Waterproofing the Heme Pocket

ROLE OF PROXIMAL AMINO ACID SIDE CHAINS IN PREVENTING HEMIN LOSS FROM MYOGLOBIN*

Received for publication, September 20, 2000, and in revised form, November 13, 2000
Published, JBC Papers in Press, November 17, 2000, DOI 10.1074/jbc.M008593200

Elaine C. Liong‡, Yi Dou, Emily E. Scott§, John S. Olson, and George N. Phillips, Jr.¶
From the Department of Biochemistry and Cell Biology and the W. M. Keck Center for Computational Biology,
Rice University, Houston, Texas 77005

The ability of myoglobin to bind oxygen reversibly depends critically on retention of the heme prosthetic group. Globin side chains at the Leu99(F4), His97(FG3), Ile99(FG5), and Leu104(G5) positions on the proximal side of the heme pocket strongly influence heme affinity. The roles of these amino acids in preventing heme loss have been examined by determining high resolution structures of 14 different mutants at these positions using x-ray crystallography. Leu99 and His97 are important surface amino acids that interact either sterically or electrostatically with the edges of the porphyrin ring. Ile99 and Leu104 are located in the interior region of the proximal pocket beneath ring C of the heme prosthetic group. The apolar amino acids Leu99, Ile99, and Leu104 “waterproof” the heme pocket by forming a barrier to solvent penetration, minimizing the size of the proximal cavity, and maintaining a hydrophobic environment. Substitutions with smaller or polar side chains at these positions result in exposure of the heme to solvent, the appearance of crystallographically defined water molecules in or near the proximal pocket, and large increases in the rate of hemin loss. Thus, the naturally occurring amino acid side chains at these positions serve to prevent hydration of the His97-Fe(III) bond and are highly conserved in all known myoglobins and hemoglobin.

Three important functions of the globin portion of myoglobin are to sequester heme, enhance coordination with the proximal histidine, and inhibit oxidation of the iron atom. In myoglobin, the heme pocket is surrounded by four of the eight globin helices. The B, C, and E helices form the top and sides of the porphyrin binding site, and the F helix forms the bottom (as oriented in Fig. 1). Heme binding causes the globin to form a more compact structure with an approximately 20% increase in helicity (1). Three edges of the porphyrin are buried in the protein interior and stabilized by hydrophobic interactions with apolar side chains that line the heme pocket (2–4). The fourth edge contains the solvent-exposed heme propionates, which interact electrostatically with polar amino acid side chains located on the surface of the protein.

The affinity of free heme for binding a single imidazole is very weak, whereas the affinity for binding a second base is very high. Thus, hexacoordinate bis-imidazole complexes are much more stable than the pentacoordinate intermediate (5–7). Although a much weaker ligand, water competes effectively against the binding of the first imidazole base in aqueous solution, causing the apparent equilibrium dissociation constant to be $K_d = 10^{-2} - 10^{-4}$ M (7). Hemo- and globin-binding imidazoles have evolved to stabilize the mono-imidazole heme complex to facilitate reversible O2 binding to the sixth coordination position. This stabilization appears to be accomplished by removing water molecules from the vicinity of the proximal coordination site.

The side chains of the proximal amino acids, Leu99(F4), His97(FG3), Ile99(FG5), and Leu104(G5), are within 4 Å of the porphyrin ring. Leu99 is located at the entrance of a hydrophobic cavity underneath pyrrole B. Xenon gas binds readily in this proximal pocket, which has been designated the Xe1 site (8). Photodissociated O2 and CO access both the distal Xe4 site and this Xe1 proximal site on nanosecond to microsecond time scales, suggesting that these spaces may be functionally significant (9–13, 34). His97 forms a hydrogen bond with the heme-7-propionate and is a hydrophobic barrier to solvation of the proximal heme pocket on the other side of the heme group. Ile99 and Leu104 are located just beneath the pyrrole rings B and C, respectively. Leu104 forms one of the internal sides of the Xe1 cavity. Hargrove et al. (3) suggested that both these amino acids may be important for keeping the internal region of the proximal pocket anhydrous and preventing disruption of the His97(F8)-Fe bond. To test these ideas, we have examined the effects of mutations at these four key positions on the crystal structure of recombinant sperm whale myoglobin.

MATERIALS AND METHODS

Preparation of Proteins—Recombinant wild-type and mutant sperm whale myoglobins were constructed, expressed, and purified as described by Springer and Sligar (15) and Carver et al. (16). Recombinant wild-type myoglobin differs from native myoglobin by the initiator methionine required for bacterial expression and an Asp122 to Asn mutation that was a result of an error in the original sequence determination (17) but do cause differences in crystallization conditions (18). These changes do not affect the function of the recombinant protein (17) but do cause differences in crystallization conditions (18). The protein samples were concentrated to about 1 mM in heme, frozen, and stored under liquid nitrogen until ready for use.

Measurement of Hemin Dissociation Rate Constants—Hemin dissociation rates were measured by monitoring the absorbance changes associated with the transfer of heme from holoprotein to excess H64Y/V68F apomyoglobin at pH 5.0 and 7.0, as described by Hargrove et al. (19). Briefly, the transfer of heme from 6 μM metmyoglobin to 30 μM H64Y/V68F apomyoglobin was measured by following the absorbance decrease at 409 nm as “green” H64Y/V68F holo-metmyoglobin was
tals grew in several weeks to sizes from 0.1 to 0.5 mm. The Ile99 to Ser mutations for which heme loss rates were given and no structure is presented were not screened for crystallization (Leu89 to Ala and Ser and His97 to Ala). The other mutations failed to crystallize with these conditions. The other mutant crystals were eluted using a linear gradient of 0.0–1.0 M sodium chloride in 20 mM sodium phosphate, 1 mM EDTA, pH 6.5, at 4 °C. The dialyzed samples were loaded onto a preparative weak cation exchange column (Custom LC) equilibrated with the dialysis buffer. The samples were eluted using a linear gradient of 0.0–1.0 M sucrose at pH 7. The values for the rate of heme loss for most of the mutants were originally reported in Hargrove et al. (3); the remainder were measured for this report (20).

Preparation of Protein Crystals—Prior to crystallization, recombinant proteins were purified further by HPLC1 (LDC/Milton Roy). Samples were dialyzed against 20 mM sodium phosphate, 1 mM EDTA, pH 8.0–9.0, with well solutions having concentrations of 2.4–3.0 M sucrose at 37 °C in either sodium acetate at pH 5.0 or potassium phosphate at pH 7. The values for the rate of hemin loss from the wild type kH values at pH 7 is due to the slowness of the reaction that requires collecting time courses for over 24 h. Much less variability is observed for those mutants with higher rates of heme loss. Differences were observed in electron density maps. The mutated side chains were changed only after \(2F_o - F_c\) and \(|F_o - F_c|\) difference Fourier maps indicated clear positions for these altered side chains. Peaks in the difference Fourier maps were modeled as water molecules using the program PEAK2 or OLE2.0. After the refinement converged in X-PLOR, the model was refined further using the conjugate gradient and full matrix least squares method in SHELXL (25). Estimated standard deviations of the atomic positions of the heme iron and the 24 atoms of the pyrrole rings and methionine bridges were calculated from refinement without the imposition of standard distances and planarity restraints, as described in Ref. 25.

Sequence Alignment—All myoglobin sequences were downloaded from the National Center of Biotechnology Information. About 460 sequences were obtained through the Entrez Protein search engine, which included all myoglobin related data from SwissProt, Protein Information Resource, Protein Data Bank, and Protein Research Foundation entries, etc. Incomplete myoglobin sequence fragments were immediately discarded. Most of the myoglobin sequences from the Protein Data Bank are from known myoglobins, and their mutants were also omitted in the sequence alignment. The first sequence alignment containing 203 sequences was performed using Clustal W (version 1.8) (26). Duplicate sequences with different entry names were recognized and sorted in a log file, and extra copies of duplicate entries were discarded. The final multiple alignment contains 90 myoglobin sequences, and the nomenclature for helical position was taken from the Protein Data Bank. Although the focus of this work is a structural interpretation of the effects of proximal pocket mutations on heme loss (Table I), it is important to note that some of these amino acid replacements do have significant effects on ligand binding. Table II presents a list of the \(O_2\) binding properties of all 14 mutants whose structures have been determined by x-ray crystallography. There is a progres-

| Myoglobin     | \(k_{\text{H}, \text{pH 5}} \) | \(k_{\text{H}, \text{pH 7}} \) |
|---------------|-----------------|-----------------|
| Wild type     | 1.0 ± 0.2       | 0.05 ± 0.04     |
| L89G          | 700             | 140             |
| L89A          | 56              | 5.7             |
| L89S          | ~500            | 15              |
| L89F          | 2.5             | 0.1             |
| L89W          | 17              | 0.4             |
| H97A          | 39              | 2.0             |
| H97V          | 23              | 0.8             |
| H97F          | 1.5             | 0.1             |
| H97Q          | 15              | 0.70            |
| H97D          | 38              | 5               |
| H99A          | 8.0             | 0.2             |
| H99V          | 2.0             | 0.7             |
| H99S          | 41.0            | 0.9             |
| L104A         | 19.2            | 0.7             |
| L104V         | 5.4             | 0.1             |
| L104F         | 6.6             | 0.2             |
| L104W         | 4.9             | 0.1             |
| L104N         | 1.7             | 0.08            |
| H93G          | 660             | 140             |

1 The abbreviation used is: HPLC, high pressure liquid chromatography.

2 M. L. Quillin, unpublished data.

3 M. B. Berry, unpublished data.
Waterproofing the Heme

Table II
Overall rate and equilibrium constants for O₂ binding to proximal mutants of sperm whale myoglobin

| Myoglobin | k₁₀² | k₂₀² | K₀₂ |
|-----------|------|------|-----|
|           | μm⁻¹ s⁻¹ | μm⁻¹ | s⁻¹ |
| Wild type | 16 ± 3 | 14 ± 3 | 1.1 ± 0.2 |
| L89G      | 13    | 9.7  | 1.3 |
| L89F      | 20    | 31   | 0.64 |
| L89W      | 20²   | 71   | 0.28 |
| L97V      | 14    | 12   | 1.1 |
| L97F      | 18    | 12   | 1.5 |
| L97Q      | 16    | 34   | 0.48 |
| H97D      | 18    | 9.5  | 1.9 |
| I99A      | 13    | 6.3  | 2.1 |
| I99V      | 18    | 8.8  | 2.1 |
| L104A     | 25    | 8.0  | 3.1 |
| L104V     | 18    | 9.0  | 2.0 |
| L104F     | 20    | 6.2  | 3.3 |
| L104W     | 16    | 2.8  | 5.8 |
| L104N     | 21    | 3.5  | 0.6 |

The rates of hemin loss from the His97 mutants show both a size and a charge dependence with the order of kₘ_H: L89G > L89S > L89A > L89W > L89F > native/wild type. The hydration and increase of polarity in the heme pocket of the H93G mutant shows packing against the heme forming a significant part of the outer covering of the proximal portion of the heme pocket. The His93(F8) is hidden in a hydrophobic cavity that is protected from the solvent molecules located on the surface of the protein. Replacing leucine at position 89 with glycine creates a clear pathway for the entry of solvent into the heme pocket (Fig. 3B). Electron density indicative of at least three internal water molecules is seen in the [Fe₃⁺ ₉W₃] omit map for L89G metmyoglobin (Fig. 2). The three internal water molecules are observed inside the Xe1 pocket of this mutant and are connected to four other water molecules located on the protein exterior (Fig. 3B).
Waterproofing the Heme

TABLE III
Statistics for data collection and refinement

| Protein  | Protein Data Bank file | a ( Å ) | b ( Å ) | c ( Å ) | d_{min} ( Å ) | R_{merge} | R_{cryst} | R_{free} |
|----------|------------------------|---------|---------|---------|---------------|-----------|-----------|---------|
| L89G     | 1CH1                   | 91.46   | 91.46   | 45.98   | 1.9           | 39596     | 16894     | 19.2    |
| L89F     | 1CH2                   | 91.43   | 91.43   | 45.91   | 1.8           | 139450    | 20451     | 18.0    |
| L89W     | 1CH3                   | 91.72   | 91.72   | 46.02   | 2.0           | 76505     | 15027     | 19.2    |
| H97V     | 1CH5                   | 91.70   | 91.70   | 46.05   | 2.1           | 70094     | 13521     | 20.6    |
| H97F     | 1CH7                   | 91.25   | 91.25   | 45.84   | 1.9           | 66812     | 15177     | 19.3    |
| H97Q     | 1CH9                   | 91.60   | 91.60   | 46.00   | 1.8           | 36034     | 19568     | 18.5    |
| H97D     | 1D7T                   | 49.16   | 49.00   | 80.01   | 1.7           | 66693     | 17917     | 22.7    |
| I99A     | 1C1K                   | 91.50   | 91.50   | 45.94   | 1.7           | 67888     | 22498     | 18.6    |
| I99V     | 1CIO                   | 91.61   | 91.61   | 46.02   | 1.6           | 90855     | 26662     | 18.1    |
| L104A    | 1C08                   | 91.54   | 91.54   | 45.89   | 1.8           | 86563     | 19086     | 21.0    |
| L104V    | 1C09                   | 91.55   | 91.55   | 46.02   | 1.6           | 108911    | 28207     | 18.7    |
| L104F    | 1C05                   | 91.60   | 91.60   | 45.90   | 2.1           | 66964     | 12814     | 19.0    |
| L104W    | 1C0W                   | 91.59   | 91.59   | 45.91   | 2.2           | 38133     | 10972     | 19.7    |
| L104N    | 1CPO                   | 91.47   | 91.47   | 45.90   | 2.0           | 60605     | 14297     | 18.6    |

TABLE IV
Statistics from refinement using SHELXL for the globin and porphyrin ring geometries

| Protein  | Bond length ( Å ) | Bond angles ( deg ) | Dihedrals ( deg ) | Improper ( deg ) |
|----------|-------------------|---------------------|-------------------|------------------|
| L89G     | 0.010             | 2.0                 | 21                | 1.4              |
| L89F     | 0.010             | 2.1                 | 20                | 1.5              |
| L89W     | 0.009             | 2.0                 | 21                | 1.6              |
| H97V     | 0.016             | 2.3                 | 23                | 1.6              |
| H97F     | 0.008             | 2.1                 | 21                | 1.5              |
| H97Q     | 0.012             | 2.1                 | 21                | 1.5              |
| H97D     | 0.010             | 2.1                 | 22                | 1.4              |
| I99A     | 0.011             | 2.1                 | 21                | 1.5              |
| I99V     | 0.016             | 2.2                 | 20                | 1.6              |
| L104A    | 0.009             | 2.1                 | 21                | 1.5              |
| L104V    | 0.012             | 2.2                 | 21                | 1.6              |
| L104F    | 0.011             | 1.9                 | 21                | 1.5              |
| L104W    | 0.008             | 2.1                 | 22                | 1.5              |
| L104N    | 0.009             | 2.0                 | 21                | 1.6              |

TABLE V
Bond lengths for proximal pocket metmyoglobin mutants

| Protein  | Iron displacement from plane ( Å ) | Fe-His^{97}Nε2 ( Å ) |
|----------|-----------------------------------|----------------------|
| Wild type| -0.16 ± 0.02                       | 2.19 ± 0.02          |
| Native   | 0.14                              | 2.14                 |
| L89G     | -0.11 ± 0.01                      | 2.15 ± 0.03          |
| L89F     | -0.16 ± 0.01                      | 2.13 ± 0.03          |
| L89W     | -0.20 ± 0.02                      | 2.16 ± 0.04          |
| H97V     | -0.09 ± 0.03                      | 2.15 ± 0.05          |
| H97F     | -0.12 ± 0.01                      | 2.16 ± 0.03          |
| H97Q     | -0.07 ± 0.01                      | 2.15 ± 0.03          |
| H97D     | -0.11 ± 0.01                      | 2.14 ± 0.03          |
| I99A     | -0.08 ± 0.01                      | 2.14 ± 0.03          |
| I99V     | -0.12 ± 0.01                      | 2.12 ± 0.02          |
| L104A    | -0.10 ± 0.01                      | 2.14 ± 0.03          |
| L104V    | -0.01 ± 0.01                      | 2.13 ± 0.02          |
| L104F    | -0.11 ± 0.01                      | 2.14 ± 0.03          |
| L104W    | -0.10 ± 0.02                      | 2.15 ± 0.05          |
| L104N    | -0.09 ± 0.02                      | 2.16 ± 0.04          |

where there is no covalent attachment between the heme and the globin. A similar dramatic increase in the rate of hemin loss is observed for the L89S mutation, where the small hydroxymethyl side chain probably also allows an influx of water and increases the polarity of the heme pocket. The L89A mutation also causes a substantial increase in k_{-1}, but the effect is significantly smaller than that seen for the L89G and L89S mutants.

Mutation of Leu^{89} to an amino acid with similar size and polarity, such as L89F, has a much smaller effect on the rate of hemin loss, particularly at pH 5. As shown in Fig. 3C, the aromatic side chain of Phe^{89} metmyoglobin forms a hydrophobic barrier to solvation which is equivalent to that of the native leucine side chain. Trp^{89} also provides a steric barrier to solvation (Fig. 3D). However, the L89W mutation causes a 17-fold increase in the rate of hemin loss at pH 5, which is most likely due to both steric hindrance between the large indole side chain and the rigid porphyrin ring and to the increased polarity of the indole side chain.

His^{97} Mutants—The location of His^{97} (FG3) is shown in Fig. 4A where the protein molecule has been rotated 90° to the right relative to its position in Fig. 1. Replacing His^{97} with small (H97A and H97V) or negatively charged (H97D and H97E) side chains results in large increases in k_{-1}. The crystal structures of H97V and H97D show that the heme-7-propionate is no longer interacting with this side chain and that a channel has been created that connects the solvent exterior to the proximal histidine (Fig. 4, B and E, respectively). In both mutants, hydration of the heme pocket is significantly smaller than that seen for the L89G mutation, which almost completely exposes the proximal histidine to water. The smaller channel and extent of hydration explain why the effects of H97V and H97D mutations on hemin loss are less than those observed for the L89G, A, and S replacements.

Replacing His^{97} with glutamine was expected to have little effect on the metmyoglobin structure because the amide side chain can still form a hydrogen bond with the heme-7-propionate. However, as shown in Fig. 4D, the glutamine side chain is not large enough to completely seal the opening to the hydrophobic cavity, resulting in significant increases in k_{-1} at both neutral and acid pH (Table I). Phenylalanine can replace His^{97} with very little effect on the rate of hemin loss at either pH. (Table I). In the crystal structure of this mutant, the phenyl ring is located in the same position as the naturally occurring imidazole ring (Fig. 4C). Phe^{97} also appears to provide some electrostatic stabilization through interaction between the positive edge of the phenyl multipole and the heme-7-propionate. Thus, the native His^{97} side chain provides both a
steric barrier to solvation and electrostatic stabilization of one of the heme propionates.

Ile99 Mutants—The side chain at the FG5 position is important in maintaining the exact position of the porphyrin ring in the heme pocket. Replacing the native isoleucine with a smaller side chain results in tilting of the B ring toward the proximal side of the heme pocket because of the loss of the C6 atom of the sec-butyl side chain (Fig. 5). Hargrove et al. (3) suggested that an apolar side chain at position 99 is important for keeping the heme pocket anhydrous and observed that the I99S mutation results in a 90-fold increase in hemin loss at pH 5. Mutation to alanine results in a more moderate 20-fold increase. Other than changes in the tilt of porphyrin ring, the crystal structures of I99A and I99V do not show significant alteration of the protein or hydration of the heme pocket with well defined internal water molecules. Thus, the causes of the high rates of hemin dissociation from the I99A, I99V, and I99S mutants are more ambiguous and could be due to the increase in internal space and/or changes in the tilt of the heme which weaken the Fe-His93 bond.

Leu104 Mutants—Replacing Leu104(G5) with alanine and valine causes moderate to small increases in the rate of hemin loss (from 20- to 5-fold, respectively). Ordered water molecules are found in the proximal portion of the heme pocket in the L104A and L104N mutants (Fig. 6A) and almost certainly account for the increases in k_{11} for these proteins. The presence of these interior water molecules was unexpected. There is no obvious connection to the solvent phase because Leu89 still blocks direct access to solvent, making visualization of the

![Fig. 2. Stereoview of the electron density for internal water molecules in the model of L89G sperm whale metmyoglobin. The |Fo - Fc| electron density map was generated by refining the structure of side chain without a side chain at position 89 and without including any internal water molecules. The observed density indicates at least three mobile water molecules, and as many as four could be present and penetrate well into the proximal pocket.](image-url)

![Fig. 3. Close-up views of the proximal heme pocket for 89(F4) substitutions in sperm whale myoglobin. The heme, proximal histidine (His93), Leu104(G5), and position 89(F4) side chains are highlighted in stick representation. Water molecules are shown in space-filling representation in dark gray. Upper left, wild type; upper right, L89G; lower left, L89F; lower right, L89W.](image-url)

Waterproofing the Heme

Leu104 Mutants—Replacing Leu104(G5) with alanine and valine causes moderate to small increases in the rate of hemin loss (from 20- to 5-fold, respectively). Ordered water molecules are found in the proximal portion of the heme pocket in the L104A and L104N mutants (Fig. 6A) and almost certainly account for the increases in k_{11} for these proteins. The presence of these interior water molecules was unexpected. There is no obvious connection to the solvent phase because Leu89 still blocks direct access to solvent, making visualization of the
internal waters difficult (Fig. 6A).

There are two possible pathways for entry of the water molecules seen in the L104A and L104N mutants. Solvent could enter the heme pocket through the distal histidine gate and then migrate to the proximal Xe1 site as is seen for the apolar ligands, O₂ and CO (9–13, 34). Alternatively, water molecules could diffuse into the protein via multiple transient channels created by thermal fluctuations in the protein matrix (32). There are two well defined surface water molecules within hydrogen bonding distance of Gln¹⁰⁵(G6), and when Leu¹⁰⁴ is replaced with Asn, a chain of three water molecules is found in this area and extends from the exterior surface of the amide side chain out into the solvent phase along the G helix.

In general, water molecules are not seen in the proximal and distal cavities of myoglobin because these spaces are lined with hydrophobic amino acids. However, in the L104A mutation, the proximal cavity is probably large enough to sequester two internal water molecules that appear to hydrate the interior edge of the porphyrin ring and the Xe1 pocket. The net result is a ~20-fold increase in the rate of heme dissociation.

The amide side chain in the L104N mutation is highly polar. Two internal water molecules are attached to this side chain in the crystal structure of the metmyoglobin mutant. Remarkably, this mutation causes little increase in the rate of heme loss at either pH 5 or 7. The water molecules in the L104N mutant appear to be attached strongly to each other and to the amide side chain by strong hydrogen bonds with O–O and N–O distances of 2.6 to 3.1 Å (Fig. 6B). Thus, the hydroscopic nature of the Asn side chain appears to prevent these internal water molecules from solvating the heme group and disrupting the proximal histidine-iron bond. In contrast, the more mobile water molecules present in the L104A mutant do facilitate disruption of the Fe-His(F8) bond (Table I).

DISCUSSION

The four proximal pocket amino acids examined appear to “waterproof” the heme pocket. They maintain a favorable apolar environment for the heme group and prevent the hydration of the proximal imidazole base. Leu⁸⁹(F4) seals the hydrophobic cavity underneath pyrrole B, preventing solvation of the Xe1 site in the proximal portion of the heme pocket. Decreasing the size or increasing the polarity of the side chain at position 89 results in dramatic increases in the rate of heme loss. There appears to be strong selective pressure to seal off the heme pocket as judged by the high degree of conservation at the F4 amino acid. The frequency of leucine at this position is 87% in the 99 known myoglobin sequences, and the remaining major variation is conservative (7% Phe). Leu(F4) is even more conserved in vertebrate hemoglobins showing a frequency of 99% in both the α (279 sequences) and β (224 sequences) subunits.

His⁹⁷(FG3) is part of the extensive hydrogen-bonding network found at the exterior portion of the proximal pocket of myoglobin. Disruption of the hydrogen-bonding network by mutation of His⁹⁷ to smaller, apolar, or negatively charged side chains has little effect on the ligand binding properties of the protein (10, 20, 34). However, loss of favorable electrostatic interactions with heme-7-propionate or opening of a channel to the proximal histidine does accelerate the rate of hemin loss up to 40-fold.

Both Ile⁹⁹ and Leu¹⁰⁴ are located in the interior region of the heme pocket beneath the heme prosthetic group and are important in maintaining the position of the porphyrin ring, which in turn influences the reactivity of the heme iron toward ligands (10, 20, 34). Mutations at both positions have significant effects on the rate of heme dissociation. Substitution with an alanine at position 104 enlarges the proximal pocket enough to allow the entry of water molecules that promote heme loss. In the L104N mutant, solvent also enters the proximal cavity, but the water molecules are sequestered by the Asn side chain. In effect, the amide group acts as a hydroscopic agent to keep the Fe-His⁺ “dry,” preventing a large increase in the rate of heme dissociation.

All of these results show that the proximal environment of the heme group must be kept anhydrous and that direct exposure to solvent must be minimized. Tilton et al. (8) have shown that there are internal cavities at both the proximal and distal sides of the heme group and visualized them by adding Xe to crystals and determining the binding sites. The most stable Xe site(Xe1) is circumscribed by Leu⁸⁹, His⁹³, Leu¹⁰⁴, and Phe¹³⁸. In the native protein, this cavity is too small and hydrophobic to accommodate water molecules. A small fraction of the apolar ligands O₂ and CO do reach this location on ns to μs time scales after laser photolysis or under high pressure (9–13, 34). However, if the Xe1 cavity is opened to solvent (L89G), enlarged (L104A), or made more polar (L104N), water does enter and destabilize the proximal His⁹³-Fe bond.

The idea that solvation of the heme pocket facilitates heme loss was first proposed to account for the instability of Hemo-globin Boras (β Leu(F4) → Arg), which exhibits abnormally high rates of heme loss, formation of semihemoglobin, and precipitation (14, 27, 33). The structures presented here show directly and convincingly that Leu⁸⁹ does protect the proximal pocket from hydration and subsequent loss of the heme pros-
thetic group. Clearly, there is strong selective pressure to prevent hydration of the heme pocket, as can be seen by the high degree of conservation of large apolar side chains at the Leu$^{89}$(F4), Ile$^{99}$(FG5), and Leu$^{104}$(G5) positions in myoglobin. At the FG5 position, the frequencies are 83% Ile, 12% Val, 3% Leu, and 2% others in the 99 known myoglobin sequence; those at the G5 position are 78% Leu, 17% Phe, and 5% other. In contrast, there is much more variability at the His$^{97}$(FG3) position. In most cases, large side chains are present at FG3 to seal off the proximal pocket, and when polar side chains are present, they interact with the heme propionates.

The results described in this report confirm experimentally the necessity of a hydrophobic pocket to bind and retain the heme prosthetic group in myoglobin and show directly by crystallography the presence of water molecules around the heme when this hydrophobic pocket is disrupted. Furthermore, the character and contributions of particular proximal side chains on heme binding and retention have been dissected.

Acknowledgments—We thank Eileen Singleton for construction, expression, and purification of the mutant proteins, Bog Stec for help with the figures, and Mark Hargrove for encouraging determination of the structures of these mutants by x-ray crystallography.

REFERENCES
1. Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., and Kutysenko, V. P. (1988) J. Mol. Biol. 202, 127–138
2. Hargrove, M. S., and Olson, J. S. (1996) Biochemistry 35, 11310–11318
3. Hargrove, M. S., Wilkinson, A. J., and Olson, J. S. (1996) Biochemistry 35, 11300–11309
4. Hargrove, M. S., Barrick, D., and Olson, J. S. (1996) Biochemistry 35, 11293–11299
5. Cole, S. J., Curthoys, G. C., Cole, and Magnusson, E. A. J. (1970) J. Am. Chem. Soc. 92, 2991–2996
6. Cole, S. J., Curthoys, G. C., Cole, and Magnusson, E. A. J. (1971) J. Am. Chem. Soc. 92, 2153–2158
7. Brault, D., and Rougee, M. (1974) Biochemistry 13, 4591–4597
8. Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) Biochemistry 23, 2849–2857
9. Scott, E. E., and Gibson, Q. H. (1997) Biochemistry 36, 11909–11917
10. Scott, E. E. (1998) Apoglobin Stability and Ligand Movements in Mammalian Myoglobins. Ph.D. Thesis, Rice University, Houston, TX
11. Chu, K., Vojtechovsky, J., McMahon, B. H., Sweet, R. M., Berendzen, J., and Schlichting, I. (2000) Nature 403, 921–923
12. Brunori, M., Vallezzi, E., Cutruzzola, F., Travaglini-Allocatelli, C., Berendzen, J., Chu, K., Sweet, R. M., and Schlichting, I. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 2058–2063
13. Ostermann, A., Waschiky, R., Parak, F. G., and Nienhaus, G. U. (2000) Nature 404, 205–208
14. Bunn, H. F., and Forget, B. G. (1986) *Hemoglobin: Molecular, Genetic, and Clinical Aspects*, W. B. Saunders, Philadelphia, PA
15. Springer, B. A., and Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 8961–8965
16. Carver, T. E., Brantley, R. E., Jr., Singleton, E. W., Arduini, R. M., Quillin, M. L., Phillips, G. N., Jr., and Olson, J. S. (1992) *J. Biol. Chem.* 267, 14443–14450
17. Springer, B. A., Sligar, S. G., Olson, J. S., and Phillips, G. N., Jr. (1994) *Chem. Rev.* 94, 699–714
18. Phillips, G. N., Jr., Arduini, R. M., Springer, B. A., and Sligar, S. G. (1990) *Proteins* 7, 358–365
19. Hargrove, M. S., Singleton, E. W., Quillin, M. L., Ortiz, L. A., Phillips, G. N., Jr., Olson, J. S., and Mathews, A. J. (1994) *J. Biol. Chem.* 269, 4207–4214
20. Liong, C. E. (1999) *Structural and Functional Analysis of Proximal Pocket Mutants of Sperm Whale Myoglobin*. Ph.D. Thesis, Rice University, Houston, TX
21. Kabsch, W. (1988) *J. Appl. Crystallogr.* 21, 916–924
22. Brünger, A. T. (1992) X-PLOR, version 3.1, Yale University Press, New Haven, CT
23. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr.* A 47, 392–400
24. Sack, J. S. (1988) *J. Mol. Graphics* 6, 244–245
25. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* 276, 319–343
26. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* 22, 4673–4680
27. Dickerson, R. E., and Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution, and Pathology*. Benjamin/Cumings Series in the Life Sciences, Benjamin/Cumings Publishing Company, Menlo Park
28. Dou, Y. (1997) *Protein Engineering Reveals Specific Roles of Amino Acid Residues in Function and Stability of Myoglobin*. Ph.D. Thesis, Case Western Reserve University, Cleveland, OH
29. Yang, F., and Phillips, G. N., Jr. (1996) *J. Mol. Biol.* 256, 762–774
30. Barrick, D. (1994) *Biochemistry* 33, 6546–6554
31. Quillin, M. L., Li, T., Olson, J. S., Phillips, G. N., Jr., Dou, Y., Ikeda-Saito, M., Regan, R., Carlson, M., Gibson, Q. H., Li, H., and Elber, R. (1995) *J. Mol. Biol.* 245, 416–436
32. Benson, E. S., Rossi Panelli, M. R., Giacometti, G. M., Rosenberg, A., and Antonini, E. (1973) *Biochemistry* 12, 2699–2706
33. Antonini, E., and Brunori, M. (1971) *Hemoglobin and Myoglobin and Their Reaction with Ligands*. North Holland Publishing Co., Amsterdam
34. Scott, E. E., Gibson, Q. H., and Olson, J. S. (2001) *J. Biol. Chem.* 276, 5177–5188