Solution and Crystal Structures of a Sperm Whale Myoglobin Triple Mutant That Mimics the Sulfide-binding Hemoglobin from Lucina pectinata*

(Received for publication, August 25, 1997, and in revised form, February 6, 1998)

Bao D. Nguyen, Xuefeng Zhao, Krishnamurthi Vyas, and Gerd N. La Mar‡
From the Department of Chemistry, University of California, Davis, California 95616

R. Ashley Lile, Eric Allen Brucker§, George N. Phillips, Jr., and John S. Olson
From the Department of Biochemistry and Cell Biology and the W. M. Keck Center for Computational Biology, Rice University, Houston, Texas 77005-1892

Jonathan B. Wittenberg
From the Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York 10461

The bivalve mollusc Lucina pectinata harbors sulfide-oxidizing chemosautrophic bacteria and expresses a monomeric hemoglobin I, HbI, with normal O₂, but extraordinarily high sulfide affinity. The crystal structure of aquomet Lucina HbI has revealed an active site with three residues not commonly found in vertebrate globins: Phe(B10), Gln(E7), and Phe(E11) (Rizzi, M., Wittenberg, J. B., Coda, A., Fasano, M., Ascenzi, P., and Bolognesi, M. (1994) J. Mol. Biol. 244, 86–89). Engineering these three residues into sperm whale myoglobin results in a triple mutant with 700- to 1000-fold higher sulfide affinity than for wild-type. The single crystal x-ray structure of the aquomet derivative of the myoglobin triple mutant and the solution ¹H NMR active site structures of the cyanomet derivatives of both the myoglobin mutant and Lucina HbI have been determined to examine further the structural origin of their unusually high sulfide affinities. The major differences in the distal pocket is that in the aquomet form the carbonyl of Gln⁶⁴(E7) serves as a H-bond acceptor, whereas in the cyanomet form the amido group acts as H-bond donor to the bound ligand. Phe⁶⁸(E11) is rotated ∼90° about χ₂ and located ∼1–2 Å closer to the iron atom in the myoglobin triple mutant relative to its conformation in Lucina HbI. The change in orientation potentially eliminates the stabilizing interaction with sulfide and, together with the decrease in size of the distal pocket, accounts for the 7-fold lower sulfide affinity of the myoglobin mutant compared with that of Lucina HbI.

Despite substantial amino acid sequence variability, the reversible binding of molecular oxygen in myoglobin (Mb) and hemoglobin (Hb) is achieved by a surprisingly invariant protein folding topology: a heme group imbedded within 7–8 packed α-helices (1, 2). Examination of the more than 300 known sequences of vertebrate Hbs and Mbs and the >130 nonvertebrate globin-like sequences demonstrates that only two residues appear to be conserved, the proximal His at helical position F-8 and the Phe parallel to the heme surface at an interhelical loop, position CD1 (1, 3). The residue considered to be pivotal in the stabilization of the bound O₂ is the one whose side chain can provide a neutral donor for a hydrogen bond with the ligand, such as His or Gln at the distal position E-7. Among vertebrate Mbs and Hbs, His predominates, with Gln occurring only in elephant Mb (4), shark Mb (5), hagfish Hb (6), and the α chain of opossum Hb (7). Although His is also found in the majority of nonvertebrate Hbs and Mbs, Gln is not uncommon and, in addition, other residues may occupy the E-7 position, e.g. Val in Aplysia (8) and several other mollusc Mbs (9), Tyr in the Mb from Paramphistomum reptile (10, 11), and Leu in the monomeric components of Glycera Hb (12). Although capable of O₂ binding, a number of the variant globins exhibit some anomalous functional properties. Parallel functional and structural studies on genetically engineered mutants of mammalian Mb have shown that changes in only a very limited number of residues in the distal pocket often transfers the essence of the unusual behavior of the natural genetic variant to the appropriate reference protein mutant (13–17).

Cytoplasmic Hbs occur in many symbiotic associations between bacteria and invertebrates or plants and may have a physiological role in the symbiosis (18–20). The bivalve mollusc Lucina pectinata, found in sulfide-rich coastal sediments, harbors sulfide-oxidizing chemosautrophic bacteria (21, 22). Its abundant cytoplasmic hemoglobins consist of three single-chain components, each having a moderately high affinity for oxygen (P₅₀ = 0.1–0.2 torr), which is achieved, however, by a very different balance of combination and dissociation rates. Reactions of the monomeric HbI with oxygen are rapid, whereas those of HbII and HbIII are extraordinarily slow (23).

* This work was supported by United States National Institutes of Health Grants HL16087 (to G. N. L.), GM35649 and HL47020 (to J. S. O.), AR40252 (to G. N. P.), and Postdoctoral Fellowship AR08355 (to E. A. B.), the States of Texas Advanced Technology Program Grant 003604-025 (to G. N. P. and J. S. O.), Robert A. Welch Foundation Grants C-412 (to J. S. O.) and C-1142 (to G. N. P.), and the W. M. Keck Center for Computational Biology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Current address: Somatogen, Inc., Boulder, CO 80301.

§ To whom correspondence should be addressed.

The abbreviations used are: Mb, myoglobin; Hb, hemoglobin; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; MbCO, carbonmonoxymyoglobin; HbI, monomeric hemoglobin I; metMbCN, cyanomet myoglobin; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional total correlation spectroscopy; ppm, parts per million; TOCSY, two-dimensional nuclear Overhauser spectroscopy; V₅₈F-Mb, Val⁵⁸(E10)→Phe-Mb; L₂₉F/H₆₄Q/V₅₈F-Mb, Leu²⁹(B₁₀)→Phe, His⁶⁸(E⁷)→Gln, Val⁵⁸(E¹¹)→Phe-Mb; HbII, monomeric hemoglobin component II; metHbICN, cyanomet-hemoglobin I.
Recent crystal structures of *Lucina* aquamortHbI and the sulfide bound form (24, 25) have revealed a distal Gln^{64}(E7) and Phe^{20}(B10), neither of which is rare among invertebrate Mbs, Hbs. However, the E-11 position is uniquely occupied by Phe rather than the aliphatic amino acids Val, Ile, or Leu. To examine the structural origin of strong sulfide binding in *Lucina* HBI, we have systematically replaced His^{64}(E7), Leu^{29}(B10), and Val^{64}(E11) with Gln, Phe, and Phe, respectively, in sperm whale myoglobin and then determined the structures of the resulting multiple mutants by NMR and x-ray crystallography. This approach is similar to that used to study the unusual functional properties of *Aplysia* Mb, *Ascaris* Hb, and elephant Mb, and provides comparison between the native protein and the synthetic mimic in two ligation/spin states and between protein structures in crystal and solution (13–17, 26, 27). The NMR studies were pursued to define more carefully the positions of the labile protons of Gln^{64}(E7) and their role as hydrogen bond donors to the bound ligand. Both the paramagnetic relaxation and induced dipolar shifts in metMbCN and metHbICN can be used to model the position of distal residues (14–17, 26, 29, 30). The combined NMR and crystallographic results reveal systematic differences between the orientation of Gln^{64}(E7) in the aquomet and cyanomet complexes of *Lucina* Hbl and L29F/H64Q/V68F-Mb, which demonstrate that Gln^{64}(E7) may serve as either a H-bond acceptor or donor to bound ligands and that a difference in the orientation of the Phe^{68}(E11) ring may account for the difference in sulfide affinity between *Lucina* Hbl and the sperm whale triple mutant Mb mimic.

**EXPERIMENTAL PROCEDURES**

**Protein Preparation**—The pUC19 plasmid containing V68F sperm whale myoglobin was made by Egeberg et al. (31). The plasmid containing L29F/H64Q/V68F sperm whale Mb was constructed from pEMBL19 containing L29F Mb using cassette mutagenesis as described in Springer et al. (32). Vectors harboring the mutated gene were transformed into *Escherichia coli* strain TB1 and expressed constitutively using a 100-liter fermentor. The resulting soluble holomyoglobin protein was purified as described previously (33, 34).

**Protein Crystallization**—Crystals of the recombinant L29F/H64Q/V68F-Mb were grown by the hanging drop method (35, 36) in a temperature-controlled environment (17 °C). Large single crystals developed out of solutions ranging from 2.4 to 2.7 M ammonium sulfate buffer at pH 9 (20 mM Tris-HCl, 1 mM EDTA). 1H NMR Measurements—All the 1H NMR spectra were collected on the GE Omega 500 MHz spectrometer. The strongly relaxed signals were optimally detected in water-eliminated Fourier transform spectra (42). Nonselective T_{1a} for the resolved strongly relaxed protons were measured via inversion-recovery experiment. Steady state NOEs were recorded as described previously (43). The phase-sensitive TOCSY (44, 45), NOESY (46), and conventional magnitude COSY (MCOSY) (47) employed the method described by States et al. (48) to provide quadrature detection in the t_{1} dimension. Solvent suppression, when required, was achieved by direct saturation in the relaxation delay period. 512 blocks were collected with 25.0 kHz spectral widths to include all resonances, and 10 kHz to improve resolution for the diamagnetic envelope. 128 to 256 scans were accumulated with repetition rate of 0.7 s^{-1} or 12 s^{-1} for each block with free induction decays of 2048 complex points. The data were processed as described previously (49); details are given in the figure captions. All two-dimensional data were processed on Silicon Graphics workstation using the software package Felix from Biosym/MSI (San Diego).

**RESULTS**

**Hydrogen Sulfide and O_{2} Binding to Lucina Hbl and Sperm Whale Mb**

A brief summary of the affinities of the myoglobin mutants and *Lucina* Hbl for O_{2} and hydrogen sulfide^{2} is shown in Table 1 (17, 23, 28). *Lucina* Hbl shows a 5,000-fold higher affinity for...
The three mutated side chains in the recombinant myoglobin do not alter the overall tertiary structure of myoglobin nor do they appear to sterically hinder the bound ligand. Two of these mutated residues, Glu<sup>64</sup>(E7) and Phe<sup>68</sup>(E11), occupy positions equivalent to those observed in the structures of corresponding H64Q-Mb and V68F-Mb single mutants (36, 55). In contrast, the aromatic ring at position 29 in the triple mutant is perpendicular to the orientation found in sperm whale L29F- and L29F/H64Q-MbCO (Fig. 2) (17, 56). A detailed comparison of the distal pockets of the L29F/H64Q/V68F-metMb<sub>H2O</sub> and Lucina metHb<sub>H2O</sub> (24) is presented in Table II and Fig. 2. The overall positions of the three mutated residues in the recombinant sperm whale metmyoglobin are close to those of the corresponding residues in the clam protein, but the exact orientations of the phenyl side chains are significantly different. The distal glutamates in both proteins form hydrogen bonds with the coordinated water molecule.

**Table I**

| Protein                        | $O_2$ binding constant ($K_{O2}$) | $H_2S$ binding constant ($K_{H2S}$) |
|--------------------------------|-----------------------------------|-------------------------------------|
| Lucina HbI                     | 2.5                               | 290                                 |
| (Phe<sup>29</sup>B10, Gln<sup>64</sup>(E7), Phe<sup>68</sup>(E11)) |                                   |                                    |
| Wild-type sperm whale Mb       | 1.1                               | 0.05                                |
| (Leu<sup>29</sup>B10, His<sup>64</sup>(E7), Val<sup>68</sup>(E11)) | 1.1                               | 0.05                                |
| L29F-Mb                        | 15                                | 0.32                                |
| H64Q-Mb                        | 1.8                               | 0.32                                |
| V68F-Mb                        | 0.48                              | 0.33                                |
| L29F/H64Q-Mb                   | 0.46                              | 3.7                                 |
| L29F/V68F-Mb                   | 74                                | 5.3                                 |
| L29F/H64Q/V68F-Mb              | 3.8                               | 37                                  |

*References for $O_2$ equilibrium constants at pH 7.0, 20 °C: Lucina HbI (23); wild-type, L29F, H64Q, and V68F Mb (28); L29F/H64Q Mb (17); L29F/V68F and L29F/H64Q/V68F Mb.*

*References for $H_2S$ equilibrium constants at pH 7.5, 20 °C: Lucina HbI (23); wild-type and mutant myoglobins.*

Heme Pocket Structure of L29F/H64Q/V68F-metMbCN and Lucina metHbICN

**Resonance Assignments**—The 500 MHz <sup>1</sup>H NMR spectra for L29F/H64Q/V68F-metMbCN, V68F-metMbCN, and Lucina-metHbICN in $D_2O$ at 25 °C are shown in Fig. 3, A-C, respectively. The signals for the heme for each protein were located and assigned by the characteristic pattern of dipolar contacts about the heme periphery among the TOCSY identified vinyl, propionate groups, and the pyrrole methyl, meso-Hs in a fashion standard for both diamagnetic and paramagnetic heme proteins. The pattern of heme chemical shifts is similar to that observed in wild-type sperm whale metMbCN (not shown). Resonances for residues near the heme are assigned to the degree possible, as limited by spectral congestion and paramagnetic relaxation, and by standard backbone dipolar connectivities together with characteristic TOCSY and/or COSY connectivities for several side chains (15, 16, 49). Other residues not addressable by these approaches are assigned on the basis of dipolar contacts to the heme and/or to other assigned residues. Inasmuch as this assignment strategy has been reported in detail for both wild-type and mutant metMbCN (17, 29, 30, 43),
two-dimensional data are shown only as relevant to the placement of perturbed distal pocket residues in both protein complexes.

The fingerprint portion of the NOESY and COSY maps (not shown) for each protein locates two extended helical fragments each, the shorter one corresponding to F helix residues F4-F9...
TABLE II
Conformational parameters in the distal pockets for the aqueo-met (crystal) and cyanomet (solution) structures of sperm whale L29F/H64Q/V68F-Mb and Lucina pectinata HbI

| Distance or angle | L29F/H64Q/V68FLucina pectinata |
|------------------|---------------------------------|
|                  | MetMbCN | MetMbH,O | LuciMetMbCN | LuciMetHbI,O |
| Fe-C$_5$(Phe$_{68}(E11)$) | 6.17 Å | 5.97 Å | 6.15 Å | 6.36 Å |
| Fe-N$_{2}$(Gln$_{64}(E7)$) | 4.64 Å | 5.98 Å | 4.52 Å | 4.32 Å |
| Fe-C$_2$(Phe$_{68}(E11)$) | 6.09 Å | 8.06 Å |
| Fe-C$_2$(Phe$_{68}(CD1)$) | 5.30 Å | 5.30 Å | 5.30 Å | 4.92 Å |
| Fe-NeHis$_{34}(F8)$ | 2.19 Å | 2.51 Å |
| Fe-OH | 2.08 Å | 2.28 Å |
| Cy$_2$(Gln$_{64}(E7)$) | 4.00 Å | 3.55 Å | 3.62 Å | 4.87 Å |
| Cy$_2$(Phe$_{68}(B10)$) | 1.71 ° | 179 ° | 87 ° | 93 ° |
| x$_1$(Phe$_{68}(B10)$) | -80 ° | -98 ° | -95 ° | -80 ° |
| x$_2$(Phe$_{68}(B10)$) | 171 ° | 179 ° | 87 ° | 93 ° |
| x$_3$(Gln$_{64}(E7)$) | -195 ° | -180 ° | -151 ° | -92 ° |
| x$_4$(Gln$_{64}(E7)$) | 50 ° | 49 ° | 67 ° | -46 ° |
| x$_5$(Gln$_{64}(E7)$) | -130 ° | 45 ° | 168 ° | -98 ° |
| x$_6$(Phe$_{68}(E11)$) | 150 ° | -172 ° |
| x$_7$(Phe$_{68}(E11)$) | 66 ° | 91 ° |

* Data taken from Rizzi et al. (24).

that includes the strongly hyperfine shifted axial His(H8). The strongly relaxed and hyperfine shifted proximal FG corner residues (His$_{57}(F3)$ and Ile$_{69}(F5)$ in Mb mutants, Arg$_{99}(F1)$ and Ile$_{101}(F4)$ in Lucina HbI) are identified by their characteristic NOE SY cross-peaks to the 5-CH$_2$. For each protein, the pattern of heme residue and inter residue dipolar contacts and paramagnetic relaxation reveal a proximal side heme pocket structure that is indistinguishable from that in the respective reference crystal structures (Fig. 1). The chemical shifts for relevant assigned resonances for the three protein complexes are listed in Table III.

The longer helical segment represents the E helix residues E7–E14 in each protein, with residues E10, E11 and E14 exhibiting the expected NOE SY cross-peaks in the heme, including that of the Phe$_{68}(E11)$ ring to the 2-vinyl group. The TOCSY/MCOSY detected fragment NH$_{C}(H)C$_{H}(H) for Gly$_{64}(E7)$ in each protein shows strong NOESY cross-peaks to a moderately relaxed and hyperfine shifted CH$_{CH}$ fragment which is assigned to the remainder of Gly$_{64}(E7)$; the C$_{H}H$-C$_{H}$ TOCSY/ COSY connectivities are not observed because the cross-peak is too close to the diagonal in each protein (Figs. 4 and 5). A strongly relaxed (T$_1$ ~12 ms) and strongly low-field hyperfine shifted labile proton, when saturated, exhibits an intense NOE to another upfield shifted and moderately relaxed (T$_1$ ~50 ms) labile proton in Lucina metHbIICN (Fig. 3E), but not detected because of overlap with the solvent signals in the triple pulse, as well as to the Gly$_{64}(E7)$ C$_{H}$s (and for the triple mutant, to Thr$_{77}(E10)$; Fig. 3D) and confirms the two labile protons as arising from the N$_{H3}$ of Gly$_{64}(E7)$. The T$_1$s via Equation 5 reveal, R$_{Fe}(N_{H}) = 4.3 ± 0.2$ Å for the low-field NH for both proteins, and for Lucina metHbIICN, the upfield resolved R$_{Fe}(N_{H}) = 4.7 ± 0.2$ Å. The resulting R$_{Fe}$ are listed in Table II.

One strongly relaxed (for C$_{H}$) and hyperfine shifted aromatic ring with NOE SY contact to 5-CH$_{2}$ and Gly$_{64}(E7)$ (not shown) which uniquely identifies Phe$_{68}(CD1)$. Another strongly hyperfine shifted and relaxed (C$_{H}$ 19 ppm, T$_1$ ~33 ms) aromatic ring exhibits NOE SY cross-peaks to Phe$_{68}(E11)$, Gly$_{64}(E7)$, Gly/Ala$_{68}(E8)$, and Phe$_{68}(CD1)$ and definitively assigns Phe$_{68}(B10)$ for each protein (Figs. 4 and 5). The chemical shifts of relevant assigned distal residues for the three proteins are included in Table III. The pattern of dipolar contacts between distal residues and the heme, and among distal residues, are predicted by the crystal structure, as shown schematically in Fig. 1. The difference that are noted, as described below, involve primarily the residues Gly$_{64}(E7)$.

Magnetic Axes Determination—All dipolar shifted residues except Phe$_{68}(E11)$ exhibit a good correlation between δ$_{dd}$(obs) and the slope of the chemical shift in a Curie (chemical shift versus $T^{-1}$) plot that is indicative of well defined orientations to the heme (not shown) (49). A variety of proximal proton dipolar shifts which correlate with their Curie slopes for each protein were used. To determine the magnetic axes, both as three-parameter searches for (α,β,γ) using the wild-type metMbCN anisotropies, δ$_{xx}$ = 2.04 × 10$^{-3}$ m$^2$/mol, δ$_{xy}$ = -0.48 × 10$^{-9}$ m$^2$/mol (29), and as five-parameter searches for both (α,β,γ) and the anisotropies. Both mutants exhibit excellent correlation between δ$_{dd}$(obs) and δ$_{dd}$(calc) with very low residual F$_{in}$ in Equation 1 (not shown) typical of previous NMR studies of mutant metMbCNs (15, 16, 29, 30, 50). The chemical shifts for numerous distinct residues are also well predicted. The resulting Euler angles using different input data sets were highly clustered for each protein, although distinct for each of the proteins. For L29F/H64Q/V68F-metMbCN, the optimized angles (and ranges for the various fits) are: a tilt of the major axes from the heme normal, β ~ 6° (6.1–6.6°), direction of tilt, α ~ -40° (~40° ~50°) and rhombic axes, κ = α + γ ~ 30° (20° ~40°) using MbCO crystal coordinates. The magnetic axes for V68F-metMbCN are α = 5° (0–10°), β = 7.5° (7°–8°), and κ = 40° (30°–50°). The optimized anisotropies for each five-parameter fit differ inconsequentially for those of wild-type metMbCN and yield Euler angles within the ranges obtained by the 3-parameter fit (not shown). When the coordinates for L29F/H64Q/V68F-metMbCN are taken from the L29F/H64Q/V68F-metMbH$_2$O crystal structure presented herein, we obtain values for the axes and anisotropies that are well within the ranges obtained using the wild-type coordinates. The magnetic axes (and ranges in their values), obtained similarly for Lucina metMbIICN, using either the Lucina metHbIICN (24) or metHbIICN (25) crystal coordinates, are β = 7° (6°–8°), α = 155° (150°–160°), κ = α + γ = 255° (250°–260°), and optimized anisotropies are unchanged from those of wild-type metMbCN (not shown).

Orientation of Distal Residues in L29F/H64Q/V68F-metMbCN—Using either the wild-type MbCO crystal coordinates with the mutated Phe$_{68}(B10)$ inserted with an orientation as found in the crystal structure of L29F/H64Q-MbCO (half-open squares in Fig. 6A), or directly the crystal coordinates of the presently characterized L29F/H64Q/V68F-metHbIICN (open squares in Fig. 6A), a qualitative, but not quantitative, correlation between δ$_{dd}$(calc) and δ$_{dd}$(obs) are observed, and the R$_{Fe}$ (C$_{H} = 4.7$ Å) is shorter than indicated by T$_1$ = 33 ms (R$_{Fe}$ = 5.1 ± 0.3 Å). However, altering x$_1$ by 5° from that in the L29F/H64Q-metMbCO and a 50° rotation of x$_2$ leads to a Phe$_{68}(B10)$ orientation whose δ$_{dd}$(calc) correlate very well with δ$_{dd}$(obs) (closed squares in Fig. 6A) and for which the R$_{Fe}$ (C$_{H} = 5.1$ Å is consistent with the T$_1$ value.

Introduction of Gly$_{64}(E7)$ into the wild-type MbCO crystal structure based on its orientation in the HbI-MbCO (or L29F/H64Q-MbCO) crystal structure results in reasonable agreement with both T$_1$ data and δ$_{dd}$(calc) (half-closed circles in Fig. 6A). However, when directly using the crystal coordinates for L29F/H64Q/V68F-metMbH$_2$O, neither the δ$_{dd}$(calc) (open circles in Fig. 6A) nor T$_1$ value for N$_{H}$ (R$_{Fe}$ = 5.9 Å) are reasonably predicted (T$_1$ = 14 ms, R$_{Fe}$ = 4.2 ± 0.2). Small rotation of x$_1$ (~5°) and x$_2$ (~8°), followed by a 180° rotation of x$_3$ for
Gln$^{64}$(E7) in the L29F/H64Q/V68F-metMbH$_2$O crystal structure leads to an excellent fit for $\delta_{dip}(\text{obs})$ versus $\delta_{dip}(\text{calc})$ (closed circles in Fig. 6A) and yields $R_p(N\_H)$ = 4.2 Å, in good agreement with the 4.2-Å value obtained from the $T_1 = 14$ ms via Equation 5. The other N\_H is correctly predicted to resonate close to the solvent signal by the optimized $\delta_{dip}(\text{calc})$. Similarly small variations in $\chi_1$, $\chi_2$ starting with the Gln$^{64}$(E7) orientation in the single mutant lead to a Gln orientation essentially the same as that obtained starting with the coordinate of the triple mutant (not shown).

The anomalous variable temperature slopes (see above), small $\delta_{dip}(\text{calc})$ which are relatively insensitive to ring orientation, and the presence of comparable magnitude but opposed sign $\delta_{dip}(\text{calc})$ and ring current shift rendered the optimization of the orientation of Phe$^{68}$(E11) ring on the basis of $\delta_{dip}(\text{obs})$ completely impractical (50). The pattern of small $\delta_{dip}$ for Phe$^{68}$(E11), nevertheless, are consistent with observed dipolar shifts and NOE contacts. The orientation for these three mutated residues in L29F/H64Q/V68F-metMbCN are shown in Fig. 2A, and the relevant side chain bond angles and distances to the iron are listed in Table II.

**Orientation of Distal Residues in Lucina MetHbICN**—The orientations for Gln$^{64}$(E7) in the Lucina metMbH$_2$O (24) or metHbIH$_2$S (25) crystal structures yielded completely unacceptable fits for $\delta_{dip}(\text{obs})$ versus $\delta_{dip}(\text{calc})$ for Lucina-metMbH$_2$O in solution (as shown by open circles and half-closed circles, respectively, in Fig. 6B); the correlation is particularly poor for the labile N\_Hs where both models predict even the wrong sign for $\delta_{dip}(\text{calc})$. Moreover, the low-field N\_H closest to the iron has $R_p(N\_H)$ = 3.5 Å and 5.2 Å in the two crystal structures while the observed $T_1 = 12$ ms predicts $R_p(N\_H)$ = 4.2 ± 0.2 Å via Equation 5. Since the Gln$^{64}$(E7) orientation in Lucina metMbH$_2$O had the N\_H oriented toward the ligated water, which is expected to be a H-bond donor, the Gln orientation with $\chi_3$ rotated by 180° to interchange the carbonyl and amide groups was considered. This altered orientation leads to different N\_H dipolar shifts (not shown) that fit $\delta_{dip}(\text{obs})$ better, but still unacceptably. Attempts to obtain a fit for $\delta_{dip}(\text{obs})$ by sequential $\chi_1$, $\chi_2$, $\chi_3$ rotation failed since $\delta_{dip}(\text{calc})$ for C$_\text{H}$s is quite insensitive to the $\chi_1$. Instead, the combination of the large low-field $\delta_{dip}(\text{obs})$ and $R_p(N\_H)$ = 4.2 Å (from the $T_1 = 12$ ms) was used to uniquely locate the low-field N\_H in the crystal coordinates, and a search made for the position in space of the other N\_H (with intra N\_H distance 1.88 Å) to satisfy both the upfield $\delta_{dip}(\text{obs})$ and $R_p(N\_H)$ = 4.6 ± 0.2 Å obtained from its $T_1 = 22$ ms. Upon obtaining a reasonable fit for the N\_Hs that satisfy both relaxation and $\delta_{dip}(\text{obs})$ constraints, a search was pursued for the range of $\chi_1$, $\chi_2$, and $\chi_3$ allowed by the fixed C$_\text{H}$ and NH$_2$ positions. The good correlation between $\delta_{dip}(\text{calc})$ and $\delta_{dip}(\text{obs})$ for this optimized Gln orientation is shown in Fig. 6B (closed circles). The resulting torsional angles for Gln$^{64}$(E7) are listed in Table II where they are compared with values in the crystal structure of Lucina metMbH$_2$O.

The orientation of Phe$^{68}$(B10) in the crystal structures of either Lucina metHbH$_2$O (24) or metHbIH$_2$S (25) yield very similar ring proton $\delta_{dip}(\text{calc})$ that correlate equally well with $\delta_{dip}(\text{obs})$ (open and closed squares in Fig. 6B, respectively). The crystallographic $R_p(N\_H)$ (C$_\text{H}$) (5.6 Å), however, is somewhat larger than the value reflected in its $T_1 = 40$ ms ($R_p(N\_H)$ = 5.2 ± 0.2 Å via Equation 5. A -5° change in $\chi_1$ alters the $\delta_{dip}(\text{calc})$ inconsequentially but results in a reasonable $R_p(N\_H)$ (C$_\text{H}$) = 5.3 Å. Phe$^{68}$(E11) exhibits only minor $\delta_{dip}(\text{obs})$ and, like in the triple mutant Mb, yields anomalous Curie plots that reflect variable population of alternate orientations (50). The crystallographic orientation, however, is consistent both with $\delta_{dip}(\text{calc})$ (not shown) and NOE symptomatic cross-peak pattern (Fig. 1). It is noted that
for Lucina metHbICN, in contrast to L29F/H64Q/V68F-metMbCN, the crystallographic distal Phe$_{64}$(CD1) placement in either crystal structure yields similarly poor fits for $\delta_{\text{obs}}$ versus $\delta_{\text{calc}}$ (open triangles). The positions of the distal residues for metHbICN and V68F-metMbCN in solution are compared in Fig. 2B, and reveal significant differences primarily for the Gln$_{64}$(E7) orientation.

**DISCUSSION**

**Solution NMR Cavity Structure**—The conventional two-dimensional NMR approach allowed unambiguous assignment for all relevant residues in the heme cavity. The intrinsic paramagnetism of cyano-metMb, instead of being an impediment to assignment, was actually an advantage since the dipolar shift with minimal relaxation provide significantly enhanced resolution for the E and F helical segments. For the proximal side of the heme, the NOESY cross-peak patterns and paramagnetic relaxation for the two Mb mutants were indistinguishable from those observed for wild-type metMbCN and predicted by the wild-type MbCO structure. A similarly strongly conserved proximal structure is confirmed for Lucina metHbICN. These conserved proximal structures allow the determination of meaningful magnetic axes for each protein. The robust nature of the magnetic axes for L29F/H64Q/V68F-metMbCN is confirmed by the indistinguishable magnetic axes determined using either the wild-type MbCO or triple mutant crystal structures.

For the distal residues, whose relaxation properties, dipolar shifts and/or NOESY cross-peak pattern differed from that expected for the reference crystal structure, the combined restraints for dipolar shifts and paramagnetic relaxation allow the determination of the their orientations in a more definitive manner than allowed by the NOESY cross-peak pattern themselves. The accuracy of the distal orientations is limited only by the basic assumption that the helical backbone positions are inconsequentially perturbed in the mutant and/or ligation variant relative to the available crystal structure. It is noted that the labile protons of the distal Gln$_{64}$(E7) are readily assigned.

### Table III

| Residues Peaks | L29F/H64Q/V68F-metMbCN | V68F-metMbCN | Lucina metHbICN |
|----------------|-------------------------|--------------|-----------------|
| Phe$_{29}$(B10) | Ring$_s^a$ | 8.28, 10.73, 19.49 | 5.38, 3.98, 4.48 | 8.70, 13.07, 18.77 |
| Phe$_{43}$(CD1) | Ring$_g^b$ | 7.48, 9.42, 11.95 | 7.55, 9.54, 11.92 | 6.40, 6.20, 7.26 |
| Gln$_{64}$(E7) | NpH | 9.16 | 8.93 | |
| | CaH | 5.88 | 4.62 | |
| | CbH$_s$ | 5.56, 4.33 | 3.03, 3.29 | |
| | CyH$_s$ | 11.12, 7.19 | 7.63, 4.97 | |
| | NcH$_s$ | 18.67, 5.1 | 14.05, −1.39 | |
| Thy$_{57}$(E10)/Ser$_{65}$(E10) | NpH | 9.34 | 8.58 | |
| | CaH | 3.34 | 3.59 | |
| | CbH$_s$ | 5.22 | 2.16 | |
| | CyH$_s$/CbH$_s$ | $−0.59$ | $−1.00$ | 3.45 |
| Phe$_{69}$(E11) | NpH | 10.61 | 9.89 | |
| | CaH | 1.80 | 3.81 | |
| | CbH$_s$ | 6.27, 5.22 | 2.84 | 9.05, 6.40 |
| | Ring$_g$ | 8.85, 6.81, 7.67 | 8.57, 8.07, 7.88 | 7.47, 7.53, 7.87 |
| Ala$_{14}$(E14)/Leu$_{17}$(E14) | NpH | 6.11 | 6.03 | |
| | CaH | 3.26 | 3.04 | |
| | CbH$_s$/CbH$_s$ | $−0.66$ | $−0.41$ | $−0.86$, $−0.69$ |
| | CyH$_s$ | $−2.01$ | |
| | CyH$_s$ | $−2.36$, $−0.83$ | |
| Leu$_{39}$/Phe$_{29}$(F4) | NpH | 7.26 | 7.89 | 5.98 |
| | CaH | 6.63 | 7.56 | |
| Ala$_{90}$/Thr$_{92}$(F5) | NpH | 10.23 | 9.64 | |
| | CaH | 6.42 | 6.26 | 5.60 |
| | CbH$_s$ | 2.53 | 2.31 | |
| His$_{92}$/Thr$_{94}$(F6) | NpH | 13.89 | 12.52 | |
| | CaH | 9.34 | 9.17 | 10.86 |
| | CbH$_s$ | 13.93, 9.64 | 12.64, 7.97 | 11.34, 9.13 |
| | CyH$_s$ | 15.83 | 10.10 | 18.79 |
| | CaH$_s$ | 4.5 | $−2.25$ | $−6.14$ |
| | NpH$_s$ | 18.23 | 15.94 | |
| His$_{97}$/Thr$_{100}$/Arg$_{99}$(FG3) | Ring$_g$/CbH$_s$ | 13.00, 6.76 | 13.80, 7.20 | 4.26, 6.61 |
| Ile$_{101}$/Thr$_{103}$/Ile$_{105}$(FG5)/Ile$_{105}$(FG4) | CaH | 3.35 | 2.98 | 4.07 |
| | CbH$_s$ | 2.68 | 1.41 | 3.07 |
| | CyH$_s$ | $−3.53$, $1.33$ | $−5.46$, 0.49 | 0.50 |
| | CyH$_s$ | $−1.38$ | $−2.33$ | |
| | CbH$_s$ | $−2.71$ | 0.83 | |
| Ile$_{106}$/Phe$_{109}$/Ala$_{109}$(G8) | NpH | 6.84 | 3.61 | |
| | CbH$_s$/CyH$_s$ | 0.24 | 0.23 | |
| | CbH$_s$/CbH$_s$ | $−0.54$ | |
| Phe$_{126}$(H15) | Ring$_g$ | 6.84, 6.97, 7.11 | 7.08, 6.62, 7.08 | |

$^a$ For Phe ring protons, the shifts are listed as: CbH$_s$, CbH$_s$, and CyH$_s$.

$^b$ For His ring proton, the shifts are listed as CbH and CbH.$_s$.

1H NMR chemical shifts of the amino acid residues in sperm whale L29F/H64Q/V68F-metMbCN and V68F-metMbCN and in Lucina metHbICN.

Chemical shifts in ppm at 25 °C (L29F/H64Q/V68F-metMbCN and Lucina metHbICN) and at 30 °C (V68F-metMbCN), pH 7.2, reference to DSS.

Structure of Sulfide-binding Globins

9523
The aromatic ring for Phe68(E11) is oriented away from the expected binding site with the ring inserting into a largely pre-existing cavity in the distal pocket (55). The absence of a methyl group at the \( \gamma_2 \) position results in less direct hindrance of the bound ligand than is observed for the naturally occurring Val side chain. Thus, the Phe68(E11) mutant behaves like V68A-metMbCN; both exhibit less tilt, \( \beta \sim 7-9^\circ \), than that in the wild-type protein (\( \beta = 16^\circ \)) (30, 57). Last, the labile Gln64(E7) protons are in a position to form hydrogen bonds with bound cyanide in the distal pockets of both L29F/H64Q/

The aromatic ring for Phe68(E11) is oriented away from the expected binding site with the ring inserting into a largely pre-existing cavity in the distal pocket (55). The absence of a methyl group at the \( \gamma_2 \) position results in less direct hindrance of the bound ligand than is observed for the naturally occurring Val side chain. Thus, the Phe68(E11) mutant behaves like V68A-metMbCN; both exhibit less tilt, \( \beta \sim 7-9^\circ \), than that in the wild-type protein (\( \beta = 16^\circ \)) (30, 57). Last, the labile Gln64(E7) protons are in a position to form hydrogen bonds with bound cyanide in the distal pockets of both L29F/H64Q/
FIG. 6. Plot of $\delta_{\text{dp}}(\text{obs})$ versus $\delta_{\text{dp}}(\text{calc})$ for key distal residues. A, L29F/H64Q/V68F-metMbCN for Phe$^{29}$(B10) (squares) and Gln$^{64}$(E7) (circles) based on the L29F/H64Q-MbCO crystal structure (half-filled marker), L29F/H64Q/V68F-metMbH$_2$O (open marker) crystal structures, and the latter crystal structures with Gln$^{64}$(E7) $\chi_3$ rotated by 180° (open marker with asterisk) as determined by NMR (closed markers); B, Lucina metHbICN for Phe$^{29}$(B10) (squares), Gln$^{64}$(E7) (circles) and Phe$^{43}$(CD1) (triangles) based on the crystal structures of Lucina metHbICN (open markers), the crystal structure of Lucina metHbICN (half-closed markers with asterisk), and as determined by NMR (closed markers).

V68F metMbCN and Lucina metHbICN. This orientation is probably similar to that in the oxy complexes.

Comparison of Distal Pockets in Solution and Crystal Structures—For all but select distal residues, the heme pocket structure for each protein is essentially the same in cyanomet and aquo-met derivatives. In the case of the sperm whale triple mutant, the major difference in the Gln$^{64}$(E7) orientation between the cyanomet and aquo-met complexes is −180° rotation about $\chi_3$ that interchanges the carboxy and amide termini (Table II, Fig. 2A). The 180° rotation of the Gln side chain terminus is completely consistent with the stabilizing roles of the side chain for the different ligands. For the cyano complex, the N$_3$H$_2$ is oriented to serve as a H-bond donor to bound cyanide, whereas in the aquo-met complex, the carbonyl oxygen is oriented to accept a H-bond from coordinated water. The small changes in $\chi_1$, $\chi_2$ for Gln$^{64}$(E7) (Table II) help to accommodate the larger cyanide ligand.

As shown in Fig. 2B, the solution orientation of Gln$^{64}$(E7) in Lucina metHbICN I differs substantially from that reported for the corresponding aquomet crystal. The altered orientation is firmly and independently established by the NOESY pattern from C$_9$H$_3$ to Phe$^{29}$(B10), the dipolar shift simulation, and the paramagnetic relaxation. However, both structures place the amino group in a position to interact with the bound ligand. The electron density in the crystal structure does not allow differentiation between the terminal carbonyl O atom and the N$_3$ atom. However, the carbonyl O atom rather than the reported NH$_2$ must be pointing toward the coordinated water to serve as H-bond acceptor. Donation of proton to the bound water would give the ligand a partial positive charge (i.e. H$_2$O$^+$ character) causing substantial electrostatic repulsion with the net +1 charge on the hemin iron atom. Thus, the solution NMR data, through the dipolar shift and paramagnetic relaxation, provide the most definitive location of the NH$_2$ group. The small movements of Phe$^{29}$(B10) and Phe$^{43}$(CD1) in Lucina HbI appear to reflect minor accommodation of the larger cyanide ligand. However, the orientations of Gln$^{64}$(E7) do differ in solution and crystal, as shown in Fig. 2B.

The orientation of the Gln$^{64}$(E7) side chain in the sulfide complex of Lucina metHbI (not shown) is intermediate between that for the aquomet crystal and that for metHbICN in solution (Fig. 2B). The CO versus NH$_2$ terminus orientation in the sulfide complex of Lucina metHbI also differs from that in metHbICN in solution by a 180° rotation at $\chi_3$. Again, this difference is consistent hydrogen bond donation to bound cyanide and acceptance from bound H$_2$S.

Implications for Sulfide Binding—All three single distal pocket substitutions in sperm whale Mb cause an enhancement of sulfide binding but have variable effects on oxygen affinity. The key mutation is H64Q, which causes a 25-fold increase in sulfide affinity by itself. The triple mutant, L29F/H64Q/V68F-metMb, has a sulfide affinity which is ~700 times greater than that of wild-type metMb and only ~7-fold less than that of Lucina HbI. Thus the triple mutant exhibits an increased sulfide binding free energy of ~3.8 kcal/mol or about ~75% of the stabilization of the bound sulfide (~5.1 kcal/mol) by native Lucina HbI relative to that by wild-type sperm whale Mb. Hence, a major portion, but not all, of the remarkable sulfide affinity of Lucina HbI can be transferred to a mammalian Mb by replacing the three residues in contact with bound sulfide with those found in Lucina HbI. The high affinity of Lucina HbI for sulfide was attributed (24, 25) to a combination of the hydrogen-bond acceptance by the Gln$^{64}$(E7) side chain carbonyl, the hydrophobic pocket provided by the “cage” generated by the three Phe side chains, B10, CD1, and E11, and highly favorable electrostatic interactions between the sulfur and the edges of the three aromatic rings. While the distal pockets for the triple mutant Mb and Lucina HbI are similar (Fig. 2), there are some significant differences that relate to their differences in affinity for sulfide.

The side chains of Gln$^{64}$(E7) and Phe$^{43}$(CD1) have similar orientations in the sperm whale triple mutant and the clam
protein (Fig. 2C). However, both Phe^{29}(B10) and Phe^{68}(E11) in the sperm whale mutant are rotated ~90° about χ₂ relative to their positions in Lucina HbI and their edges are closer to the iron atom in the mollusc Hb (Fig. 2C). As a result of the differences, Lucina hemoglobin has a significantly larger ligand-binding site than that found in the myoglobin mutants, which should facilitate the binding of the large H₂S ligand. The change in orientation of Phe^{29}(B10) should have little effect on stabilization of bound sulfide since the positive edge of the ring multipole is pointing toward the bound ligand in both proteins. In Lucina HbI, the positive edge of the Phe^{68}(E11) ring also points toward the bound ligand (Fig. 2C). This orientation of the Phe(E11) side chain is caused by direct steric interactions Phe^{29}(B9) in Lucina HbI (25). Stabilization of bound sulfide by Phe^{68}(E11) ring does not appear to occur in the myoglobin triple mutant since the E-11 side chain is allowed to take a position more perpendicular to the plane of the heme since the B9 position is occupied by a smaller Ile residue. A more complete and quantitative interpretation of the sulfide binding data will require detailed volume and electrostatic calculations for the entire set of mutants listed in Table I; such studies are now in progress.

Acknowledgments—We thank Eileen Singleton for the expression and purification of the triple mutant myoglobin, Mike Berry for help in the protein's crystallization and subsequent structural refinement, and Jun Qin for assistance with NMR experiments.

REFERENCES
1. Lesk, A. M., and Chothia, C. (1980) J. Mol. Biol. 136, 225–270
2. Bashford, D., Chothia, C., and Lesk, A. M. (1987) J. Mol. Biol. 196, 199–216
3. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
4. Romero-Herrera, A. E., Goodman, M., Dene, H., Knight, D. E., and Braunitzer, G. (1982) J. Mol. Biol. 258, 529–544
5. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
6. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
7. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
8. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
9. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
10. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
11. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
12. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
13. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
14. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
15. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
16. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
17. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
18. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
19. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
20. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
21. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
22. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
23. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
24. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
25. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
26. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
27. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
28. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
29. Vinogradov, S. N., Walr, D. A., Pohajdak, B., Moens, L., and Kapp, O. H. (1993) Biochem. Physiol. 106, 1–26
Solution and Crystal Structures of a Sperm Whale Myoglobin Triple Mutant That Mimics the Sulfide-binding Hemoglobin from *Lucina pectinata*

Bao D. Nguyen, Xuefeng Zhao, Krishnamurthi Vyas, Gerd N. La Mar, R. Ashley Lile, Eric Allen Brucker, George N. Phillips, Jr., John S. Olson and Jonathan B. Wittenberg

*J. Biol. Chem.* 1998, 273:9517-9526.
doi: 10.1074/jbc.273.16.9517

Access the most updated version of this article at [http://www.jbc.org/content/273/16/9517](http://www.jbc.org/content/273/16/9517)

Alerts:
- When this article is cited
- When a correction for this article is posted

[Click here](http://www.jbc.org/content/273/16/9517.full.html#ref-list-1) to choose from all of JBC's e-mail alerts

This article cites 53 references, 13 of which can be accessed free at [http://www.jbc.org/content/273/16/9517.full.html#ref-list-1](http://www.jbc.org/content/273/16/9517.full.html#ref-list-1)