9-2000

Crystal Structure of a Nonsymbiotic Plant Hemoglobin

Mark S. Hargrove
Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames IA 50011, USA

Eric Allen Brucker
Baxter Hemoglobin Therapeutics, Boulder, CO 80301, USA

Boguslaw Stec
Rice University

Gautam Sarath
University of Nebraska - Lincoln, Gautam.sarath@ars.usda.gov

Raúl Arredondo-Peter
Centro de Investigación sobre Fijación del Nitrógenó, Apartado Postal 565-A, Cuernovaca, Morelos, Mexico

See next page for additional authors

Published in *Structure* 8:9 (September 2000), pages 1005–1014. doi:10.1016/S0969-2126(00)00194-5

Submitted June 12, 2000; revised July 19, 2000; accepted July 19, 2000; published September 1, 2000; published online September 8, 2000.

Crystal structure of a nonsymbiotic plant hemoglobin

Mark S Hargrove[a]*, Eric Allen Brucker[b], Boguslaw Stec[c], Gautam Sarath[d], Raúl Arredondo-Peter[e], Robert V Klucas[d], John S Olson[c], and George N Phillips, Jr[c]

[a] Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames IA 50011, USA
[b] Baxter Hemoglobin Therapeutics, Boulder, CO 80301, USA
[c] Department of Biochemistry and Cell Biology WM Keck Center for Computational Biology, Rice University, Houston, TX 77005, USA
[d] Department of Biochemistry, University of Nebraska–Lincoln, Lincoln, NE 68588, USA
[e] Centro de Investigación sobre Fijación del Nitrógenó, Apartado Postal 565-A, Cuernovaca, Morelos, Mexico
* Corresponding author, email: msh@iastate.edu

Abstract

Background: Nonsymbiotic hemoglobins (nsHbs) form a new class of plant proteins that is distinct genetically and structurally from leghemoglobins. They are found ubiquitously in plants and are expressed in low concentrations in a variety of tissues including roots and leaves. Their function involves a biochemical response to growth under limited O_2 conditions.

Results: The first X-ray crystal structure of a member of this class of proteins, riceHb1, has been determined to 2.4 Å resolution using a combination of phasing techniques. The active site of ferric riceHb1 differs significantly from those of traditional hemoglobins and myoglobins. The proximal and distal histidine sidechains coordinate directly to the heme iron, forming a hemichrome with spectral properties similar to those of cytochrome b_5. The crystal structure also shows that riceHb1 is a dimer with a novel interface formed by close contacts between the G helix and the region between the B and C helices of the partner subunit.

Conclusions: The bis-histidyl heme coordination found in riceHb1 is unusual for a protein that binds O_2 reversibly. However, the distal His73 is rapidly displaced by ferrous ligands, and the overall O_2 affinity is ultra-high ($K_D \approx 1$ nM). Our crystallographic model suggests that ligand binding occurs by an upward and outward movement of the E helix, concomitant dissociation of the distal histidine, possible repacking of the CD corner and folding of the D helix. Although the functional relevance of quaternary structure in nsHbs is unclear, the role of two conserved residues in stabilizing the dimer interface has been identified.

Keywords: Anaerobic response; Hemoglobin; Nonsymbiotic; Plant; Structure

Introduction

For many years it was thought that hemoglobins in plants were found only in the root nodules of the legume family. These leghemoglobins have structures very similar to mammalian myoglobin and play an important role in nitrogen fixation. They prevent inhibition of the nitrogenase complex by binding O_2 tightly and facilitating its diffusion to nitrogen-fixing bacteroids [1]. A rationale for the origin of the leghemoglobins was not clear until the discovery in the late 1980s of nonsymbiotic hemoglobins (nsHb) in both nodulating and non-nodulating plant families [2–4]. The first nonsymbiotic hemoglobin was identified in barley based on its similarity to one of the non-legume, symbiotic plant hemoglobins [5]. The nucleotide sequence of the barley protein facilitated the identification of nsHbs in virtually all other plant species examined, including soybean, *Arabidopsis*, maize, and rice [6–8]. However, until now, the three-dimensional structure of members of this class of plant proteins was unknown.

Two discrete classes of nsHbs have been found in a wide variety of plant tissues, suggesting more than one physiological function [7,8]. Expression is constitutive in some plants and induced by hypoxia in others. Sowa *et al.* [9] demonstrated that nsHb is necessary for the survival of corn-cell suspensions under hypoxic conditions and suggested that these proteins help to maintain energy charge in the absence of oxidative phosphorylation. On the basis of these observations three possible functions have been proposed for nsHbs. First, they might be O_2-transport proteins that scavenge oxygen under hypoxic conditions and supply it for aerobic respiration. Second, they might act as terminal oxidases that, by virtue of their ultra-high O_2 affinity, facilitate glycolytic generation of ATP by removing NADH under microaerobic conditions. Third, they could be O_2-sensing proteins that activate other proteins with regulatory functions [6–9]. Evaluation of these hypotheses requires determination of the structures of the nsHbs and detailed analyses of their functional properties.

The spectral properties of both the ferric and ferrous forms of all the nsHbs resemble those of cytochrome b_5, indicating that the heme iron is six-coordinated in both oxidation states [8,10]. Mutagenesis suggests that the distal histidine (His73 in riceHb1) is coordinated directly to the iron atom in both the deoxy and ferric forms (Table 1).

Table 1. Rate and equilibrium constants for oxygen binding to selected hemoglobins[a]

Protein	k_{on} (μM^{-1} s^{-1})	k_{off} (s^{-1})	K (μM^{-1})
swMb	17	15	1.1
swMb H64L	98	4100	0.023
Lba	130	5	23
Lba H61L	400	24	17
riceHb1	68	0.038	1800
riceHb1 H73L	620	51	12

[a] Sperm whale Mb (swMb) and its H64L mutant demonstrate the degree of bound oxygen stabilization resulting from the His64 hydrogen bond [40]. Lba and its homologous distal histidine mutant demonstrate that bound oxygen is probably not stabilized by electrostatic interactions [14]. RiceHb1 has an affinity for oxygen 100 times greater than Lba and 1000 times greater than swMb, and His73 in riceHb1 stabilizes bound oxygen to a degree similar to that of swMb [8].

In contrast, almost all animal hemoglobins and plant leghemoglobins contain a pentacoordinate heme in the deoxy state and coordinate water when oxidized. One major goal of our crystallographic study was to confirm more directly that the iron atom coordinates both the proximal and distal histidines in riceHb1.

Even though the iron atom is six-coordinate, nsHb is still able to bind O_2 rapidly and with unusually high affinity compared with most other hemoglobins and myoglobins (Table 1). We previously hypothesized that the distal histidine rapidly dissociates from the iron atom but remains close enough to stabilize bound O_2 by formation of a strong hydrogen bond with either the N_δ–H or N_ε–H atoms of the imidazole sidechain [8]. This mechanism requires significant changes in the orientation of sidechains in the active site. Again, a structure of at least one form of a plant nsHb is needed to verify these ideas and to propose a more specific structural mechanism.

This work describes the three-dimensional structure of riceHb1, the first nonsymbiotic plant hemoglobin to be examined successfully by X-ray crystallography. The resultant model has been compared in detail with the crystal structures of sperm whale myoglobin (Mb), lupin and soybean leghemoglobins (Lbs), and several nonvertebrate hemoglobins including that from *Caudina arenicola*, which is the only other example of a hemoglobin with bis-histidyl heme coordination [11]. These comparisons provide a tentative molecular mechanism for oxygen binding.

Results

General features of riceHb1

The structure of ferric riceHb1 was solved to 2.4 Å using a combination of phases determined by anomalous scattering and one mercury isomorphous derivative. The model shows that the protein is a homodimer containing two 165-residue subunits, each having a heme *b* prosthetic group with bonds between the heme iron and the distal and proximal histidines (Figure 1).

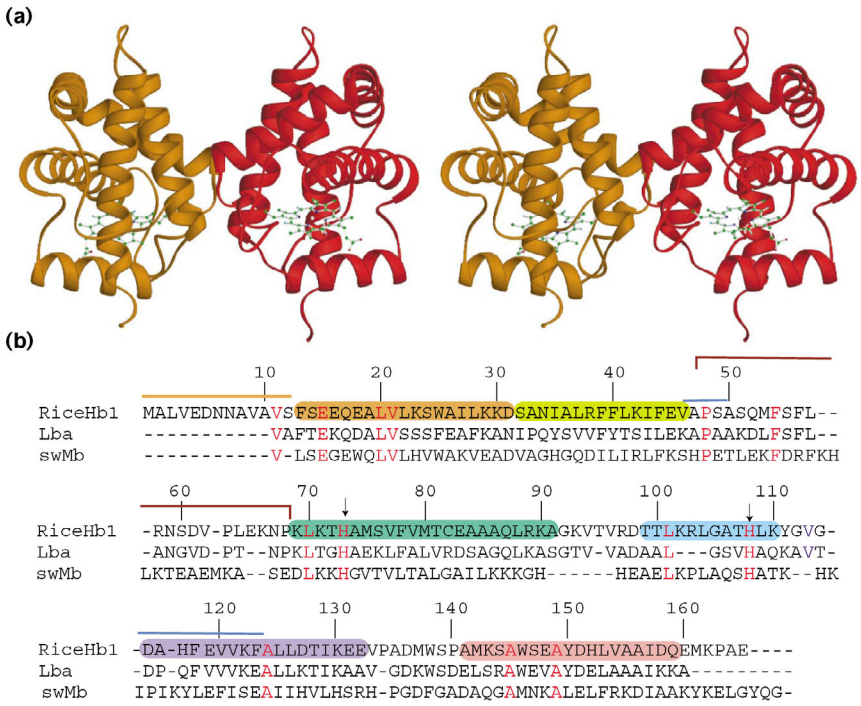


Figure 1. The structure of riceHb1. **(a)** Stereoview ribbon structure of dimeric riceHb1 with each helix labeled according to the typical globin fold. **(b)** Sequence alignment of riceHb1, Lba, and sperm whale Mb. The residues corresponding to riceHb1 helices are labeled and colored; the residue numbering is for riceHb1, with the initiator methionine as 0. Also indicated are the pre-A helix N-terminal extension, and the CD region. Residues involved in contacts at the subunit interface are overlined in blue. Identical residues are in red.

The tertiary structure consists of six helices that correspond to the A, B, E, F, G, and H helices of typical hemoglobins (Figure 1). In Figure 1b the primary sequence of riceHb1 is compared with those of sperm whale Mb and Lba, and the residue positions are marked with respect to their helical positions in the overall globin fold. RiceHb1 and Lba share 40% sequence identity, whereas the identity between riceHb1 and sperm whale Mb or between Lba and Mb is much less (~20%). The extended polypeptide linking the B and E helices (the CD region) in riceHb1 is similar in length to those found in most globins, but shows a structure quite different from that found in Lba, Mb, and the β subunits of human hemoglobin. A C helix is present in riceHb1 but it is very short and distorted, encompassing only two turns (residue 48–54). The D-helix region is totally unwound and forms an extended, poorly ordered loop. The heme pocket contains a well-ordered porphyrin ring and two histidine sidechains coordinated to the iron atom. These sidechains have been identified previously as His73 on the distal side and His108 on the proximal side by sequence alignments and mutagenesis [8]. In addition to the amino acids and prosthetic groups, the final riceHb1 model contains 208 crystallographically defined water molecules.

The average temperature factor for all atoms in each amino acid is plotted versus residue number in Figure 2. In general, the core of the tertiary structure is well defined and, with two exceptions, is covered with high quality electron-density maps. The pre-A segment (residues 1–13) is highly mobile and disordered, and the CD corner (residues 56–67) is also poorly defined. Unexpectedly, many of the disordered sidechains in the CD corner, namely Met53, Phe54, Phe56, Leu57, Val62, Pro63 and Leu64, are apolar. The models for these two regions describe the most dominant conformers and illustrate the volume and probable location of the polypeptide fragments. The backbone stereochemical parameters for the model of riceHb1 are very good, despite the two regions of extensive disorder. More than 97% of the torsional angles are in the allowed regions of a Ramachandran plot (Table 2).

Dimer interface

The subunit interface is formed by contacts between five amino acids at the end of the B helix and beginning of the C helix, and a nine-residue section at the beginning of the G helix (Figure 1 and Figure 3). The total surface area buried as a result of dimerization in riceHb1 is 554 Å^2 per subunit, which is at the lower end of what is expected for a stable dimer [12]. The dimer model is quite symmetrical, even though it was refined without noncrystallographic-symmetry constraints. The root mean square deviation (rmsd) of the positions of the two sets of Cα atoms representing individual subunits is ≤ 0.3 Å when the disordered residues (1–13 and 56–67) are omitted. The latter regions of both subunits are in close proximity and form an extended, highly mobile boundary. The disordered regions are involved in different symmetry-related crystal contacts, suggesting that crystallization does not cause the disorder. This observation also suggests that it is unlikely that crystallization causes dimerization.

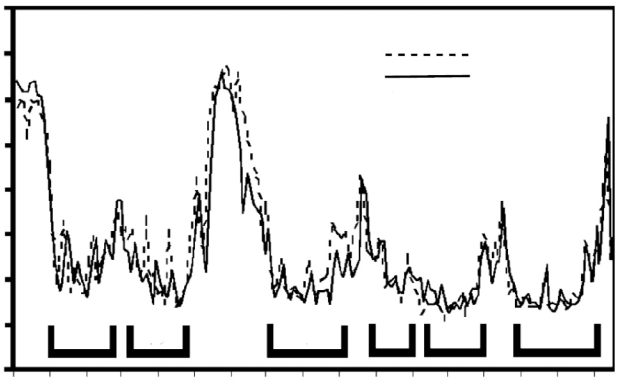


Figure 2. Temperature-factor plot. Average temperature (B) factors for each amino acid in both subunits are presented versus residue number. The high mobility regions are residues 1–12 and 57–67. There is much lower mobility at major helices. The last residue (166) is the heme group.

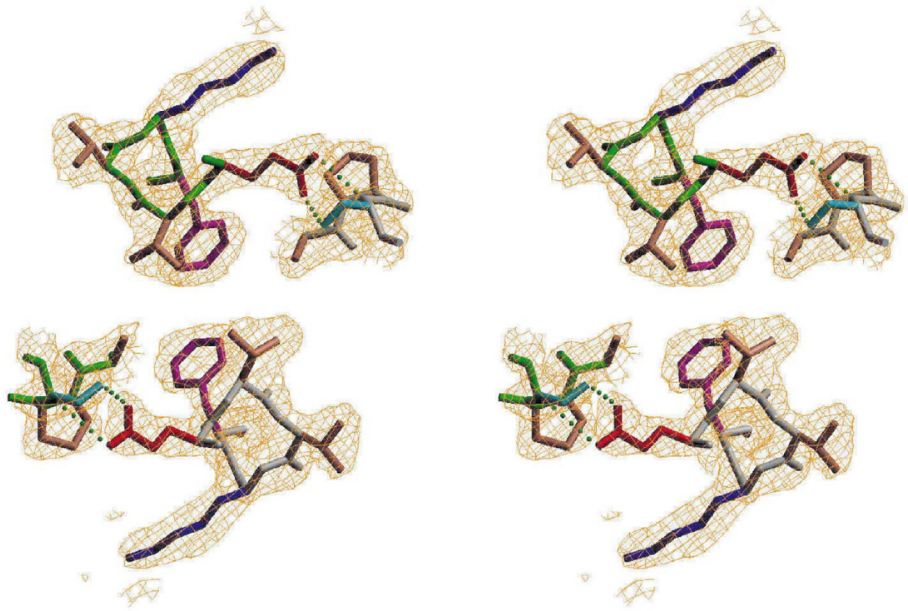
The amino acid sidechains that form the dimer interface are shown in Figure 3a. The closest contacts are a pair of 2.9 Å hydrogen-bond interactions between the Ser49 sidechain of one subunit and the Glu119 sidechain of the other. The other sidechains in this region (Phe123, Val120, and Val46) form a hydrophobic pocket that is capped at each end by the Glu119–Ser49 hydrogen bond. The amino acid sequence in this region is shown in Figure 3b along with the homologous sequences from

Table 2. Refined model parameters for riceHb1

Residues	330
Protein atoms	2576
Water molecules	208
No. reflections in the refinement	19,025
R factor*	0.208
R_{free}	0.261
B factors (Å^2)	
Mainchain†	54.2 (44.4)
Sidechain†	58.6 (50.6)
Rmsd‡	
Bonds (Å)	0.009
Angles (°)	2.11
Planarity (°)	1.85
Ramachandran plot (%)	
Preferred	86.3
Allowed	10.8
Generously allowed	2.9
Disallowed	0.0

* Calculated on all reflections (F>0σ(F)) in 20–2.3 Å resolution shell. † Average values calculated with disordered regions omitted are given in parentheses. ‡ Rmsd, root mean square deviation.

(a)



(b)

	E V A P S A S E I A P S A S E I A P S A Q	H F E V V K F A L L H F E V T R F A L L H F E V T K F A L L
	E I A P S A K	H F E V T R F A L L
	E K A P A A K E K A P A A K E I A P A A K	Q F V V V K E A L L Q F V V V K E A L V H F P V V K E A I L

Figure 3. The dimer interface of riceHb1. **(a)** A stereo image of the electron density at the dimer interface shows that Ser49 and Glu119 of each residue form an electrostatic interaction that caps a hydrophobic core comprised of the sidechains of Val120, Val46, and Phe123. **(b)** Sequence alignment of dimeric nsHbs and monomeric symbiotic plant Hbs shows that the residues making up this capped hydrophobic core are conserved in the nsHbs. The conserved apolar and polar residues are outlined in blue and yellow, respectively.

several other monomeric and dimeric plant hemoglobins. All nsHbs included in Figure 3b are reported to be dimeric and have serine and glutamate at positions homologous to those found at positions 49 and 119 in riceHb1. The symbiotic hemoglobin from *Parasponia* is also dimeric and contains the consensus sequence found in the nsHbs 2 and 13. In monomeric symbiotic leghemoglobins the amino acids homologous to Ser49 and Glu119 of riceHb1 are always hydrophobic. In addition Phe123 in nsHb is replaced by glutamate at the corresponding position in the leghemoglobins, which should disrupt the hydrophobic core and prevent dimer formation.

Comparison of riceHb1 to Lba and Mb

The riceHb1 subunit shares significant structural identity with Lba (40%) [14]. An overlay of the backbone ribbon structures of these two proteins includes the heme groups and distal and proximal histidines and is shown in Figure 4. Lba contains 144 residues, 21 less than riceHb1. The elongated N and C termini of riceHb1 account for 13 of these residues. The F helix of riceHb1 contains 12 amino acids. In contrast, the F helix of Lba contains only seven residues; however, it should be noted that the F helices of most other hemoglobins are 10–12 residues in length [15,16].

The most striking difference between riceHb1 and Lba is the unraveled and outward position of the CD corner in the nonsymbiotic hemoglobin (right-hand portion of Figure 4). This unfolded conformation seems to be dictated by a position of the E helix in riceHb1 that is bent toward the center of the porphyrin ring. The bend in the E helix is caused by direct coordination of His73 with the iron atom. The amino acid is the seventh along the E helix and can be designated His(E7) or the distal histidine. The distance between the heme iron and the Cα atom of the distal histidine changes from 8.7 Å in Lba to 6.3 Å in riceHb1. The shift in the E helix, coordination of His73, and the elongated F helix cause the heme group in riceHb1 to be 'pulled' ~1 Å toward the solvent compared with its position in Lba (Figure 4).

The distal portion of the heme pocket of riceHb1 is relatively large and, with the exception of His73, contains no hydrophilic residues. As shown in Figure 5, Nε of the distal histidine His73(E7) coordinates directly to the heme iron atom with a bond length of 2.2 Å. The other residues nearest to the ligand-binding site in riceHb1 are Phe40(B10), Phe54(CD1) and Val77(E11).

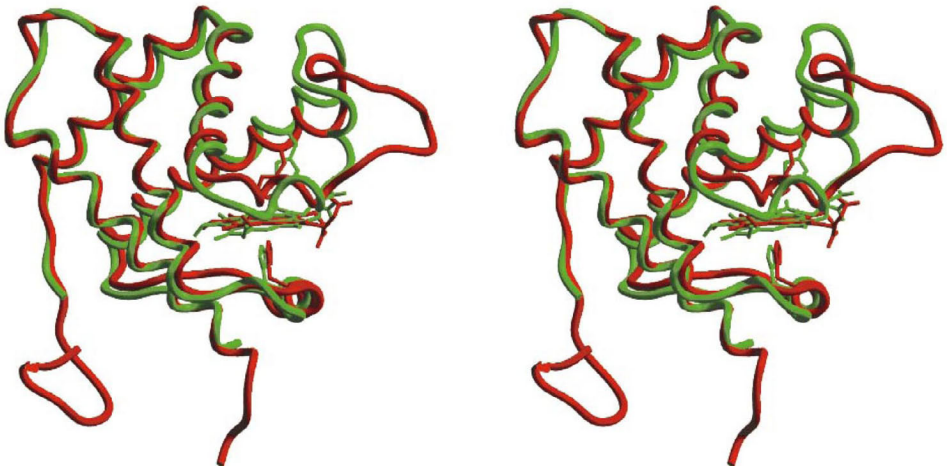


Figure 4. Comparison of riceHb1 and soybean Lba. A stereo overlay of a riceHb1 subunit (red) with Lba (green) shows that the tertiary structures of these two proteins are very similar, with the exception of the N-terminal half of the E helix. The heme groups of these two proteins are translated such that the heme in riceHb1 is ~1 Å closer to solvent than that of Lba. The residues of the CD corner extend deep into the solvent. The Lba structure used in this figure has accession code 1bin in the Protein Data Bank (PDB).

Phe(CD1) is conserved in all known reversible oxygen-binding globins, and the E11 residue in almost all globins is either valine or leucine. Both of these residues confer resistance to autooxidation and hemin dissociation and help to orient bound O_2 for hydrogen bonding to the distal histidine [17–19].

The edge of the Phe40(B10) sidechain is located very close to the Cγ atom of His73(E7) as a result of coordination of the latter residue with the heme iron atom. As shown in Figure 5, the quality of the electron-density map assigned to Phe40 is significantly lower than those of neighboring amino acids. The observed disorder suggests at least three distinct positions for the phenyl sidechain. However, attempts to model these conformers lead to collisions with neighboring amino acids. Because phenylalanine is found at position B10 in all nonsymbiotic plant hemoglobins, it is likely that this disorder represents a functionally significant, unfavorable steric interaction that promotes dissociation of the coordinated distal histidine to allow O_2 binding. Preliminary mutagenesis studies indicate that the rate of His73 dissociation decreases from ~1900 s^{-1} in wild-type riceHb1 to ~200 s^{-1} when Phe40(B10) is replaced with smaller apolar residues (MSH and MD Goodman, unpublished results).

A comparison of the active sites of oxidized riceHb1, lupin LbNO, and sperm whale MbO_2 is shown in Figure 6. The structure of the lupin protein was used for comparison because the NO ligand is similar in size to O_2, which facilitates comparison with MbO_2. The Cε–Nε bond of the distal imidazole in riceHb1 lies parallel to the O–N and O–O bonds of the bound ligands in the lupin Lb and Mb structures. In both Lb and Mb the distal histidine is too far away to interact directly with the heme group, but the imidazole sidechains in both proteins form a hydrogen bond with the bound ligand. In both leghemoglobin and riceHb1 the plane of the proximal imidazole is staggered with respect to the pyrrole nitrogens, and the iron atom is slightly displaced (~0.2

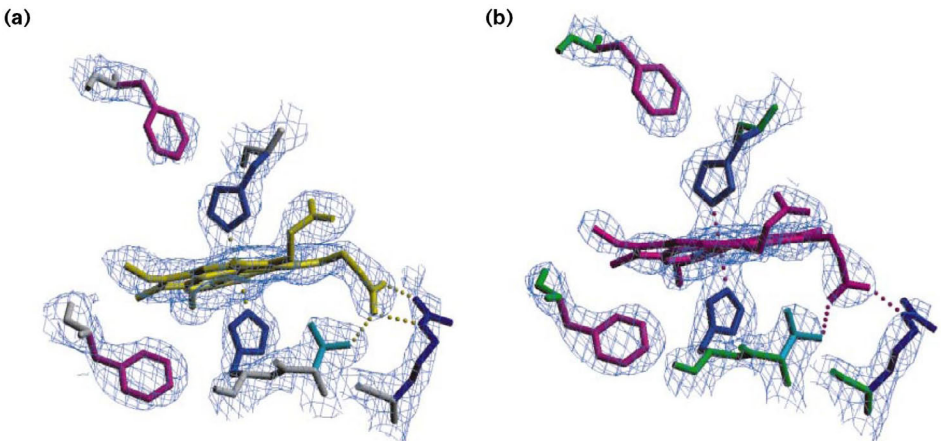


Figure 5. The bis-imidazole heme complex of riceHb1. The model of the riceHb1 heme pocket has two histidine residues coordinating the heme iron. The $2F_o–F_c$ electron-density map of the heme pocket shows that the sidechain of Phe40 is not well localized compared with the other residues. Panels **(a)** and **(b)** present the A and B subunits, respectively.

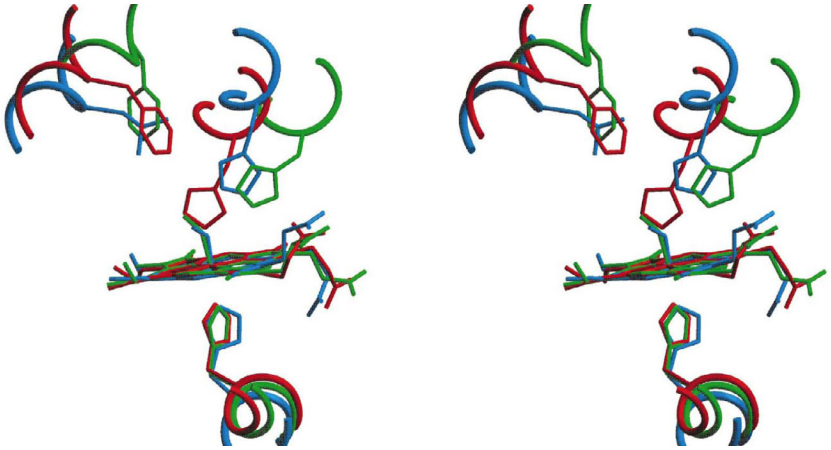


Figure 6. Comparison of the distal pockets of riceHb1, Mb, and lupin leghemoglobin. The heme groups of riceHb1 (red), lupin Lb (green), and sperm whale Mb (blue) were superposed. The proximal and distal histidine residues together with the residues at position B10 (Leu29 in sperm whale Mb, Phe40 in riceHb1 and Phe39 in lupin Lb) are shown together with the heme group and NO ligand of Lb, and the heme group and O_2 in sperm whale Mb. His73 in riceHb1 must move to accommodate ligand binding in this protein, and is probably discouraged from moving inward by the presence of Phe40. The PDB files used in this figure have accession codes 1gdl [39] (leghemoglobin) and 1a6m (sperm whale Mb).

Å) toward the distal pocket (Figure 6). These structural results explain the high reactivity of the iron atom in riceHb1 toward all ligands, including O_2 and the distal histidine (Table 1).

Discussion

A model for ligand binding to nsHb

The bis-imidazole complex found in both oxidized and reduced riceHb1 suggests that there should be a large equilibrium and kinetic barrier to ligand binding. However, as shown in Table 1, riceHb1 has a remarkably high affinity for O_2 that is ~100-fold greater than that of Lba. The association rate constant for O_2 binding to riceHb1 is also large. Thus, the distal histidine is readily displaced.

In principle, dissociation of the coordinated distal histidine could involve either inward or outward rotation about the Cα–Cβ bond. If the movement were toward the interior of the protein, the Phe(B10) sidechain would have to be displaced from its position in the ferric crystal structure, causing a significant expansion of the region around the B helix. Inward movement would also place the polar imidazole sidechain in the hydrophobic interior of the distal pocket. Previous work with myoglobin mutants has shown that placing asparagine or histidine residues in the interior of the distal pocket (e.g., at CD1, E11, B10, and G8) results in very unstable apoglobins [20]. Therefore, it seems more likely that dissociation of His73 occurs by outward movement of the imidazole sidechain with concomitant upward and outward movement of the E helix. A model for this displacement is shown in Figure 6; O_2 binding would cause His73 and the E helix in riceHb1 (red) to move to positions similar to those seen for the distal histidine and E helix (blue) in sperm whale Mb. In this mechanism His73 is 'pushed' by steric hindrance from Phe40(B10) toward the heme propionates, but stays close enough to the iron atom to form a strong hydrogen bond with bound O_2.

Evidence in favor of this outward dissociation mechanism is fivefold. First, the Phe40(B10) sidechain is disordered and the average position is unfavorably close to the edge of the coordinated imidazole (Figure 5). Second, mutation of Phe40→Leu causes the rate of His73 dissociation to decrease ~tenfold from ~1900 s^{-1} to ~200 s^{-1}, suggesting strongly that the kinetic barrier to dissociation is lowered by steric hindrance with the larger, native sidechain (MSH and MD Goodman, unpublished results). Third, movement of the distal histidine toward the CD corner is less likely to be hindered because this region is already disordered in the structure of the bis-coordinated riceHb1 complex (Figure 1 and Figure 4). Fourth, outward and upward movement of the E helix would allow the short D helix and the end of the C helix to re-form in response to ligand binding (like that of Lba in Figure 4). Fifth, the distal histidine in riceHb1 stabilizes bound O_2 by a factor of ~1000 as judged from the effects of the His73→Leu mutation (Table 1).

This model for ligand binding is nearly the opposite of that predicted from the structures of *Caudina arenicola* Hb-C and Hb-D [11]. Hb-C is a monomeric hemichrome hemoglobin resulting from coordination of the distal and proximal histidines in a manner similar to that observed for riceHb1. The Hb-D chain shares 65% sequence identity with Hb-C but is dimeric in the cyanomet form. A comparison of the structures of these two proteins shows that a short D helix is present in the Hb-C hemichrome form, whereas it is absent in the cyanide-bound Hb-D structure. In contrast, there is no D helix in the ferric hemichrome state of riceHb1. Thus, the pattern for formation and disappearance of the D helix seems to be variable with respect to hexacoordination versus pentacoordination in plant and invertebrate hemoglobins.

As shown in Table 1, replacement of the distal histidine with leucine in sperm whale Mb and in riceHb1 causes dramatic increases in the rate constants for O_2

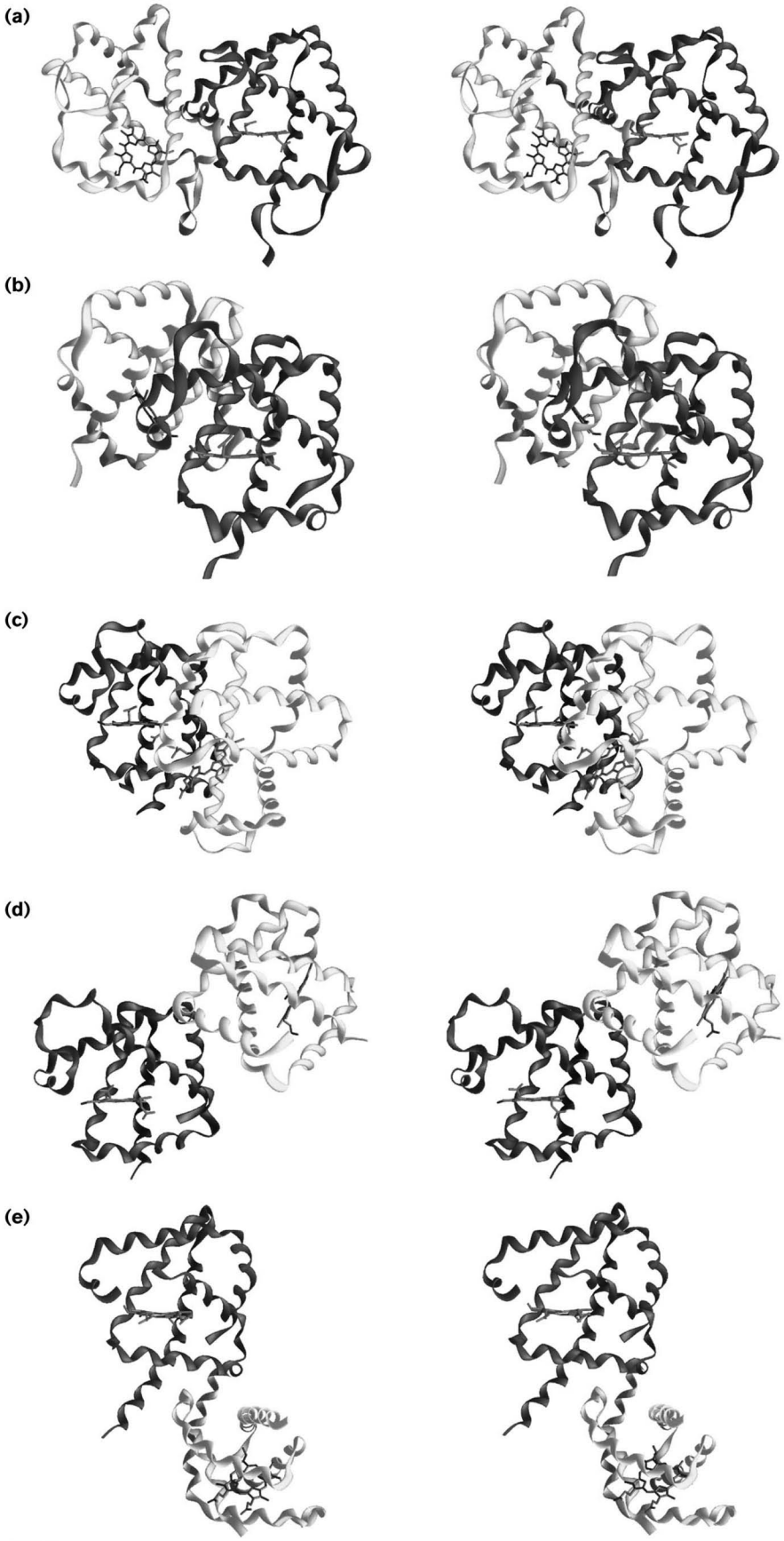


Figure 7. Structures of selected hemoglobin dimer interfaces. A stereo image of **(a)** riceHb1 is compared with that of **(b)** the human hemoglobin $\alpha_1\beta_1$ dimer (PDB code 3hhb), **(c)** the dimeric hemoglobin from *Scapharca inaequivalvis* (4hbi), **(d)** *Urechis caupo* hemoglobin (1ith), and **(e)** *Vitreoscilla* hemoglobin (1vhb). To aid comparison, each protein is presented so that the dark-gray subunit is in the same orientation. The dimer interfaces in all of these proteins are distinct from one another.

dissociation, k_{O2}. A much smaller increase in k_{O2} is observed for the His61→Leu mutation in Lba, presumably because the distal histidine is further away from the bound ligand in this protein and forms a weaker hydrogen bond (Figure 6). The His73→Leu mutation in riceHb1 also causes a tenfold increase in the association rate constant for O_2, k'_{O2}. The absolute value of k'_{O2} for Leu73 riceHb1 is very large and similar to that for Leu(E7) Lba. In each protein there is little or no barrier to ligand entry when the distal histidine is removed.

Comparison of K_{O2}, the equilibrium association constants, for three Leu(E7) mutants shows that the intrinsic affinities of the iron atoms for O_2 in both riceHb1 and Lba are very large. In contrast, the reactivity of the iron atom in Leu(E7) Mb is very low. In the absence of stabilization by the distal histidine, K_{O2} for Mb is roughly 1000-fold lower than those seen for the corresponding mutant plant proteins (Table 1). The high iron reactivity in Lba has been attributed to the staggered orientation of the proximal histidine with respect to the heme pyrole nitrogens. This staggered conformation allows in-plane movement of the iron atom with minimum steric hindrance between the edges of the imidazole ring and the coordinating nitrogen atoms of the porphyrin ring [21]. In contrast, the proximal imidazole in Mb is in an eclipsed position that inhibits ligand binding [22]. The orientation of the proximal imidazole in riceHb1 is identical to that observed in Lba (Figure 4 and Figure 6), accounting for its high reactivity towards ligands even when the coordinated distal histidine is present.

Dimer organization

The orientation of the subunits in riceHb1 represents a new hemoglobin quaternary structure and adds to the growing list of nonvertebrate hemoglobins with structures different from those of mammalian hemoglobins [15]. Figure 7 compares the riceHb1 dimer to three other dimeric structures of nonvertebrate hemoglobins and the $\alpha_1\beta_1$ dimer of human hemoglobin.

The contact between the two riceHb1 subunits involves 554 Å^2 of buried surface area per subunit between the CD region and the G helix. This is nearly 300 Å^2 less than that associated with the $\alpha_1\beta_1$ dimer of human hemoglobin (820 Å^2) [23], and much less than that of the cooperative homodimeric hemoglobin from the clam *Scapharca inaequivalvis* (980 Å^2) [24]. However, the buried surface area associated with riceHb1 dimer formation is similar to the buried surface areas of hemoglobins from the marine worm *Urechis caupo* (520 Å^2) [25] and from *Vitreoscilla* (430 Å^2) [26]. One possible explanation for the quaternary structure is a need for cooperative ligand binding. The dimeric hemoglobin from *Scapharca* does bind ligands cooperatively [15], but cooperative ligand binding has not been observed for riceHb1 or its homologs.

The equilibrium dissociation constant for dimerization of ferric riceHb1 is 86 μM (MD Goodman and MSH, unpublished results), which indicates weak interactions compared with those for the $\alpha_1\beta_1$ dimer of human hemoglobin [27]. This large K_D suggests that quaternary structure might not be physiologically relevant because the intracellular concentrations of these proteins are probably very low. However, there are examples of many hemoglobins (including those from *C. arenicola*) with quaternary structures that are dependent on the oxidation and coordination state of the heme iron 11 and 28. Until the equilibrium constants for association have been determined under these different conditions, firm conclusions cannot be drawn about the importance of quaternary structure for riceHb1 function.

Biological implications

Three hypotheses have been suggested for the function of nonsymbiotic plant hemoglobins (nsHbs). First, the nsHbs might be O_2-transport proteins that sequester O_2 under hypoxic conditions and facilitate its diffusion to cells that require aerobic mitochondrial respiration. Second, the nsHbs might be NADH oxidases that facilitate the glycolytic generation of ATP under microaerobic conditions. Third, the nsHbs might act as O_2-sensing proteins that undergo a substantial conformational change in response to ligand binding or release, and this conformational change activates other proteins or enzymes that have specific regulatory functions.

The high affinity of riceHb1 for O_2 and its relatively low concentration inside plant tissues argues against a role in the transport or facilitated diffusion of residual O_2 [29]. The tissue distribution of these proteins is not known well enough to rule out very small regions of high concentrations; however, there are no reports of 'red' regions in the roots or leaves of plants that express these nsHbs. The nsHbs are certainly not present at the levels of leghemoglobins found in the root nodules of nitrogen-fixing plants. Finally, even if present at high concentration, the O_2 dissociation rate constant for riceHb1 is too small to be physiologically relevant for O_2 delivery.

RiceHb1 does not act like a simple high-affinity, terminal oxidase because its rate of autooxidation is very slow at room temperature ($t_{1/2} \geq 1$ h). In order to consume O_2, $nsHbO_2$ needs to react with some other oxidizing compound and a cognate flavoprotein reductase to oxidize NADH. Examples of proteins that carry out this activity are the *Alcaligenes*, *Escherichia coli*, and yeast flavohemoglobins. These proteins function as NO dioxygenases and contain a heme domain, which reacts successively with O_2 and NO, and a flavoprotein domain that reacts with NADH 30 and 31. It is possible that the nsHbs are NO dioxygenases that lack a flavoprotein domain but have a corresponding reductase that is expressed by a separate gene. The affinities and rate con-

stants for O_2 binding to riceHb1 are similar to those of the flavohemoglobins [32]. However, all of the known flavohemoglobin NO dioxygenases have conserved Gln(E7) and Tyr(B10) residues in the distal pockets of the globin portions of the enzymes, whereas all of the nsHbs have histidine and phenylalanine, respectively, at these positions.

The bis-imidazole coordination in riceHb1 ensures that large tertiary conformational changes will occur in response to ligand binding. Comparison with the structure of Lba in Figure 4 suggests that deoxygenation and hemochrome formation results in the unfolding of the CD corner, an event that could easily trigger changes in affinity for the binding of auxiliary signal-transduction proteins if that binding occurs in the region between residues 50 and 70. Thus, nsHbs could readily act as sensors for extreme microaerobic or anaerobic conditions.

The diverse tissue localization displayed by the nsHbs and the presence of multiple genes in some plant species imply that the proteins have more than one function. There is clearly a link between the expression of nsHbs and the ability of a plant to survive under conditions of O_2 stress [9]. However, the underlying physiology of this linkage is still unclear. Thus, it might be that these proteins carry out two or even all three of the proposed functions; namely, O_2 storage and transport, NO and NADH oxidation, and O_2 sensing.

Materials and methods

Protein production and crystallization

The origin of the cDNA used for the expression of riceHb1 is described by Arredondo-Peter *et al.* [8]. This cDNA was inserted into the expression vector pET28a (Novagen, Inc.) where its expression was induced by T7 polymerase. BL21DE3 *E. coli* cells carrying this plasmid were grown at 37°C to an optical density of 0.6 at 600 nm, at which time protein expression was induced with 1 mM IPTG. The cells were then incubated at 37°C for a further 4 h, harvested, and lysed using lysozyme in 20 mM Tris pH 8.0. The resulting supernatant was bright red, containing soluble heme-bound riceHb1. The protein was purified using phenyl sepharose hydrophobic resin, and DEAE ion-exchange chromatography [8]. The protein binds phenyl sepharose (CL4B, Pharmacia) in 1.8 M $(NH_4)_2SO_4$, 20 mM Tris pH 8.0, and elutes in 0.7 M $(NH_4)_2SO_4$ at the same pH. RiceHb1 binds DEAE in 20 mM Tris pH 8.0, and elutes in the same buffer containing 50 mM NaCl.

Crystals of ferric riceHb1 were grown using the vapor-diffusion, hanging-drop method at room temperature. The well solution contained 2.2 M $(NH_4)_2HPO_4$, pH 7.8, and 20% sucrose. Protein (2 µl; equilibrated in 10 mM potassium phosphate, pH 7.0) at 2.2 mM (per heme) was mixed with 2 µl of well solution to form the drop on a siliconized cover slip. Crystals were seen as hexagonal columns in about five days.

Data collection and structure solution

Our in-house data were collected at 100K using a Siemens rotating anode, Rigaku Raxis II imaging plate, and an Oxford Cryostream cooling apparatus. These data sets were used to identify the riceHb1 space group as either $P3_121$ or $P3_221$, with unit-cell constants of a and b=126 Å, c=56 Å, α and β=90°, and γ=120°. The identification of $P3_121$ as the final space group resulted from analysis of electron density from the iron edge data described below.

Two diffraction data sets were used to solve the structure of this protein. Scaling and indexing of both data sets was accomplished using DENZO and Scalepack [33]. The highest resolution data were collected at beamline X4A of the National Synchrotron Light Source at Brookhaven National Laboratory. This was an anomalous data set using X-rays at 1.74 Å to maximize the anomalous signal from the heme iron. These data were 80% complete with a limiting resolution of 2.2 Å (Table 3). The anomalous difference Patterson Harker section calculated from the data set produced peaks that were initially attributed to a single iron atom per asymmetric unit.

The second set of data were collected in-house on a mercury derivative resulting from soaking a crystal for 1 h in 5 mM parachloromercuri-benzoate. The data were 95% complete to a limiting resolution of 3.2 Å (Table 3). The difference Patterson map calculated from these data were consistent with a single mercury binding site in each molecule (the mercury is bound to the Cys83 sidechain in this protein), suggesting a single molecule in the asymmetric unit. The combination of this data set with the iron edge data followed by solvent flattening using the PHASES software package [34] resulted in electron density with a heme group and some identifiable helices. An initial model for riceHb1 was built using these data and the program O [35]. However, further refinement was unable to lower the R value below ~40%.

Because the solvent content of the crystal assuming one molecule in the asymmetric unit was rather high (~70% with a Matthews coefficient of 5.6 A^3/Da), our data were re-evaluated with an effort to locate additional molecules. Upon closer inspection, the mercury derivative data were modestly nonisomorphous (the cross R_{merge} with the iron data was ~30%) and a more careful look at the difference Fourier maps revealed a weaker second mercury site. Introduction of two iron and two mercury sites followed by solvent flattening produced an interpretable electron-density map that showed two molecules in the asymmetric unit, with the two iron atoms clearly localized at the heme sites of both molecules.

Model building and refinement

The model of the dimer was built by duplicating a partially refined subunit and fitting it as a rigid body to the electron density. The process was facilitated by knowing the position of the iron atom from heavy-atom work. The program Chain [36] was used for subsequent fitting and corrections. This dimeric asymmetric unit was subjected to eight macrocycles of X-PLOR refinement [37]. An individual macrocycle consisted of simulated annealing (a slow-cooling protocol from 2000K to 100K), energy minimization, and temperature-factor refinement, followed by manual model building.

Table 3. Reduction statistics for diffraction data[a]

	Iron-edge	Mercury
Wavelength (Å)	1.74	1.54
Resolution range (Å)	20–2.3	30–3.2
Overall completeness*(%)	80 (46)	95
I/σ(I)*	14.2 (2.3)	
Unique reflections*	21,422 (1634)	8311
R_{merge}*†	0.069 (0.223)	0.093
Phasing power‡	2.02	1.97

[a] Two data sets were used in the solution of the structure of riceHb1. An iron-edge anomalous diffraction experiment was combined with an in-house mercury derivative. The iron-edge anomalous data were used as the native data set in the final refinement.

* Values in parentheses are for the 2.3–2.4 Å high-resolution shell.

† $R_{merge}=\Sigma_{hkl}\Sigma_i|I_{mean} - I_i|/\Sigma_{hkl}\Sigma_i I_i$.

‡ Phasing power$=\Sigma_{hkl}|F_{AHcalc}|/\Sigma_{hkl}|E|$, where F_{AHcalc} is the structure for anomalous or heavy-atom scattering and E is lack of closure.

Throughout this process, R_{free} (calculated from 10% of the data) was used to monitor progress. Initially, the noncrystallographic-symmetry constraints were used but were released for final refinement because the disordered regions proved to be quite different. The final model contains 208 water molecules, has an R factor 20.8%, R_{free} of 26.1%, and has good stereochemistry (97% of the backbone conformations are in the allowed regions of the Ramachandran plot; see Table 2).

There are two regions of significantly increased mobility and possible disorder (Figure 2). The representative models for these regions were obtained by trial and error model building guided by visual map interpretations. A family of related structural models was constructed using stereo chemical correctness and positive difference densities. The final model was selected on the basis of the best improvement in R factor (removal of the fragments leads to a higher R factor by ~0.5% R) and the production of the least amount of negative electron difference density near the proposed fragment. The higher than average temperature factors for these regions are due to the lack of using other conformers and the refinement of disordered atoms with full occupancy.

Analysis of buried surface area was carried out using the buried surface area algorithm in the CNS [38] software package with a probe radius of 1.4 Å[12]. Figures were generated using the programs Ribbons [39] (Figure 1 and Figure 7), and SETOR [40] (Figure 3, Figure 4, Figure 5 and Figure 6).

Accession numbers

This model has been assigned entry number 1d8u in the Protein Data Bank.

Acknowledgements

This work was supported by Robert A Welch Foundation grants C-1142 (GNP) and C-612 (JSO); National Institutes of Health grants AR40252 (GNP), HL47020 (JSO) and GM35649 (JSO); a grant from the National Library of Medicine (T15 LM07093) and the National Science Foundation (BIR-94-13229) (BS); and USDA award number 99-35306-7833 (MSH).

References

1. C.A. Appleby, Leghemoglobins and *Rhizobium* respiration, *Annu. Rev. Plant Physiol.* **35** (1984), pp. 443–478.
2. C.A. Appleby, J.D. Tjepkema and M.J. Trinick, Hemoglobin in a nonleguminous plant, *Parasponia:* possible genetic origin and function in nitrogen fixation, *Science* **220** (1983), pp. 951–953.
3. D. Bogusz and W.J. Peacock *et al.*, Functioning haemoglobin genes in non-nodulating plants, *Nature* **331** (1988), pp. 178–180.
4. C.A. Appleby, D. Bogusz, E.S. Dennis and W.J. Peacock, A role for haemoglobin in all plants, *Plant Cell Environ.* **11** (1988), pp. 359–367.
5. E.R. Taylor, X.Z. Nie, W.M. Alexander and R.D. Hill, A cereal haemoglobin gene is expressed in seed and root tissue under anaerobic conditions, *Plant Mol. Biol.* **24** (1994), pp. 853–862.
6. C.R. Andersson, E.O. Jensen, D.J. Llewellyn, E.S. Dennis and W.J. Peacock, A new hemoglobin gene from soybean; a role for hemoglobin in all plants, *Proc. Natl Acad. Sci. USA* **93** (1996), pp. 5682–5687.
7. B. Trevaskis and W.J. Peacock *et al.*, Two hemoglobin genes in *Arabidopsis thaliana*: the evolutionary origins of leghemoglobins, *Proc. Natl Acad. Sci. USA* **94** (1997), pp. 12230–12234.
8. R. Arredondo-Peter and R.V. Klucas *et al.*, Rice hemoglobins. Gene cloning, analysis, and O_2-binding kinetics of a recombinant protein synthesized in *Escherichia coli*, *Plant Physiol.* **115** (1997), pp. 1259–1266.
9. A.W. Sowa, S.M.G. Duff, M.G. Guy and R.D. Hill, Altering hemoglobin levels changes energy status in maize cells under hypoxia, *Proc. Natl Acad. Sci. USA* **95** (1998), pp. 10317–10321.
10. M.G.D. Duff, J.B. Wittenberg and R.D. Hill, Expression, purification, and properties of recombinant barley (*hordeum* sp.) hemoglobin, *J. Biol. Chem.* **272** (1997), pp. 16746–16752.
11. D.T. Mitchell, G.B. Kitto and M.L. Hackert, Structural analysis of monomeric hemichrome and dimeric cyanomet hemoglobins from *Caudina arenicola*, *J. Mol. Biol.* **251** (1995), pp. 421–431.
12. J. Janin, S. Miller and C. Chothia, Surface, subunit interfaces and interior of oligomeric proteins, *J. Mol. Biol.* **204** (1988), pp. 155–164.
13. J.B. Wittenberg, B.A. Wittenberg, Q.H. Gibson, M.J. Trinick and C.A. Appleby, The kinetics of the reactions of *Parasponia andersonii* hemoglobin with oxygen, carbon monoxide, and nitric oxide, *J. Biol. Chem.* **261** (1986), pp. 13624–13631.
14. M.S. Hargrove and G. Sarath *et al.*, Characterization of recombinant soybean leghemoglobin a and apolar distal histidine mutants, *J. Mol. Biol.* **266** (1997), pp. 1032–1042.
15. M. Bolognesi, D. Bordo, M. Rizzi, C. Tarricone and P. Ascenzi, Nonvertebrate hemoglobins: structural bases for reactivity, *Prog. Biophys. Mol. Biol.* **68** (1997), pp. 29–68.
16. R.E. Dickerson and I. Geis In: P. Hagopian, Editor, *Hemoglobin: Structure, Function, Evolution, and Pathology*, *Benjamin/Cummings Series in the Life Sciences*, Benjamin/Cummings Publishing Company, Menlo Park, CA (1983), pp. 65–112.
17. K.D. Egeberg, B.A. Springer, S.G. Sligar, T.E. Carver, R.J. Rohlfs and J.S. Olson, The role of Val68(E11) in ligand binding to sperm whale myoglobin. Site-directed mutagenesis of a synthetic gene, *J. Biol. Chem.* **265** (1990), pp. 11788–17795.
18. M.S. Hargrove, A.J. Wilkinson and J.S. Olson, Structural factors governing hemin dissociation from metmyoglobin, *Biochemistry* **35** (1996), pp. 11300–11309.
19. R.E. Brantley Jr., S.J. Smerdon, A.J. Wilkinson, E.W. Singleton and J.S. Olson, The mechanism of autooxidation of myoglobin, *J. Biol. Chem.* **268** (1993), pp. 6995–7010.
20. M.S. Hargrove, S. Krzywda, A.J. Wilkinson, Y. Dou, M. Ikeda-Saito and J.S. Olson, Stability of myoglobin: a model for the folding of heme proteins, *Biochemistry* **33** (1994), pp. 11767–11775.
21. E.H. Harutyunyan and J.C. Wilson *et al.*, The structure of deoxy- and oxyleghaemoglobin from lupin, *J. Mol. Biol.* **251** (1995), pp. 104–115.
22. D. Barrick, Replacement of the proximal ligand of sperm whale myoglobin with free imidazole in the mutant His93→Gly, *Biochemistry* **33** (1994), pp. 6546–6554.
23. G. Fermi, M.F. Perutz, B. Shaanan and R. Fourme, The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution, *J. Mol. Biol.* **175** (1984), pp. 159–174.
24. A. Pardanani, A. Gambacurta, F. Ascoli and W.E. Royer, Mutational destabilization of the critical interface water cluster in *Scapharca* dimeric hemoglobin: structural basis for altered allosteric activity, *J. Mol. Biol.* **248** (1998), pp. 729–739.
25. P.R. Kolatkar and R.P. Phizackerley *et al.*, Structure determination and refinement of homotetrameric hemoglobin from *Urechis caupo* at 2.5 Å resolution, *Acta Crystallogr. B* **48** (1992), pp. 191–199.
26. C. Tarricone, A. Galizzi, A. Coda, P. Ascenzi and M. Bolognesi, Unusual structure of the oxygen-binding site in the dimeric bacterial hemoglobin from *Vitreoscilla* sp, *Structure* **5** (1997), pp. 479–507.
27. S.J. Edelstein, M.J. Rehmar, J.S. Olson and Q.H. Gibson, Functional aspects of the subunit association–dissociation equilibria of hemoglobin, *J. Biol. Chem.* **245** (1970), pp. 4372–4381.
28. R.L. Galick, B.J. Miller and A.F. Riggs, The hemoglobins of *Phoronopsis viridis*, of the primitive invertebrate phylum Phoronida: characterization and subunit structure, *Arch. Biochem. Biophys.* **194** (1979), pp. 13–23.
29. R. Arredondo-Peter, M.S. Hargrove, J.F. Moran, G. Sarath and R.V. Klucas, Plant hemoglobins, *Plant Physiol.* **118** (1998), pp. 1121–1125.
30. H. Zhu and A.F. Riggs, Yeast flavohemoglobin is an ancient protein related to globins and a reductase family, *Proc. Natl Acad. Sci. USA* **89** (1992), pp. 5015–5019.
31. P.R. Gardner, A.M. Gardner, L.A. Martin and A.L. Salzman, Nitric oxide dioxygenase: an enzymic function for flavohemoglobin, *Proc. Natl Acad. Sci. USA* **95** (1998), pp. 10378–10383.
32. A.M. Gardner, L.A. Martin, P.R. Gardner, Y. Dou and J.S. Olson, 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** (2000), pp. 12581–12589.
33. Z. Otwinowski and W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* **276** (1997), pp. 307–326.
34. W. Furey and S. Swaminathan, PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules, *Methods Enzymol.* **277** (1997), pp. 307–326 Chapter 31.
35. T. Jones, J. Zou and M. Kjeldgaard, Improved methods for building protein models in electron-density maps and location of errors in these models, *Acta Crystallogr. A* **47** (1991), pp. 110–119.
36. J.S. Sack and F.A. Quiocho, CHAIN: a crystallographic modelling program, *Methods Enzymol.* **277** (1997), pp. 158–173.
37. A.T. Brünger, Crystallography & NMR system, *Acta Crystallogr. D* **54** (1998), pp. 901–921.
38. A.T. Brünger, X-PLOR version 3.1: a system for X-ray crystallography and NMR, Yale University Press, New Haven, CT (1992).
39. M Carson, Ribbons 2.0, *J. Appl. Crystallogr.* **24** (1991), pp. 961–985.
40. S. Evans, SETOR: hardware-lighted three-dimensional solid model representations of macromolecules, *J. Mol. Graph.* **11** (1993), pp. 134–138.
41. E.H Harutyunyan and W.J Wilson *et al.*, The binding of carbon monoxide and nitric oxide to leghaemoglobin in comparison with other haemoglobins, *J. Mol. Biol.* **264** (1996), pp. 152–161.
42. B.A. Springer, S.G. Sligar, S.J. Olson and G.N. Phillips Jr., Mechanisms of ligand recognition in myoglobin, *Chem. Rev.* **94** (1994), pp. 699–714.