **104,** 109–144 (1976).
16. Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. RNA double-helical fragments at atomic resolution: II. The crystal structure of sodium guanylyl-3′,5′-cytidine nonahydrate. *J. Mol. Biol.* **104,** 145–167 (1976).
17. Alkorta, I., Elguero, J. & Denisov, G. S. A review with comprehensive data on experimental indirect scalar NMR spin–spin coupling constants across hydrogen bonds. *Magn. Reson. Chem.* **46,** 599–624 (2008).
18. Dingley, A. J., Peterson, R. D., Grzesiek, S. & Feigon, J. Characterization of the Cation and Temperature Dependence of DNA Quadruplex Hydrogen

Bond Properties Using High-Resolution NMR. *J. Am. Chem. Soc.* **127,** 14466–14472 (2005).

19. Dingley, A. J. *et al.* Internucleotide Scalar Couplings Across Hydrogen Bonds in Watson−Crick and Hoogsteen Base Pairs of a DNA Triplex. *J. Am. Chem. Soc.* **121,** 6019–6027 (1999).
20. Majumdar, A., Kettani, A. & Skripkin, E. Observation and measurement of internucleotide 2JNN coupling constants between 15N nuclei with widely separated chemical shifts. *J. Biomol. NMR* **14,** 67–70 (1999).
21. Hennig, M. & Geierstanger, B. H. Direct Detection of a Histidine−Histidine Side Chain Hydrogen Bond Important for Folding of Apomyoglobin. *J. Am. Chem. Soc.* **121,** 5123–5126 (1999).
22. Eletsky, A. *et al.* Direct NMR observation and DFT calculations of a hydrogen bond at the active site of a 44 kDa enzyme. *J. Biomol. NMR* **24,** 31–39 (2002).
23. Presta, L. G. & Rose, G. D. Helix signals in proteins. *Science* **240,** 1632–1641 (1988).
24. Forood, B., Feliciano, E. J. & Nambiar, K. P. Stabilization of alpha-helical structures in short peptides via end capping. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 838–842 (1993).
25. Aurora, R. & Rose, G. D. Helix capping. *Protein Sci.* **7,** 21–38 (1998).
26. Harper, E. T. & Rose, G. D. Helix stop signals in proteins and peptides: the capping box. *Biochemistry* **32,** 7605–7609 (1993).
27. Preimesberger, M. R. *et al.* Anatomy of an alpha-helix N-cap: Direct NMR detection of unusual bifurcated hydrogen bonding in the TXXH motif.

28. Rice, S. L., Preimesberger, M. R., Johnson, E. A. & Lecomte, J. T. J. Introduction of a covalent histidine–heme linkage in a hemoglobin: A promising tool for heme protein engineering (manuscript submitted to JIB).
29. Johnson, E. A. *et al.* 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 (2014).
30. Vu, B. C., Vuletich, D. A., Kuriakose, S. A., Falzone, C. J. & Lecomte, J. T. J. Characterization of the heme-histidine cross-link in cyanobacterial hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002. *J. Biol. Inorg. Chem.* **9,** 183–194 (2004).
31. Hoy, J. A., Kundu, S., Trent, J. T., Ramaswamy, S. & Hargrove, M. S. The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. *J. Biol. Chem.* **279,** 16535–16542 (2004).
32. Falzone, C. J., Christie Vu, B., Scott, N. L. & Lecomte, J. T. J. The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state. *J. Mol. Biol.* **324,** 1015–1029 (2002).
33. Scott, N. L. *et al.* Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry* **49,** 7000–7011 (2010).
34. Wenke, B. B., Lecomte, J. T. J., Héroux, A. & Schlessman, J. L. 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 (2014).

35. Pesce, A. *et al.* A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO J.* **19,** 2424–2434 (2000).

36. Vu, B. C., Nothnagel, H. J., Vuletich, D. A., Falzone, C. J. & Lecomte, J. T. J. Cyanide binding to hexacoordinate cyanobacterial hemoglobins: hydrogen-bonding network and heme pocket rearrangement in ferric H117A *Synechocystis* hemoglobin. *Biochemistry* **43,** 12622–12633 (2004).

37. Trent, J. T., Kundu, S., Hoy, J. A. & Hargrove, M. S. Crystallographic analysis of *synechocystis* cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. *J. Mol. Biol.* **341,** 1097–1108 (2004).

38. Vu, B. C., Jones, A. D. & Lecomte, J. T. J. Novel histidine-heme covalent linkage in a hemoglobin. *J. Am. Chem. Soc.* **124,** 8544–8545 (2002).

39. Nothnagel, H. J. *et al.* 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 (2011).

40. Preimesberger, M. R., Pond, M. P., Majumdar, A. & Lecomte, J. T. J. 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 (2012).

41. Preimesberger, M. R., Wenke, B. B., Gilevicius, L., Pond, M. P. & Lecomte, J. T. J. Facile heme vinyl posttranslational modification in a hemoglobin. *Biochemistry* **52,** 3478–3488 (2013).

42. Lecomte, J. T. J. & Moore, C. D. Helix formation in apocytochrome b5: the role of a neutral histidine at the N-cap position. *J. Am. Chem. Soc.* **113,** 9663–9665 (1991).
43. Falzone, C. J., Mayer, M. R., Whiteman, E. L., Moore, C. D. & Lecomte, J. T. Design challenges for hemoproteins: the solution structure of apocytochrome b5. *Biochemistry* **35,** 6519–6526 (1996).
44. Lecomte, J. T., Scott, N. L., Vu, B. C. & Falzone, C. J. Binding of ferric heme by the recombinant globin from the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochemistry* **40,** 6541–6552 (2001).
45. Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from Escherichia coli, using two-dimensional heteronuclear NMR techniques. *Protein Sci.* **2,** 543–558 (1993).
46. Wishart, D. S. *et al.* 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6,** 135–140 (1995).
47. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 (1995).
48. Goddard, T. D. & Kneller, D. G. *SPARKY 3*.
49. Bax, A. *et al.* Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol.* **239,** 79–105 (1994).
50. Falzone, C. J. & Lecomte, J. T. J. Assignment of the 1H, 13C, and 15N signals of Synechocystis sp. PCC 6803 methemoglobin. *J. Biomol. NMR* **23,** 71–72 (2002).
51. Pond, M. P., Vuletich, D. A., Falzone, C. J., Majumdar, A. & Lecomte, J. T. J. (1)H, (15)N, and (13)C resonance assignments of the 2/2 hemoglobin from

the cyanobacterium *Synechococcus* sp. PCC 7002 in the ferric bis-histidine state. *Biomol. NMR Assign.* **3,** 211–214 (2009).

52. Pond, M. P., Majumdar, A. & Lecomte, J. T. J. Influence of heme post-translational modification and distal ligation on the backbone dynamics of a monomeric hemoglobin. *Biochemistry* **51,** 5733–5747 (2012).
53. Preimesberger, M. R., Pond, M. P., Majumdar, A. & Lecomte, J. T. J. 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 (2012).
54. Banci, L., Bertini, I., Rosato, A. & Scacchieri, S. Solution structure of oxidized microsomal rabbit cytochrome b5. Factors determining the heterogeneous binding of the heme. *Eur. J. Biochem. FEBS* **267,** 755–766 (2000).
55. Majumdar, A., Kettani, A., Skripkin, E. & Patel, D. J. Observation of internucleotide NH...N hydrogen bonds in the absence of directly detectable protons. *J. Biomol. NMR* **15,** 207–211 (1999).
56. Witanowski, M., Stefaniak, L., Januszewski, H., Grabowski, Z. & Webb, G. A. Nitrogen-14 nuclear magnetic resonance of azoles and their benzo-derivatives. *Tetrahedron* **28,** 637–653 (1972).
57. Blomberg, F., Maurer, W. & Rüterjans, H. Nuclear magnetic resonance investigation of 15N-labeled histidine in aqueous solution. *J. Am. Chem. Soc.* **99,** 8149–8159 (1977).
58. Bachovchin, W. W. & Roberts, J. D. Nitrogen-15 nuclear magnetic resonance spectroscopy. The state of histidine in the catalytic triad of alpha-

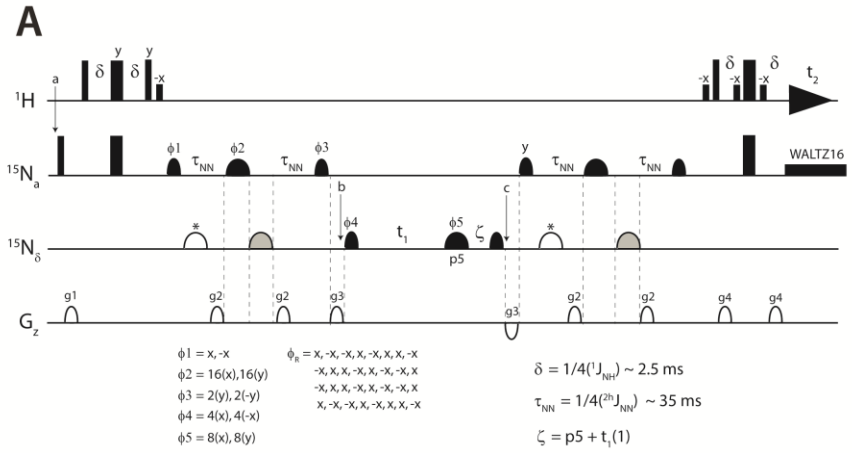
lytic protease. Implications for the charge-relay mechanism of peptide-bond cleavage by serine proteases. *J. Am. Chem. Soc.* **100,** 8041–8047 (1978).

59. Vila, J. A. Limiting values of the 15N chemical shift of the imidazole ring of histidine at high pH. *J. Phys. Chem. B* **116,** 6665–6669 (2012).
60. Bachovchin, W. W. Nitrogen-15 NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: evidence for a moving histidine mechanism. *Biochemistry* **25,** 7751–7759 (1986).
61. Lodi, P. J. & Knowles, J. R. Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. *Biochemistry* **30,** 6948–6956 (1991).
62. Bhattacharya, S., Sukits, S. F., MacLaughlin, K. L. & Lecomte, J. T. The tautomeric state of histidines in myoglobin. *Biophys. J.* **73,** 3230–3240 (1997).
63. Plesniak, L. A., Connelly, G. P., Wakarchuk, W. W. & McIntosh, L. P. Characterization of a buried neutral histidine residue in Bacillus circulans xylanase: NMR assignments, pH titration, and hydrogen exchange. *Protein Sci.* **5,** 2319–2328 (1996).
64. Day, R. M. *et al.* Tautomerism, acid-base equilibria, and H-bonding of the six histidines in subtilisin BPN' by NMR. *Protein Sci.* **12,** 794–810 (2003).

Supporting Information for

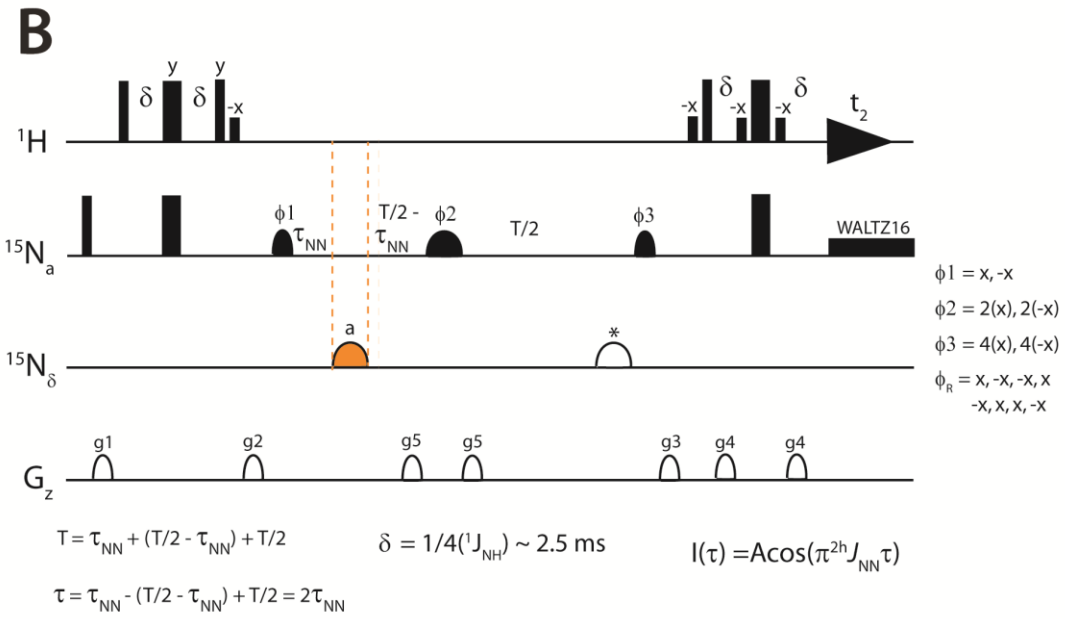
The NMR hydrogen bond scalar coupling $^{2h}J_{NN}$ as a sensitive probe for helix capping N-H···N hydrogen bonds in proteins

Supporting Information Methods



Methods S7.1. HNN-COSY pulse sequence. A. Narrow and wide black bars correspond to π/2 and π pulses, respectively. The RF power levels of these hard rectangular pulses were ~25 kHz (^{1}H) and 6.6 kHz (^{15}N). Pulses applied to amide ^{15}N (120 ppm) and histidine ^{15}Nδ1 (250 ppm) nuclei are denoted by $^{15}N_a$ and $^{15}N_\delta$, respectively. The ^{1}H carrier was positioned on the water signal, set at 4.76 ppm. The ^{15}N carrier was positioned at 120 ppm at point a and switched to ~250 ppm at point b. Water magnetization was maintained at equilibrium using low power 1.5 ms rectangular (flip-back) π/2 ^{1}H pulses. Proton excitation is followed by selective $^1J_{NH}$ INEPT transfer using ~2.5 ms δ delays and a 1-ms Q5 $^{15}N_a$ shaped pulse (narrow black bell) to generate transverse ^{15}N magnetization (HzNy). During the τ_{NN} period, the two-bond trans hydrogen bond scalar coupling, $^{2h}J_{NN}$, is preserved using 2 ms ReBurp $^{15}N_a$ (wide black bell) and 1.5 ms G3 $^{15}N_\delta$ (wide gray bell) shaped pulses, the latter applied downfield at ~250 ppm. The G3 $^{15}N_\delta$ pulse marked with an asterisk (wide white bell) is required for Bloch-Siegert shift compensation. At point b, the ^{15}N carrier position is switched to ~245 ppm, and the 1 ms Q5 $^{15}N_\delta$ pulse produces transverse ^{15}Nδ magnetization, which is subsequently frequency-labeled during the t1 period. The ζ delay is for first order phase correction in t1. After ^{15}Nδ chemical shift encoding, the ^{15}N carrier is switched back to 118 ppm (point c) and the coherence transfer steps are reversed (HzNzNδy → HzNxNδz →HzNy → HxNz → Hy). The ^{15}N shaped pulses applied during the reverse $^{2h}J_{NN}$ and $^1J_{NH}$ INEPT periods are identical to those used for the forward transfers. ^{15}N WALTZ16 decoupling is applied within the amide region during direct ^{1}H acquisition. Ultimately, the observed signal correlates the amide ^{1}H with histidine ^{15}Nδ1 within an N-H···Nδ1 H-bond. Delay times and phase cycling program are described in the figure. Representative acquisition parameters for a 1 mM protein sample: number of transients = 320, steady state scans = 32,

^{1}H spectral width = 12 kHz, ^{1}H acquisition time = 75 ms, ^{15}N spectral width = 32 ppm, ^{15}N acquisition time = 15 ms, T_1 relaxation delay = 1.1 s.



Methods S7.1. B. Quantitative $^{2h}J_{NN}$ spin-echo difference HSQC: Hard rectangular ^{1}H, ^{15}N, and ^{15}N decoupling and shaped pulses are denoted as above. ^{1}H and ^{15}N carrier frequencies were 4.76 and ~120 ppm, respectively. After an initial $^{1}J_{NH}$ INEPT transfer period (δ = 2.5 ms), amide ^{15}N magnetization (HzNy) is allowed to decay during the constant-time relaxation period, T (100–160 ms). During the relaxation delay, *J*-modulation of the amide ^{15}N magnetization is achieved by variable timing (τ_{NN}) of an off-resonance (250 ppm) 1.5-ms $^{15}N_{\delta}$ G3 inversion pulse (wide orange bell marked a). The open pulse demarcated with an asterisk is used for Bloch-Siegert shift correction. With this spin-echo difference scheme, an experiment that uses a very short τ period (e.g., 1 ms) leads to nearly complete refocusing of the small $^{2h}J_{NN}$ couplings during the constant-time period T (HzNycos($\pi J\tau$) is near unity). Therefore, maximum amide ^{1}H intensity is observed following the reverse INEPT (HzNy $\rightarrow$ HxNz $\rightarrow$ Hy) period. In contrast, for longer τ_{NN} periods, the $^{2h}J_{NN}$ coupling remains active over increasing fractions of T. This leads to the buildup of unobservable HzNxNδzsin($\pi J\tau$) magnetization, eliminated by a combination of phase cycling and Z-gradients. Thus in the latter case, vanishing or negative HzNycos($\pi J\tau$) magnetization is detected. A $^{2h}J_{NN}$-modulation series comprised of ~12–20 1-D experiments, each with the same constant time T, but with the variable period, τ, systematically incremented in random order. Additionally, the first few experiments were duplicated at the end of the series as a test for data reproducibility and to assess the uncertainty in peak intensities. Delay

definitions, the phase cycling scheme, and $^{2h}J_{NN}$ fitting equation are presented in the figure.

SI Figure List

1. Cartoon depiction of helix capping H-bond "strain" and "relaxation" in GlbNs, as probed by $^{2h}J_{NN}$

2. Ser76 N-H···Nδ1 His79 H-bond in CtrHb, initial observation of $^{1h}J_{HN}$ and $^{2h}J_{NN}$ trans hydrogen bond scalar couplings

3. $^{2h}J_{NN}$ modulation curves for amide N-H···Nδ1 histidine hydrogen bonds in the conserved T/S/NXXH motif of hemoglobins and HPX motif of cytochrome b_5

4. NH HSQC, long-range HMQC, and HNN-COSY spectra used for Asp82 N-H···Nδ1 His80 H-bond assignment within ferric *bis*-histidine cytochrome b_5

Supporting Information Figures

Figure S7.1. Cartoon depiction of helix capping H-bond "strain" and "relaxation" in GlbNs, as probed by $^{2h}J_{NN}$

a. Canonical *b* heme cofactor (present in non crosslinked hemoglobins and cytochrome b_5)
b. Covalent heme modification observed in GlbN-A: the native crosslink joins His117Nε-2-Cα-heme atoms and occurs on the opposite side of the proximal heme pocket with respect to the N80 N-H···Nδ1 H83 helix capping H-bond.
c. Engineered covalent heme modification present in L79H/H117A GlbN-B. A Leu79His replacement is sufficient to produce a nonnative His79 Nε-2-Cα-heme linkage. The N80-H83 helix capping H-bond is adjacent to His79 in both sequence and space, and therefore $^{2h}J_{NN}$ is expected to be a sensitive reporter for H-bond perturbation. This linkage is analogous to that engineered in the Leu75His *C. eugametos* CtrHb-CN-B variant.
d. Nonnative covalent heme modification present in L79H GlbN-AB. This *bis*-alkylated protein contains both native His117-2-Cα and engineered His79-4-Cα-heme covalent linkages.

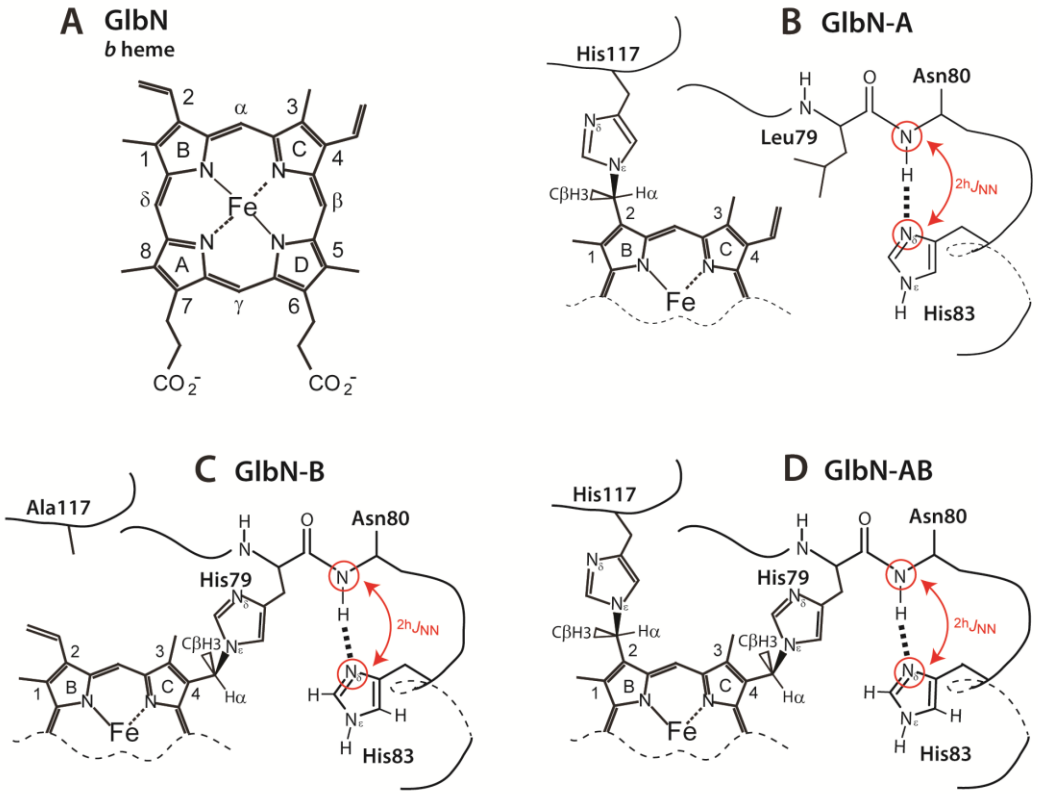


Figure S7.2. *C. eugametos* CtrHb-CN: initial experiments showing $^{1h}J_{HN}$ and $^{2h}J_{NN}$ couplings due to the Ser76 N-H···Nδ1 His79 H-bond

a. Initial evidence for homonuclear ^{15}N-^{15}N hydrogen bond scalar coupling. High-resolution ^{1}H-^{15}N HSQC (downfield ^{1}H region, ^{15}N t1 acquisition time = 300 ms) showing partially resolved Ser76 NH ^{15}N doublets. No other amide correlations in the spectrum display this splitting.
b. ^{15}N 1-D slice of the Ser76 major NH signal (^{1}H = 11.17 ppm).
c. High resolution ^{1}H-^{15}N HSQC (downfield ^{1}H region, ^{15}N t1 acquisition time as in a.) collected with a frequency-shifted ^{15}N decoupling pulse (histidine region, 255 ppm) during the ^{15}N amide t_1 period. With this pulse sequence modification, the ^{15}N-^{15}N hydrogen bond scalar coupling is refocused and Ser76 NH signals collapse into singlets.

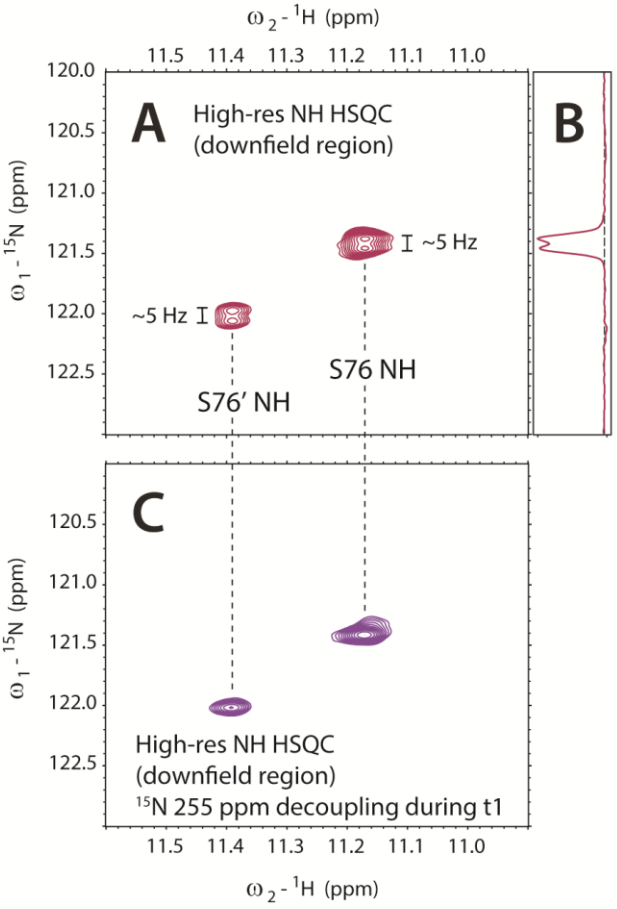


*figure continued on next page

d. High-sensitivity ^{1}H-^{15}N long-range HMQC (downfield ^{15}N, ^{1}H region), with ^{15}N decoupling (210 ppm, 298 K) during ^{1}H acquisition. The detection of the very weak doublet provided initial evidence for the identity of Ser76 N-H···Nδ1 His79 H-bonding partners. Here, the one-bond hydrogen bond scalar coupling, $^{1h}J_{HN}$, is responsible for coherence transfer. The directly detected Ser76 amide proton signal remains split due to $^{1}J_{NH}$ scalar coupling with its attached ^{15}N.

e. As a test for the above hypothesis, a ^{1}H-^{15}N long-range HMQC spectrum (downfield ^{15}N, ^{1}H region) was acquired with frequency-shifted ^{15}N decoupling (122 ppm, 313 K) during ^{1}H acquisition. The $^{1h}J_{HN}$–mediated correlation between Ser76 amide ^{1}H and His79 Nδ1 nuclei is observed as in d. However, in this instance, the $^{1}J_{NH}$ coupling is refocused and the Ser76 amide ^{1}H collapses into a singlet.

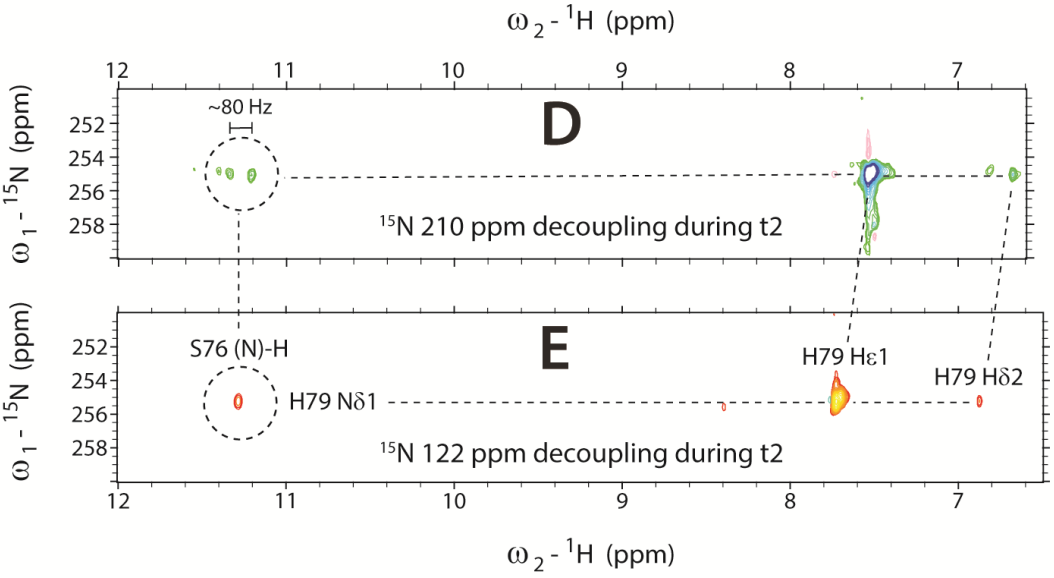


Figure S7.3. $^{2h}J_{NN}$ modulation curves and fits obtained from quantitative spin-echo difference 1-D HSQC experiments: comparison of protein N-H···Nδ1 helix capping hydrogen bonds

Data points from an individual experiment series were normalized and plotted (black circles). Best fit curves are shown with red lines. In most cases, the fitting errors were < 1%, however the errors reported for each $^{2h}J_{NN}$ value combines an estimate for the uncertainty in peak intensity (~5 %) as estimated from duplicate data points. The black horizontal line marks zero intensity, and the vertical dashed line corresponds to the null time $^{2h}J_{NN}$ = 5 Hz ($1/2J$ = 100 ms).

a. *C. eugametos* cyanomet CtrHb Ser76 N-H···Nδ1 His79 (major isomer, 313 K)
b. *C. eugametos* cyanomet CtrHb Ser76 N-H···Nδ1 His79 (minor isomer, 313 K)
c. *C. eugametos* cyanomet CtrHb, Ser76 N-H ···Nδ1 His79 (major isomer, 283 K)
d. *Synechocystis bis*-histidine ferric GlbN, Asn80 N-H···Nδ1 His83 (non crosslinked, 313K)
e. *Synechococcus bis*-histidine ferric GlbN-A, Asn80 N-H···Nδ1 His83 (native His117-heme crosslink, 313K)
f. *Synechocystis* L79H/H117A *bis*-histidine ferric GlbN-B, Asn80 N-H···Nδ1 His83 (engineered His79-heme crosslink, 313K)
g. *C. eugametos* L75H cyanomet CtrHb-B Ser76 N-H···Nδ1 His79 (engineered His75-heme crosslink, 313 K)
h. *Synechocystis* L79H *bis*-histidine ferric GlbN-AB, Asn80 N-H···Nδ1 His83 (His117-heme, His79-heme double crosslink, 313K)
i. *Synechocystis* cyanomet GlbN, Asn80 N-H···Nδ1 His83 (non crosslinked, 313K)
j. *Synechocystis* cyanomet GlbN-A, Asn80 N-H···Nδ1 His83 (native His117-heme crosslink, 313K)
k. *C. reinhardtii* cyanomet CtrHb, Asn87 N-H···Nδ1 His90 (major isomer, 313 K)
l. *C. reinhardtii* cyanomet CtrHb, Asn87 N-H···Nδ1 His90 (minor isomer, 313 K)
m. *Synechocystis* L79H/H117A cyanomet GlbN-B, Asn80 N-H···Nδ1 His83 (engineered His79-heme crosslink, 313K)
n. *Synechocystis* L79H cyanomet GlbN-AB, Asn80 N-H···Nδ1 His83 (engineered His117-heme, His79-heme double crosslink, 313K)
o. *Synechococcus bis*-histidine ferric GlbN, Asn80 N-H···Nδ1 His83 (non crosslinked, 313K)

p. *Synechococcus bis*-histidine ferric GlbN-A, Asn80 N-H···Nδ1 His83 (native His117-heme crosslink, 313K)
q. Rat microsomal cytochrome b_5, Asp 82 N-H···Nδ1 His80 (ferric *bis*-histidine, 313K)
r. Rat microsomal cytochrome b_5, Asp 82 N-H···Nδ1 His80 (ferrous *bis*-histidine, 313K)
s. Rat microsomal apocytochrome b_5, Asp 82 N-H···Nδ1 His80 (apoprotein, 298K)

*See next page for figure

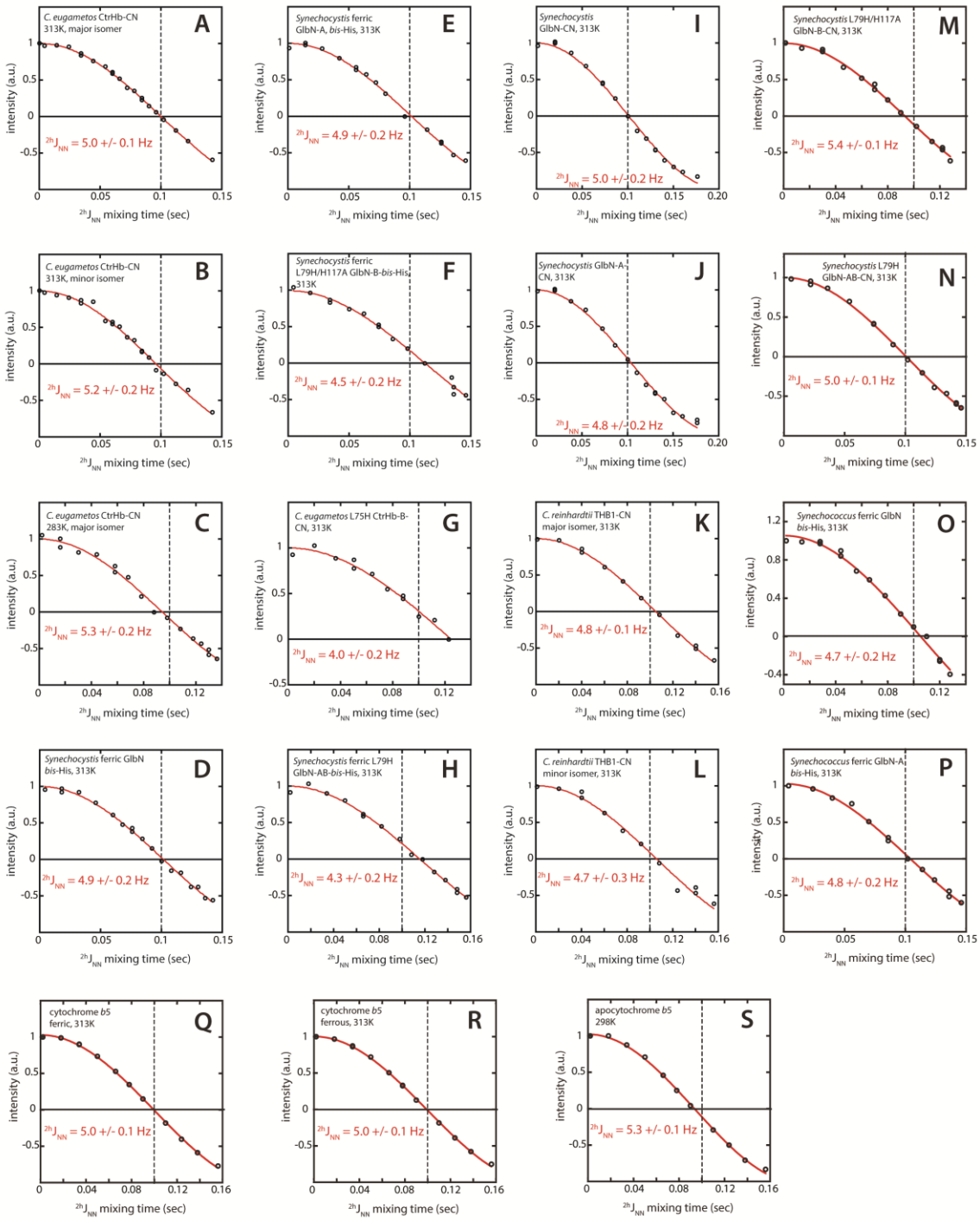
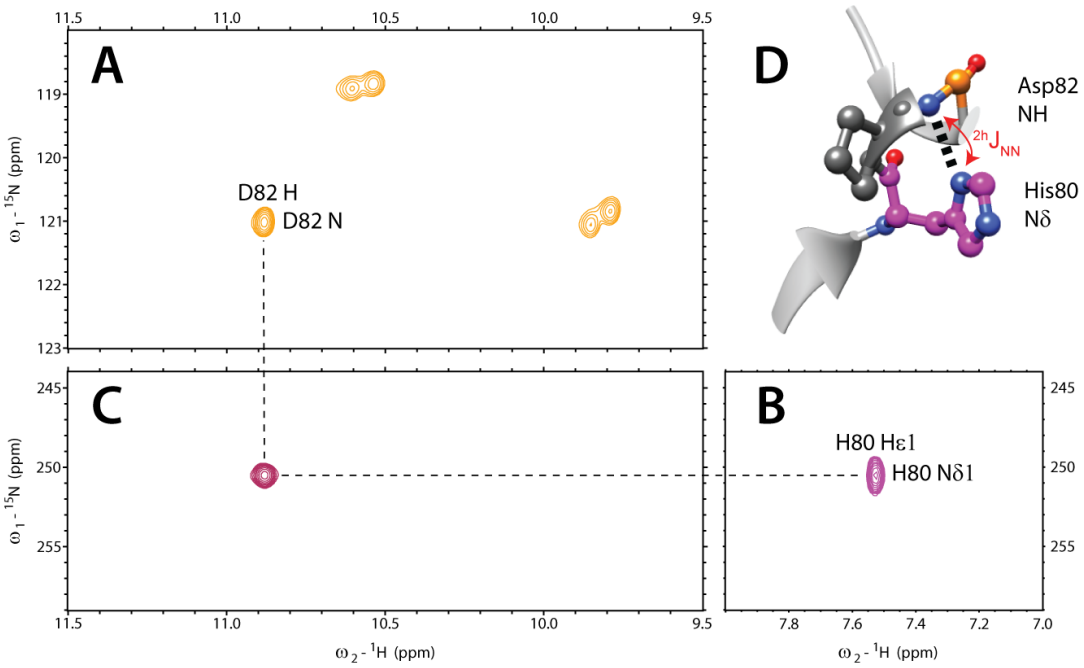

A
C. eugametos CtrHb-CN 313K, major isomer
$^{2h}J_{NN}$ = 5.0 +/- 0.1 Hz
B
C. eugametos CtrHb-CN 313K, minor isomer
$^{2h}J_{NN}$ = 5.2 +/- 0.2 Hz
C
C. eugametos CtrHb-CN 283K, major isomer
$^{2h}J_{NN}$ = 5.3 +/- 0.2 Hz
D
Synechocystis ferric GlbN bis-His, 313K
$^{2h}J_{NN}$ = 4.9 +/- 0.2 Hz
E
Synechocystis ferric GlbN-A, bis-His, 313K
$^{2h}J_{NN}$ = 4.9 +/- 0.2 Hz
F
Synechocystis ferric L79H/H117A GlbN-B-bis-His 313K
$^{2h}J_{NN}$ = 4.5 +/- 0.2 Hz
G
C. eugametos L75H CtrHb-B-CN, 313K
$^{2h}J_{NN}$ = 4.0 +/- 0.2 Hz
H
Synechocystis ferric L79H GlbN-AB-bis-His, 313K
$^{2h}J_{NN}$ = 4.3 +/- 0.2 Hz
I
Synechocystis GlbN-CN, 313K
$^{2h}J_{NN}$ = 5.0 +/- 0.2 Hz
J
Synechocystis GlbN-A-CN, 313K
$^{2h}J_{NN}$ = 4.8 +/- 0.2 Hz
K
C. reinhardtii THB1-CN major isomer, 313K
$^{2h}J_{NN}$ = 4.8 +/- 0.1 Hz
L
C. reinhardtii THB1-CN minor isomer, 313K
$^{2h}J_{NN}$ = 4.7 +/- 0.3 Hz
M
Synechocystis L79H/H117A GlbN-B-CN, 313K
$^{2h}J_{NN}$ = 5.4 +/- 0.1 Hz
N
Synechocystis L79H GlbN-AB-CN, 313K
$^{2h}J_{NN}$ = 5.0 +/- 0.1 Hz
O
Synechococcus ferric GlbN bis-His, 313K
$^{2h}J_{NN}$ = 4.7 +/- 0.2 Hz
P
Synechococcus ferric GlbN-A bis-His, 313K
$^{2h}J_{NN}$ = 4.8 +/- 0.2 Hz
Q
cytochrome b5 ferric, 313K
$^{2h}J_{NN}$ = 5.0 +/- 0.1 Hz
R
cytochrome b5 ferrous, 313K
$^{2h}J_{NN}$ = 5.0 +/- 0.1 Hz
S
apocytochrome b5 298K
$^{2h}J_{NN}$ = 5.3 +/- 0.1 Hz
intensity (a.u.)
$^{2h}J_{NN}$ mixing time (sec)

Figure S7.4. HSQC, long-range HMQC, and HNN-COSY correlation spectra used for Asp82 N-H···Nδ1 His80 H-bond assignment in ferric *bis*-histidine cytochrome b_5

a. ^{1}H-^{15}N HSQC (downfield ^{1}H region) showing the Asp82 NH resonance. The major and minor isomer (6:4) amide signals are indistinguishable.
b. Histidine selective ^{1}H-^{15}N long-range HMQC (^{15}N downfield region) showing the single Nδ1-Hε1 correlation expected from a neutral His80 Nε2-H tautomer.
c. HNN-COSY spectrum. A single crosspeak is observed, which correlates the Asp82 amide ^{1}H with His80 ^{15}Nδ1 via the two-bond HBC, $^{2h}J_{NN}$.
d. Together, the NMR data are clear evidence for N-H···N hydrogen bonding in the HPX 3_{10} helix capping motif (modeled using cytochrome b_5, PDB: 4HIL).



Vita

Matthew R. Preimesberger was born in Appleton, Wisconsin, in 1983. As a child, he developed interests in the natural sciences and mathematics. He also loved to play baseball, run, and wrestle as a young man. Upon graduating from Appleton East High School in 2002, he enrolled that fall at Lakeland College, in Sheboygan, Wisconsin. While double-majoring in chemistry and mathematics, he developed an interest for biophysical chemistry and spectroscopy. In the autumn of 2007, he was accepted into the Program in Molecular Biophysics, at Johns Hopkins University, in Baltimore, Maryland. In 2008, he joined the laboratory of Dr. Juliette Lecomte to study NMR spectroscopy, protein structure, and heme metalloenzymology.