THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 270, No. 35, Issue of September 1, pp. 20754–20762, 1995
Printed in U.S.A.

Crystal Structure of Asian Elephant (*Elephas maximus*) Cyano-Metmyoglobin at 1.78-Å Resolution

Phe^{29}(B10) ACCOUNTS FOR ITS UNUSUAL LIGAND BINDING PROPERTIES*

(Received for publication, April 24, 1995)

Daniel A. Bisig, Ernesto E. Di Iorio, Kay Diederichs‡, Kaspar H. Winterhalter, and Klaus Piontek§

From the Laboratory of Biochemistry I, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland and the ‡Institut für Biophysik, Universität Freiburg, D-79104 Freiburg, Germany

The crystal structure of Asian elephant cyano-metmyoglobin which has a glutamine instead of the usual distal site histidine has been determined to high resolution. In addition to this replacement, the substitution of a conserved leucine residue in position 29(B10) at the distal side by a phenylalanine was unambiguously identified based on the available electron density. The suspicion, that there were errors in the original sequence which has caused some confusion, is thus confirmed. Comparison with other myoglobin structures in various ligated forms reveals an essentially unchanged tertiary structure in elephant myoglobin despite the two amino acid substitutions in the heme pocket. Our current structural model shows that the $N_\epsilon 2$ atom of Gln^{64}(E7) has moved with respect to the corresponding nitrogen position of His^{64}(E7) in the CO complex of sperm whale myoglobin. The newly assigned residue Phe^{29}(B10) penetrates into the distal side of the heme pocket approaching the ligand within van der Waals distance and causing a much more crowded heme pocket compared to other myoglobins. Kinetic properties of Asian elephant myoglobin, wild type, and recombinant sperm whale myoglobins are discussed in relation to the structural consequences of the two amino acid substitutions H64Q and L29F.

Myoglobin (Mb),[1] a small, mostly α-helical, globular hemeprotein, is most abundant in heart, but skeletal muscles also contain variable amounts of this protein. Mb is commonly said to act as oxygen storage, even though it is less present in skeletal muscles of young animals, which are more active, than older ones. Furthermore, the amount of myoglobin present in the heart of an animal is sufficient to supply oxygen only for a few beats (1). Mb is also reported to facilitate O_2 unloading from hemoglobin and O_2 diffusion to the mitochondrial surface (2). Oxidative phosphorylation appears also to be mediated by Mb (2). At our present level of understanding, we cannot decide with certainty about all possible roles of Mb, and more work is needed to fully clarify the situation.

In addition to their physiological implications, studies on Mb are also of general relevance to our understanding of the relationships between structures, dynamics, and function in proteins. Mb is the first protein for which a three-dimensional structure has been determined to atomic resolution by means of x-ray crystallography (3). Its folding is nearly identical with that of hemoglobin subunits, even of highly unrelated organisms like the mollusc *Scapharca inaequivalvis* (4) or the larvae of the insect *Chironomus thummi thummi* (5). Nearly all experimental approaches can be applied to Mb, due to its relatively small size, to the presence of a heme, and its availability in large quantities. Indeed, Mb is probably one of the best characterized proteins and therefore can be used as a reference for comparative analysis on other more complex systems.

The interplay between the overall dynamic properties and the structural features of the heme pocket in modulating the reactivity of the heme iron is an important element for our understanding of the structure, dynamics, and function relationships in hemeproteins. Simple heme compounds without a protein matrix exhibit a preferred binding to CO *versus* O_2 with a ratio of several ten thousands, while a situation more favorable for oxygen affinity is found in myoglobins, allowing these proteins to function as efficient oxygen carriers. Individual binding properties arise from the specific protein environment in the vicinity of the prosthetic group. Of particular interest in this respect is the distal site pocket which harbors the bound ligand. Much attention has been drawn to the role of the distal residue 64 at the E7 helical position, since it directly interacts with the bound ligand. In most hemeproteins, ranging from hemoglobin (6) to lignin peroxidase (7), a histidine residue is found in this place. Of all known vertebrate myoglobin sequences, only a few have no histidine occupying position E7, one of them being elephant myoglobin which has a glutamine (8). The role of the distal residue E7 is probably to provide a hydrogen bond to the bound oxygen ligand, favoring its binding against CO and reducing autooxidation of the heme iron (9). Further effects on ligand access and binding by other distal side amino acids correlate to their specific steric and electrostatic properties.

The primary structure determination of Asian elephant myoglobin (8) revealed that the distal side residues are conserved apart from the E7 Gln substitution. Against expectations, kinetic studies on elephant myoglobin show an unchanged CO association constant if compared to other vertebrate myoglobins (10). Therefore, it was concluded that the interaction of bound ligand with the E7 residue remains essentially unchanged. The recent progress in molecular biology permits us to probe the function of certain residues by site-directed mu-

* 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.

The atomic coordinates and structure factors of EmMbCN (code 1EMY) have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

§ To whom correspondence and reprint requests should be addressed: Laboratorium für Biochemie I, Eidgenössische Technische Hochschule, Universitätstrasse 16, CH-8092 Zürich, Switzerland. Tel.: 41-1-632-3141; Fax: 41-1-632-1121; E-mail: piontek@wawona.vmsmail.ethz.ch.

[1] The abbreviations used are: Mb, myoglobin; SWMb, sperm whale myoglobin; σ, standard deviation; EMb, elephant myoglobin; EmMbCN, elephant cyano-metMb.

tagenesis. A single mutant of sperm whale myoglobin with a Gln residue at position E7 instead of the common His has been constructed (11). The substitution was expected to produce a kinetic behavior similar to elephant myoglobin. Instead, a 5-fold decrease of the O_2 binding constant and a 3-fold increase in the rate of autooxidation were observed. Obviously, this single amino acid change in the distal side, deduced from the reported primary sequence, is not sufficient to explain the kinetic properties of Asian elephant myoglobin which are similar to myoglobins having the usual histidine E7. Yu *et al.* (12) and Vyas *et al.* (13) have reported proton NMR spectra of Asian elephant myoglobin. These data were interpreted in terms of the presence of an additional phenylalanine in the heme crevice close enough to the prosthetic group to interact with the bound ligand. On the basis of the available sequence of EMb (8), this feature cannot be explained unless a major reshuffling of the fold is assumed. Among point mutations at position 29 (B10), a recombinant L29F SWMb was produced which is characterized by a dramatic increase in O_2 affinity and a significant decrease of autooxidation (14). A double mutant L29F/H64Q construct shows rate constants for O_2 and CO binding and autooxidation which are very similar to EMb, indicating that the two effects of the single mutation cancel each other. On the other hand, these two mutations do cause a 4-fold decrease of the CO dissociation rate resulting in a significantly increased CO equilibrium constant (see accompanying paper by Zhao *et al.* (44)).

In the absence of detailed three-dimensional structural information on EMb, its peculiar CO kinetics and light absorption properties over broad temperature ranges (15) have been tentatively ascribed to the crowding of its distal heme crevice proposed by the La Mar group (12, 13). The high susceptibility of the prosthetic group to structural fluctuations of the protein moiety, indicated by the thermal evolution of the kinetic and optical spectroscopy properties of EMb, was also tentatively related to the tight packing of its heme pocket.

The determination of the crystal structure of elephant myoglobin has been undertaken to clarify the situation at the distal side residues and to correlate the structural information with available kinetic and optical spectroscopy data. The results presented here show that the primary structure reported by Dene *et al.* (8) contains a crucial error and a phenylalanine is present at position 29 (B10) of EMb. An additional small sequencing error was found at position 27 (B8) where a Thr was observed in our structure instead of a Phe as reported by Dene *et al.* (8). But, this newly assigned Thr residue appears of no functional consequences. Phe^{29}(B10) is in close contact with the bound ligand and occupies the cavity in which the free CO is found in the 40 K crystal structure of photolyzed SWMb recently reported by Schlichting *et al.* (16).

EXPERIMENTAL PROCEDURES

Sample Preparation—Elephant myoglobin was prepared according to a previously described method (17) with minor modifications. About 200 g of Asian elephant skeletal muscle were chopped and washed with cold 20 mM Tris/HCl buffer (pH 8.4, 1 mM EDTA). The meat was homogenized in twice the volume of the washing buffer in a Waring blender. After slow addition of ammonium sulfate to 65% saturation, while keeping the pH above 8 by addition of ammonia, the homogenate was stirred for 1 h at 4 °C. The precipitated proteins were separated by centrifugation for 20 min at 4 °C and 27,000 × *g*. The supernatant was concentrated by ultrafiltration (Amicon YM3 membrane, 4 °C, 2 bar) and subsequently dialyzed against washing buffer. At this stage, the met form of myoglobin was produced by oxidizing Fe^{2+} to Fe^{3+} with a 2-fold excess of ferricyanide in the presence of KCN. To remove the excess ferricyanide and cyanide, as well as large molecular weight proteins, the protein solutions were concentrated and applied on a Sephadex G-100 column equilibrated and eluted with washing buffer. Contrary to the original procedure (17), the subsequent step of anion exchange chromatography with a DEAE-Sephacell column was omitted. Instead, the protein was further purified by preparative isoelectric focusing (18) in order to obtain myoglobin which was pure enough for crystallization.

Crystallization—Deep red crystals of elephant cyano-metMb (EmMbCN) were obtained by hanging drop vapor diffusion (19). Initial crystallization conditions were found utilizing the fast screening method (20) with 0.5 ml of reservoir liquid equilibrated against 3 μl of protein solution mixed with the same volume of reservoir liquid. Well diffracting crystals grew at room temperature in 45% polyethylene glycol 1000, buffered with 100 mM Tris/HCl to pH 8.5 in the presence of 150 mM magnesium acetate. The protein concentration was 10 mg/ml.

Macroseeding was important to obtain large (0.8 × 0.4 × 0.3 mm^3) single crystals. Seeding crystals have been obtained by hair-seeding experiments using a rabbit whisker. The crystals belong to space group $P2_12_12_1$ with unit cell dimensions a = 33.51 Å, b = 58.59 Å, c = 70.44 Å with one molecule per asymmetric unit and a V_m value of 2.03 $Å^3$/Da which is calculated according to Matthews (21). The crystals diffract to at least 1.70 Å on a conventional x-ray source.

Data Collection—Diffraction data were obtained from one single crystal measured at 4 °C on a Stoe imaging plate system (crystal to detector distance 70 mm) equipped with a rotating anode tube providing monochromatic CuK_α radiation. The overall R_{merge} value (based on I) was 4.5% (all data) for 13,649 unique reflections derived from 68,768 observations up to 1.78 Å. The data are 93.6% complete to a Bragg spacing of 1.75 Å and 98.5% complete to a Bragg spacing of 1.85 Å. Data processing and scaling was performed by the programs XDS and XSCALE both written by Wolfgang Kabsch, EMBL/Germany (22).

Phase Determination—The structure was solved by the molecular replacement technique utilizing the sperm whale metmyoglobin structure (Brookhaven Protein Data Bank entry 1MBN) as a search model. For the molecular replacement method, the program package Merlot (23) was used. The solution of the rotation function was obtained with the Crowther fast rotation function (24). For 8.0 Å–3.0 Å resolution data, a unique peak was found with a height of 7.6 times the standard deviation of the map. The three translational parameters were determined with the Crowther and Blow translation function (25) resulting in one peak in each Harker section varying from 8.1 to 8.7 times the standard deviation. These six parameters were further improved by a rigid body R-factor minimization procedure. The initial R-factor was 43.7% for 6.0 Å–3.0 Å resolution data as calculated with the program SCALENEW from the CCP4 suite of programs (26).

Model Building and Structure Refinement—According to the published (8) amino acid sequence, a first partial model of EmMbCN was built into several 6.0 Å–3.0 Å resolution omit difference maps. The resulting R-factor was 36.2% for 2σ data in the above resolution range. This initial atomic model was further improved with refinement utilizing simulated annealing within the program X-PLOR (27) employing the slow cooling protocol suggested by Weis *et al.* (28). After the first round, the R-factor improved to 19.9%. In total, six rounds of simulated annealing, including phase expansion to 2.0 Å and individual B-factor refinement, were carried out. Water molecules were included at a resolution of 2.3 Å. The criteria for the assignment of a solvent molecule was that the corresponding density in a $F_o - F_c$ map had to be at least contoured at a 2σ level and 1σ for a $2F_o - F_c$ map. In addition, the potential water molecules had to be at reasonable hydrogen bonding distances either to a protein molecule or to another solvent molecule. At a resolution of 2.0 Å, the cyanide ligand was modelled when the R-factor had dropped to 17.9%. Further refinement was performed by restrained least squares refinement to the final resolution range of 1.78–8.0 Å, using the program PROLSQ/PROTIN (29). During the whole refinement procedure, no restraints were imposed on the Fe-N bonds of the heme moiety and the C-N bond of the distal ligand. The refinement program package X-PLOR was run on a Cray YMP/464, and the CCP4 version of PROLSQ/PROTIN was implemented on a VAX 6000. Fitting of the atomic model to the electron density maps was performed on a Silicon Graphics 4D/220 GTX using the program CHAIN (30). X-PLOR requires parameter and topology files to define the standard geometry and the energy parameters of amino acid residues (31). The parameter files include information needed to calculate the empirical functions during simulated annealing and energy minimization. The topology files describe atom types, bond angles, covalent bonds, etc. In the standard protein topology file (TOPH19X.PRO), the histidine entries have been removed. Instead, single protonated histidines from TOP19X.HEME were used. In addition, aliphatic hydrogen atoms were excluded from calculations and their contributions were incorporated into the energy parameters of the carbon atoms. The peptide bonds of the polypeptide chain are defined in a second topology file (TOPH19.PEP). The charges of the amino acid residues Arg, Lys, Asp, and Glu were always turned off during simulated annealing and energy

TABLE I
Final refinement statistics of EmMbCN

Diffraction data			
Resolution	8.0–1.78 Å		
Number of reflections ($F > 2\ \sigma$)	12,966		
R-factor[a]	0.153		
Total number of atoms	1466		
Stereochemistry (root mean square deviations from ideality)			
		Delta[b]	Sigma[c]
Bonding distancess (weight = 2.00)			
Bonding length (1–2 neighbor)		0.019 Å	0.020 Å
Angle related distance (1–3 neighbor)		0.039 Å	0.030 Å
Intraplanar distance (1–4 neighbor)		0.051 Å	0.040 Å
Planar groups (weight = 2.00)			
Deviation from plane		0.017 Å	0.020 Å
Chiral centers (weight = 2.00)			
Chiral volume		0.148 Å^3	0.130 Å^3
Nonbonded contacts (weight = 1.00)			
No crystallographic symmetry			
Single torsion		0.192 Å	0.300 Å
Multiple torsion		0.197 Å	0.300 Å
Possible hydrogen bond		0.244 Å	0.300 Å
Crystallographic symmetry operations			
Multiple torsion		0.219 Å	0.300 Å
Possible hydrogen bond		0.290 Å	0.300 Å
Torsion angles (weight = 1.00)			
Planar (0, 180)		3.141°	3.0°
Staggered (±60, 180)		16.973°	15.0°
Orthonormal (±90)		38.438°	20.0°
Thermal factors (weight = 1.00)			
Main chain bond (1–2 neighbor)		1.801 Å^2	1.000 Å^2
Main chain angle (1–3 neighbor)		2.579 Å^2	1.500 Å^2
Side chain bond		3.502 Å^2	1.000 Å^2
Side chain angle		5.518 Å^2	1.500 Å^2

[a] $R = \Sigma\ ||F_o| - |F_c||/\Sigma\,|F_o|$, where F_o stands for the observed structure factor amplitudes and F_c for calculated structure factor amplitudes.
[b] This column shows the observed σ value.
[c] This column shows the target σ value.

minimization. Before the molecular dynamic simulations were started, energy minimization calculations were performed to relax the structure of the model and to obtain a good geometry without strain. The "repel" nonbonded energy function was used to remove close contacts in the structure. Subsequent to each refinement step, the ideality of the resulting structure was checked and a root mean square deviation between the original and the refined structure was calculated using routines present in the X-PLOR program. Molecular dynamic simulations were performed using the dynamics Verlet method implemented in X-PLOR (32). Initial atomic velocities were obtained from Maxwellian distribution. After refinement, electron density maps ($2F_o - F_c$ and $F_o - F_c$) were calculated with the CCP4 program package (26).

RESULTS

Accuracy and Completeness of the Model—The model of EmMbCN was refined to an R-factor of 15.3% including reflections for $F > 2\sigma$ in the resolution range 8.0 Å–1.78 Å with a good agreement between stereochemistry and ideality (Table I). Only three amino acids, two of them being glycine residues, are found in the forbidden zone of a Ramachandran plot (Fig. 1). Most of the residues cluster within the region expected for an α-helical structure. An estimate for the upper limit of the root mean square error of a crystal structure is provided by a Luzzati plot (33). For EmMbCN, this value is about 0.1 Å.

The whole polypeptide chain of 153 amino acids, 1 protoporphyrin IX molecule, 1 cyanide ligand, and 255 water molecules were included in the model. During map interpretation, it became obvious that the originally reported sequence (8) is erroneous at least for two amino acids. Of most significance here is that the reported Leu^{29}(B10) is rather a phenylalanine as can unambiguously be seen from a difference omit map (Fig. 2). Also, residue 27 (B8), a phenylalanine according to Dene *et al.* (8), is now assigned to a threonine. The difference omit map for this residue (Fig. 3) shows density for either a valine or a threonine. The proximity of the threonine hydroxyl group to a carbonyl atom and to a neighboring water molecule, permitting

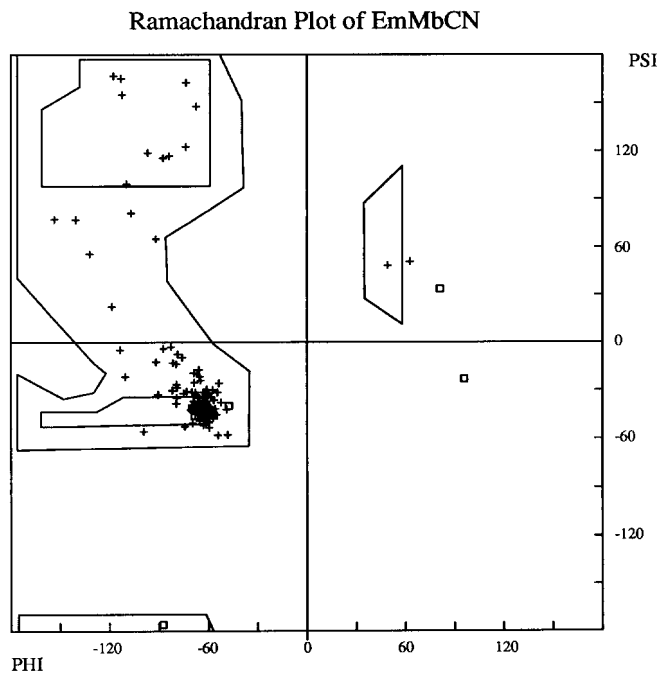

FIG. 1. **Plot of the main chain ϕ-ψ angles for Asian elephant cyano-metmyoglobin.** Glycine residues are indicated by *squares*.

the formation of hydrogen bonds, imposes the assignment of the latter residue type. A few other instances were encountered where the side chain density was very diffuse or absent for some of the terminal side chain atoms. Whether this is due to disorder or additional sequencing errors could not be decided, especially since these side chains belong to large, charged

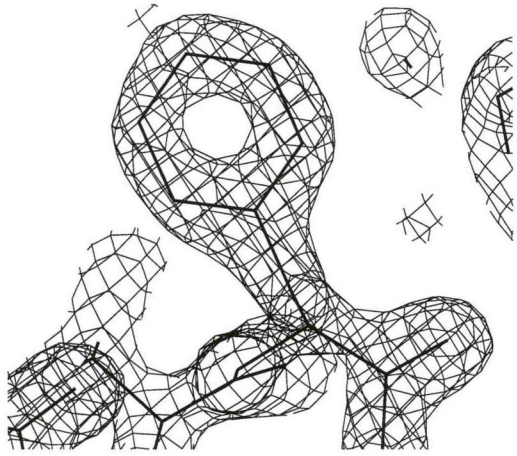

FIG. 2. **Phe[29](B10) in its electron density map ((2 $|F_o|$ − $|F_c|$) α_c) contoured at 2.0 standard deviations.** The contribution of this residue for the phase calculation was omitted.

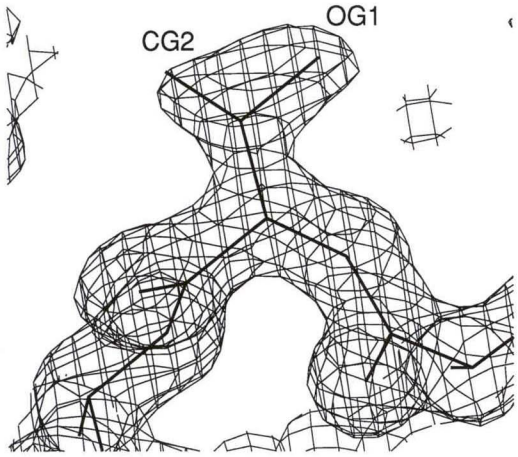

FIG. 3. **Thr[27](B8) in its electron density map ((2 $|F_o|$ − $|F_c|$) α_c) contoured at 1.5 standard deviations.** The contribution of this residue for the phase calculation was omitted.

amino acids, like Lys, which are situated at the surface of the protein.

Overall Description of the Model—The overall fold of EmMbCN is very similar to that of other myoglobins (Fig. 4). Eight helical segments are folded in a globular structure with a hydrophobic pocket which host the coordinated prosthetic heme group, as found in other members of the globin family. A superposition (Fig. 5) of the C_α atoms of EmMbCN with the aquo-met-SWMb model, the structure used as search model during molecular replacement, shows that the root mean square deviation for the two structures is 0.59 Å. Similarly, small values are also observed by comparison with other myoglobin models suggesting a very high structural conservation. This had to be expected from the high sequence identity between elephant and sperm whale Mb (81.7%). Differences larger than the mean are found in the N-terminal part of the F-helix. This region forms one face of the heme pocket and shows the highest temperature factors (Fig. 6), besides the N- and C-terminal residues, indicating a mobile area in myoglobin. In addition, large root mean square deviations between the two myoglobin structures are seen in a loop preceding the H-helix, also characterized by relatively high temperature

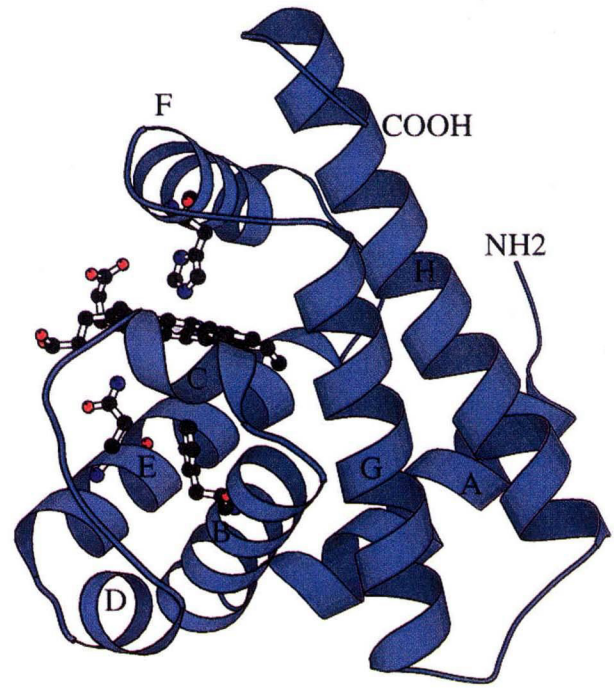

FIG. 4. **Ribbon plot showing the fold of Asian elephant myoglobin.** α-Helices are represented by *spirals* and labeled with *uppercase letters*. This plot was created using the program MOLSCRIPT (43).

factors. Comparison of EmMbCN with other myoglobin structures reveals a much similar pattern of the overall *B*-factor distribution.

Active Site—A stereo picture of the heme group and the surrounding environment with its corresponding electron density is depicted in Fig. 7. The iron of the heme is displaced by 0.06 Å from the least squares plane formed by the four pyrrole nitrogen atoms, in agreement with other ligated myoglobin structures (34) (Table II). The sixth coordination site of the iron is occupied by the carbon atom of the cyanide ion with a Fe-C distance of 1.85 Å. The Fe-C-N angle is 177°, the cyanide group being tilted from the heme normal by 7.5°. A rather close distance of 2.96 Å is found between the N atom of the cyanide and the $C_\gamma 2$ atom of Val[68](E11).

Despite the two amino acid replacements (H64Q and L29F) in the heme pocket, the conformational changes are small if compared to aquo-met-SWMb. High structural similarity is found for those residues which are identical (Fig. 8). Even in the case of the distal residue 64 (E7) where a glutamate substitutes the commonly found histidine, the length of the hydrogen bonds formed by the $N_\epsilon 2$ atom and the N or O counteratom, respectively, of the ligand differs only marginally (2.87 Å *versus* 2.78 Å). Therefore, it appears that His and Gln stabilize the bound ligand in a similar manner. On the other hand, Lys[45](CD3) in EmMbCN being replaced by an Arg in SWMb forms a hydrogen bond between its N_ζ atom and the $O_\epsilon 1$ atom of the distal residue Gln[64](E7). No hydrogen bond between Arg[45](CD3) and His[64](E7) is found in SWMb. Residue Arg[45](CD3) in the carbonmonoxy sperm whale myoglobin x-ray structure shows two conformational states (34). We did not detect any evidence for an additional conformation of the Lys[45](CD3) side chain in the electron density.

The most significant finding in our crystal structure concerns the amino acid replacement of Leu at position 29 (B10) by a phenylalanine which causes drastic steric restraints on the movement of the ligand. This residue is situated in a hydrophobic patch formed by Phe[29](B10), Leu[32](B13), Phe[33](B14), Leu[40](BC5), Phe[43](CD1), Phe[46](CD4), Leu[61](E4), Val[68](E11),

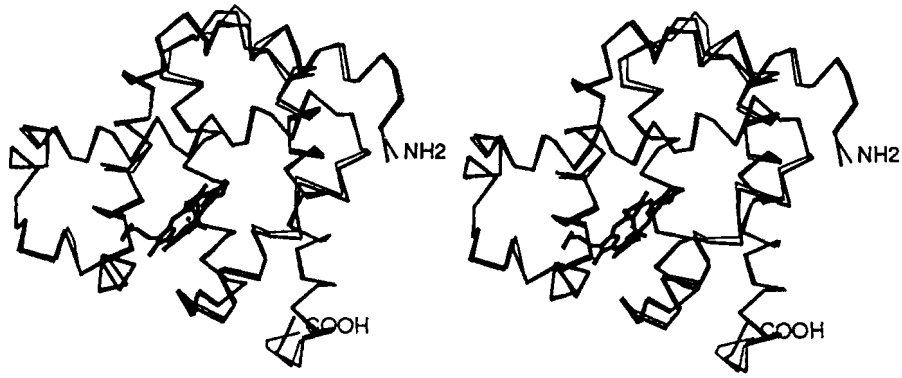

FIG. 5. **C_α stereo plot of EmMbCN (*thick lines*) superimposed with SWMb (*thin lines*).**

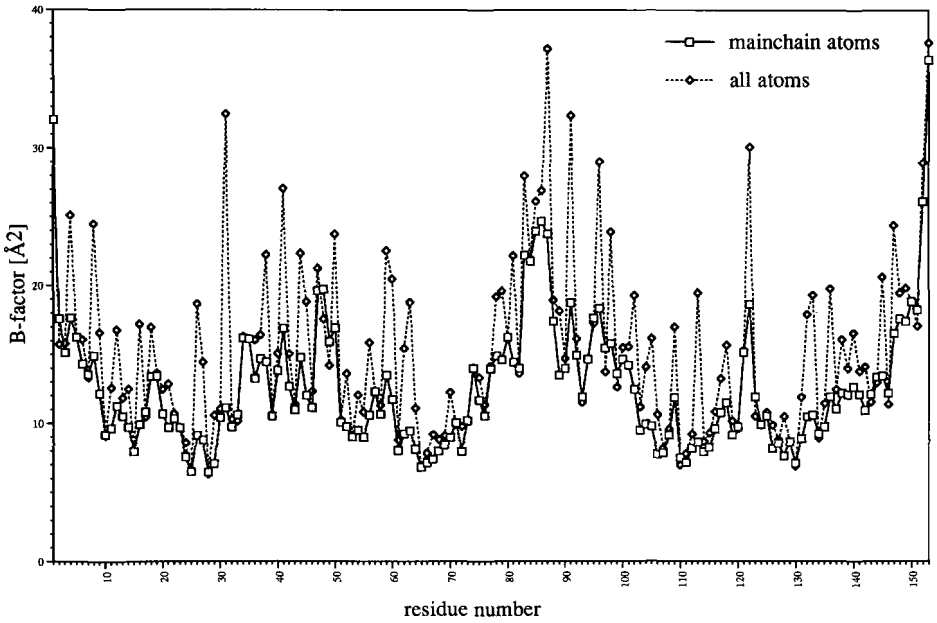

FIG. 6. ***B*-factor distribution of Asian elephant cyano-metmyoglobin.** Presented are the mean isotropic temperature factors for the main chain atoms and for all atoms of a residue.

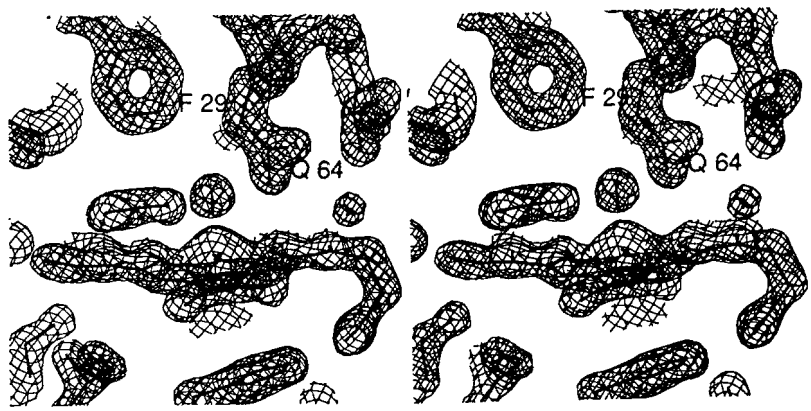

FIG. 7. **Electron density map ((2 $|F_o|$ − $|F_c|$) α_c) of the heme pocket of Asian elephant myoglobin.** Gln^{64}(E7) forms a hydrogen bond to the carbon atom of the CN ligand. The ligand is 3.5 Å away from Phe^{29}(B10).

and Ile^{107}(G8) (Fig. 9). The side chain of Phe^{29}(B10) penetrates into the distal side of the heme pocket approaching the CN ligand up to 3.5 Å. This situation causes the crevice to be more crowded than in myoglobins having a leucine at position 29 (B10).

A comparison with the carbonmonoxy structure of sperm whale myoglobin (Brookhaven Protein Data Bank entry 1MBC), where the Fe^{2+}-CO unit is iso-structural with the Fe^{3+}-CN unit in EmMbCN, reveals even more structural similarities. The proximal sides show virtually no differences in the two structures, especially since the proximal histidines are practically superimposed. On the distal side, the $N_\epsilon 2$ atom of Gln^{64}(E7) in EmMbCN has moved 0.9 Å away from the corresponding nitrogen atom of His^{64}(E7) in SpWCO toward the propionate of the A-pyrrole and at the same time closer to the heme plane (Fig. 10). Even so, the combined movement results in an unchanged distance between the heme iron and the $N_\epsilon 2$ atom in both structures (Table II). In addition, the distances between the two $N_\epsilon 2$ atoms of the distal residues and the ligand atoms N or O, respectively, are virtually the same due to the different ligand bent angles.

A more interesting and appropriate myoglobin structure for comparison is the double mutant H64Q/L29F of SWMb due to the presence of the same distal side residues as in EmMbCN.

TABLE II
Heme-ligand geometries

Angled brackets indicate a mean value averaged over several atoms.

	Distance (Å) or angle (°)				
	EmMbCN	CO-L29F/H64Q	CO-SWMb	CO[a]-SWMb	Deoxy-SWMb
Proximal residue					
Fe-N_ϵ2	2.11	2.29	2.17	2.25	2.20
Distal residue					
Fe-N_ϵ2	4.83	4.64	4.80	4.65	4.46
N_ϵ2-ligand N (respectively O)	2.87	3.25	2.88	3.90[b]	2.78[c]
N_ϵ2-ligand C	3.45	3.37	3.22	5.01[b]	
Residue 29					
C_ζ (respectively CD2)-ligand N (respectively O)	3.53	3.17	4.40	3.79[b]	5.67[c]
C_ζ (respectively CD2)-Fe	6.20	5.91	8.32	7.26	7.34
Ligand					
IR angle[d]	5.9	12.9	32	91[b]	
Bend angle[e]	177	162	160	111[b]	
Tilt angle[f]	7.5	9.2	3.9	39[b]	17.9[c]
Fe-C	1.85	1.85	1.91	4.14[b]	
Fe-N (respectively O)	2.98	2.92	2.94	3.60[b]	2.15[c]
Heme					
Fe-⟨N_p plane⟩	0.06	0.05	0.04	0.19	0.25
⟨Fe-N_p⟩	1.96	1.95	1.96	1.98	2.03
⟨Np-N_p⟩	2.82	2.76	2.78	2.78	2.84

[a] This refers to the photolyzed form of the ligand.
[b] Assumes the O of CO is nearest the iron.
[c] These entries refer to the O of a water molecule at the active site.
[d] The IR angle lies between the C-N (respectively C-O) bond and the normal to the heme plane.
[e] The bend angle is between the iron, the nearer ligand atom, and the farther ligand atom.
[f] The tilt angle lies between the line Fe-ligand and the normal to the mean heme plane.

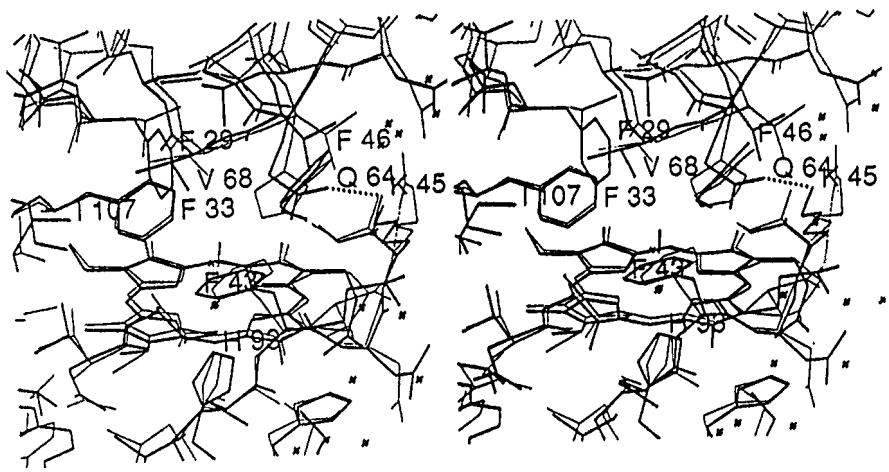

FIG. 8. **Stereo picture of the superposition of the heme pocket of Asian elephant cyano-metmyoglobin (*thick lines*) and sperm whale aquo-metmyoglobin (*thin lines*).** The distance between the N_ϵ2 atom of the distal residue 64 (E7) and the ligand is very similar. Residue 45 (CD3) being a lysine in EmMbCN forms a hydrogen bond (*dashed line*) to atom O_ϵ1 of the distal residue. Arg^{45}(CD3) in SWMb instead forms a hydrogen bond (*dashed line*) to a heme propionate group.

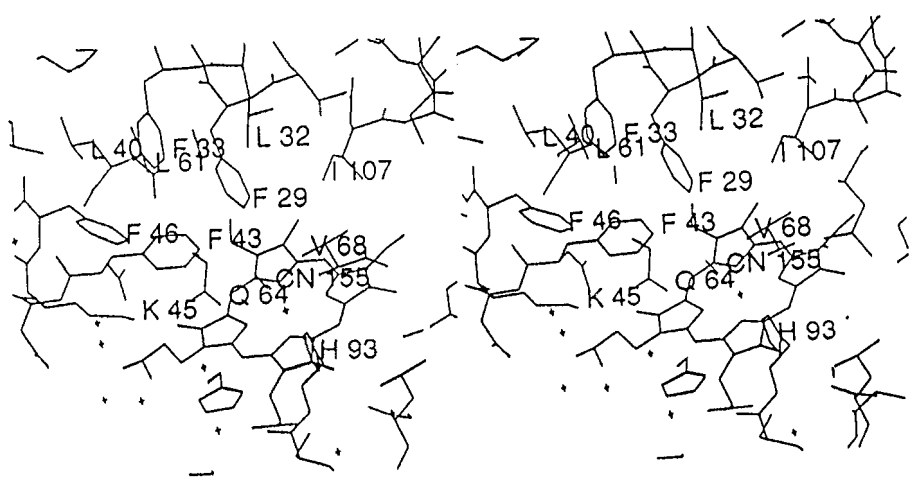

FIG. 9. **Stereo picture of the heme pocket in Asian elephant myoglobin.** The hydrophobic residues forming the upper part of the distal side are depicted. Phe^{29}(B10) penetrates further into the heme pocket than leucine which is usually found at this position in myoglobins.

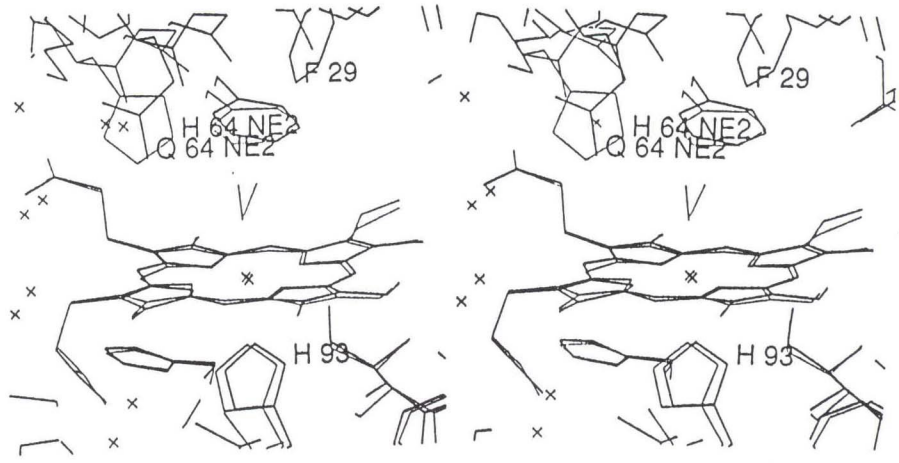

FIG. 10. **Stereo view of the superposition of the heme environment of Asian elephant myoglobin (*thick lines*) and sperm whale CO myoglobin (*thin lines*).** The $N_\epsilon 2$ atom of residue 64 (E7) has moved with respect to the corresponding nitrogen position in the CO complex of sperm whale myoglobin. Even so, the distance to the heme iron remains unchanged.

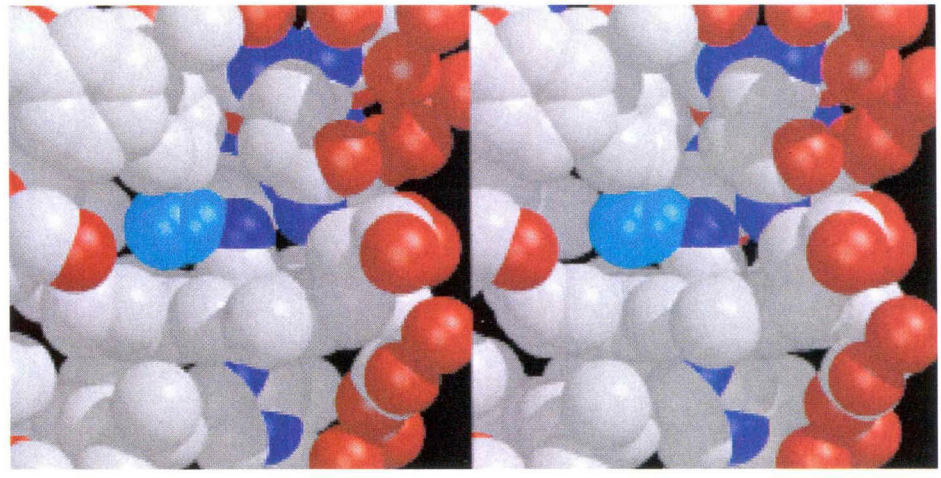

FIG. 11. **Stereo picture of a space-filling model of the photodissociated carbonmonoxy ligand of sperm whale myoglobin in the environment of Asian elephant myoglobin.** The fictitious CO ligand is colored in *cyan*. The dissociating CO ligand collides with the side chain of Phe^{29}(B10).

The carbonmonoxy mutant crystal structure has been determined recently and was kindly supplied to us by Dr. George Phillips at Rice University, Houston, TX. Zhao *et al.* (44) describe some functional and spectroscopic properties of the double mutant and of the two single mutants (L29F and H64Q) and compare them with those of the wild type and EMb. Residues Gln^{64}(E7) and Phe^{29}(B10) have been tested earlier by site-directed mutagenesis for their effects on ligand binding and autooxidation (11, 14, 35, 36). The models of EmMbCN and of the double mutant were superimposed revealing a very similar structure. The two Phe^{29}(B10) practically lie on top of each other while the Gln^{64}(E7) residues show marginal differences in their positions. In both myoglobin structures, the tilt angles of the ligands are very much alike (Table II). The bend angle is smaller in the H64Q/L29F model with respect to the situation in EmMbCN reducing the distance to Phe^{29}(B10) in the former structure.

Of special interest regarding ligand binding and dissociation is the recently published crystal structure of the photolyzed carbonmonoxy-SWMb at 40 K (16). This structure (the coordinates of the CO-photolyzed, CO-ligated, and deoxy-SWMb structures at 40 K were kindly provided to us by Joel Berendzen, Los Alamos National Laboratory, Los Alamos, NM) allowed us to investigate the relative position of Phe^{29}(B10) in EmMbCN with respect to a fictitious dissociated ligand as in sperm whale myoglobin. Fig. 11 depicts a space-filling model of the photodissociated carbonmonoxy ligand of SWMb in the EmMbCN environment which shows the dissociated ligand colliding with the side chain of Phe^{29}(B10). This indicates that, besides electrostatic effects on the Fe-C bond, an additional steric interaction, introduced by Phe^{29}(B10), influences ligand binding. Particularly relevant in this respect are the results of photolysis measurements and molecular dynamic simulations performed by Gibson *et al.* (37) on mutagenized SWMb at position 29 (B10). They show that this residue strongly influences geminate ligand recombination.

DISCUSSION

The originally reported primary sequence of Asian elephant myoglobin by Dene *et al.* (8) has caused some confusion about the structure-function relationships concerning the distal side residues of myoglobins. Despite the replacement of the essential His residue at position 64 (E7) by a Gln, this protein displays essentially unchanged O_2 binding and autooxidation behavior. Based on the assumption that only this single substitution has taken place on the distal side and the fact that a H64Q point mutant of sperm whale myoglobin exhibits much lower oxygen affinity and greater susceptibility against autooxidation, the findings above were not explainable. ^{1}H NMR investigations of the heme cavity (12) have been interpreted in such a way that even a dramatic rearrangement of the CD corner, a movement which brings the conserved residue Phe^{46}(CD4) in a position only about 5 Å away from the heme iron, has been postulated. Other investigations, including site-directed mutations on sperm whale myoglobin, have been undertaken which finally led to the assumption that a second, so far unidentified, amino acid replacement had to be considered, but a definite answer could not been given. All these attempts to explain the peculiar functional properties of EMb were ob-

viously not sufficient to construct a sound picture of the distal side pocket.

The crystal structure of Asian elephant myoglobin was determined to resolve this confusion. Recently, the presence of an additional phenylalanine has been suspected based on NMR data (13) which has been confirmed by our work. The assumption that introducing an additional large hydrophobic side chain into the active site causes more crowding was confirmed by our structure, but not the assumed structural perturbation of the standard globin fold (12). Superpositioning of the EmMbCN model with other myoglobins reveals only minor alterations of the backbone. These deviations appear to be structurally negligible and are probably caused by different crystal contacts (38). It appears that individual ligand binding properties are the result of local amino acid substitutions rather than a modulated fold in EMb.

The CN ligand in EmMbCN possesses an almost perpendicular orientation with respect to the heme plane, whereas the bend angles (Table II) of the CO ligands in the wild type SWMb and the H64Q/L29F double mutant are nearly identical and deviate significantly from 180°, suggesting that its orientation is directed by the chemical nature of the ligand. Furthermore, the comparison between the elephant myoglobin and the double mutant allows us to draw the conclusion that the different position of the residue 64 (E7) $N_\epsilon 2$ atoms in EmMbCN and SWMbCO are not caused by the hydrogen bond between Gln^{64}(E7) and Lys^{45}(CD3).

We report here on the first cyano-met complex of a native myoglobin solved by x-ray crystallography. The unfavorable contact between the CN ligand and the highly conserved Val^{68}(E10) is imposed by the almost perpendicular orientation of the ligand with respect to the heme plane. Such a close contact is not found in the CO and O_2 complex of sperm whale myoglobins which might indicate that the role of residue 68 (E10) is to discriminate between various ligands.

On the distal side, the additional Phe residue could be located at position 29 (E10) as unequivocally interpreted from its corresponding electron density. This residue penetrates into the distal side, pointing toward the heme iron which causes crowding, therefore hindering ligand access.

Implications for the Low Temperature Behavior—The very complex CO binding kinetics of EMb at cryogenic temperatures reported by Cupane *et al.* (15) were tentatively explained by the concept of a more crowded heme pocket in this protein compared to other myoglobins. This interpretation was based on the then available NMR data relative to the cyano-met derivative of EMb (12, 13), pointing to the presence of an additional phenylalanine at a small enough distance from the bound ligand to allow van der Waals interactions. The imperfect primary structure of EMb reported by Dene *et al.* (8) did not allow an identification of the additional phenylalanine detected by the group of La Mar (12, 13). Therefore, the assumptions made by Cupane *et al.* (15), although consistent with the experimental results, remained at a speculative level. The solution of the crystal structure of EMb allows now a more detailed interpretation of its low temperature kinetic behavior. Two of the four kinetic processes, detected in the time courses for CO binding to EMb below 200 K, have been attributed to the interaction between the photolyzed CO and a then generic distal residue. Fig. 11 shows that a very likely situation immediately after photolysis is the formation of a "special pair" between the CO molecule and Phe^{29}(B10). Under these conditions, the ligand cannot escape too far from the heme iron state, and recombination must occur very rapidly, as indeed observed (process I in Cupane *et al.* (15)). The stability of the special pair is expected to diminish with rising temperatures, due to progressively larger vibrations of the phenyl ring. This explains why process I dominates at very low temperatures, with a maximal amplitude at ~100 K and disappears completely at ~200 K. The second kinetic process (II in Cupane *et al.* (15)), scarcely populated and detected only up to ~100 K, can now be attributed to a stronger, relative to that in the special pair, but less probable interaction of the photolyzed ligand with Phe^{29}(B10) which can only take place as long as its aromatic ring is practically immobile. The above interpretation is also in agreement with the interaction of the photolyzed CO with Phe π electrons proposed by Straub and Karplus (39), on the basis of molecular dynamic simulations, to explain the red shift of the CO stretching frequency in the photolyzed relative to the bound state. The "dynamic gate" proposed by Cupane *et al.* (15) to explain process IV, populated only between ~110 K and ~280 K with its highest intensity at ~220 K, can be ascribed to an exchange of the photolyzed CO between Phe^{29}(B10) and (an)other phenylalanine side chain(s) within the same hydrophobic patch. From there, the ligand could either bounce back in the vicinity of the heme iron and bind to it or diffuse away. A final point that can now well be rationalized is the low entropic contribution to the activation barrier for the only geminate CO binding step (process III) observed throughout the 40 K to 300 K range examined by Cupane *et al.* (15). According to Frauenfelder and Wolynes (40), this is due to the small free volume available to the photolyzed ligand at the distal side of the heme pocket as a result of the L29F substitution.

Interestingly enough, a situation remarkably similar to that observed at cryogenic temperatures in the time courses for CO recombination to EMb after photolysis, is also present in lignin peroxidase, where again a Phe residue is present at the distal site, in van der Waals contact with the bound ligand (7, 41). Further support to the above interpretation of the low temperature CO binding kinetics comes from the results obtained by Gibson *et al.* (37) on SWMb mutants in which Leu^{29}(B10) is replaced by Ala, Val, or Phe. Based on laser photolysis measurements of the NO and O_2 derivatives on picosecond and nanosecond time scales, these authors have shown that the extent and rate of geminate ligand recombination is determined by the size of residue 29 (B10). Furthermore, as recently directly observed by x-ray crystallography at 40 K (16) and depicted in Fig. 11, the molecular dynamic simulations (37) show that immediately after photolysis the ligand moves toward residue 29 (B10). This implies that motion of the photolyzed ligand away and toward the heme iron is determined by this residue.

Relationship between Structural and Functional Properties at Physiological Temperature—The rather similar kinetic behavior of EMb at 293 K in comparison with other mammalian myoglobins, like those form horse, pig, sperm whale, and human, can be rationalized by steric and stabilizing electrostatic effects resulting from the combination of the two point mutations H64Q and L29F. Assuming a similar orientation and position of a O_2 ligand in EMb as present in oxy-SWMb (42), the hydrogen bond between $N_\epsilon 2$ of Gln^{64}(E7) and the oxygen would be about 3.8 Å compared to 2.77 Å in the $SWMbO_2$ structure. The much weaker hydrogen bonding between oxygen and the distal Gln is largely compensated by favorable electrostatic interactions with the positive edge of the phenylalanine Phe^{29}(B10) multipole and the displacement of water molecules by this large side chain. In the H64Q recombinant SWMb structure, the weak hydrogen bonding properties of Gln^{64}(E7) would produce a reduced discrimination of O_2 and CO binding against water and therefore result in an increased autooxidation rate. This effect is compensated in EMb with respect to the situation in SWMb as a consequence of the water-excluding

steric effect of Phe^{29}(B10) producing a slightly higher autooxidation rate. The slow CO dissociation rate displayed by EMb can also be rationalized on the basis of a stabilizing interaction between Phe^{29}(B10) and the bound CO. For the kinetic data at room temperature used in our discussion, we refer to the accompanying paper by Zhao *et al.* (44).

The structure determination of Asian elephant myoglobin has clarified the situation of the heme pocket in this peculiar protein which has caused much confusion in the past. Obviously, nature is able to achieve solutions for a given biological problem by different approaches. It appears that even a conserved amino acid composition in an active site environment does not have to be the ultimate solution to achieve a certain vital protein function. An interesting result of this study is also that various chemical changes do not have necessarily the effect of overall structural alterations. Rather, the combination of several local chemical modifications leaves the functional properties of a protein conserved.

Acknowledgments—We thank the Dept. of Pathology of the Veterinary Clinic, canton Zürich, for supplying us with skeletal muscle from a deceased elephant of the local zoo. We are indebted to the groups of George Phillips and John Olson for providing the structure of the H64Q/L29F SWMb mutant and to Joel Berendzen for giving us the atomic coordinates of the low temperature SWMb structure in the deoxy, CO, and photolyzed states. Thomas Choinowski is thanked for helpful discussion concerning protein purification and crystallization.

REFERENCES

1. Millikan, G. A. (1939) *Physiol. Rev.* **19,** 503
2. Wittenberg, J. B., and Wittenberg, B. A. (1990) *Annu. Rev. Biophys. Chem.* **19,** 217–241
3. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) *Nature* **181,** 662–666
4. Royer, W. E., Jr., Hendrickson, W. A., and Chiancone, E. (1989) *J. Biol. Chem.* **264,** 21052–21061
5. Steigemann, W., and Weber, E. (1979) *J. Mol. Biol.* **127,** 309–338
6. Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. T. (1960) *Nature* **185,** 416–422
7. Piontek, K., Glumoff, T., and Winterhalter, K. (1993) *FEBS Lett.* **315,** 119–124
8. Dene, H., Goodman, M., and Romero-Herrera, A. (1980) *Proc. R. Soc. Lond. B Biol. Sci.* **207,** 111–127
9. Brantley, R. E., Jr., Smerdon, S. J., Wilkinson, A. J., Singleton, E. W., and Olson, J. S. (1993) *J. Biol. Chem.* **268,** 6995–7010
10. Romero-Herrera, A. E., Goodman, M., Dene, H., Bartnicki, D. E., and Mizukami, H. (1981) *J. Mol. Evol.* **17,** 140–147
11. Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., and Olson, J. S. (1989) *J. Biol. Chem.* **264,** 3057–3060
12. Yu, L. P., La Mar, G. N., and Mizukami, H. (1990) *Biochemistry* **29,** 2578–2585
13. Vyas, K., Rajarathnam, K., Yu, L. P., Emerson, S. D., La Mar, G. N., Krishnamoorthi, R., and Mizukami, H. (1993) *J. Biol. Chem.* **268,** 14826–14835
14. 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
15. Cupane, A., Leone, M., Vitrano, E., Cordone, L., Hiltpold, U. R., Winterhalter, K. H., Yu, W., and Di Iorio, E. E. (1993) *Biophys. J.* **65,** 2461–2472
16. Schlichting, I., Berendzen, J., Phillips, G. N., Jr., and Sweet, R. M. (1994) *Nature* **371,** 808–812
17. Wittenberg, J., and Wittenberg, B. (1981) *Methods Enzymol.* **76,** 29–54
18. Bisig, D. A. (1993) *Crystal Structure Analysis of Indian Elephant Myoglobin at 2.0 Å Resolution.* Diploma thesis, Swiss Federal Institute of Technology, Zürich
19. Mc Pherson, A. (1982) *Preparation and Analysis of Protein Crystals*, Wiley, New York
20. Jancarik, J., and Kim, S. (1991) *J. Appl. Crystallogr.* **24,** 409–411
21. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
22. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 67–71
23. Fitzgerald, P. (1988) *J. Appl. Crystallogr.* **21,** 273–278
24. Crowther, R. A. (1972) *The Molecular Replacement Method* (Rossmann, M. G., ed) pp. 173–178, Gordon and Breach, New York
25. Crowther, R., and Blow, D. (1967) *Acta Crystallogr.* **23,** 544–548
26. *SERC Collaborative Computational Projects in Crystallography*, Daresbury Laboratory, Daresbury, Warrington, UK
27. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science* **235,** 458–460
28. Weis, W. I., Brünger, A. T., Skehel, J. J., and Wiley, D. C. (1990) *J. Mol. Biol.* **212,** 737–761
29. Hendrickson, W. (1985) *Methods Enzymol.* **115,** 252–270
30. Sack, J. S. (1988) *J. Mol. Graphics* **6,** 224–225
31. Brünger, A. T. (1990) *X-PLOR, version 2.1, A System for Crystallography and NMR*, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT
32. Brünger, A. T. (1992) *X-PLOR, version 3.0, A System for Crystallography and NMR*, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT
33. Luzzati, V. (1952) *Acta Crystallogr.* **5,** 802–810
34. Kuriyan, J., Wilz, S., Karplus, M., and Petsko, G. A. (1986) *J. Mol Biol.* **192,** 133–154
35. Springer, B. A, Sligar, S. G., Olson, J. S., and Phillips, G. N., Jr. (1994) *Chem. Rev.* **94,** 699–714
36. Quillin, M. L., Arduini, R. M., Olson, J. S., and Phillips, G. N., Jr. (1993) *J. Mol. Biol.* **234,** 140–155
37. Gibson, Q. H., Regan, R., Elber, R., Olson, J. S., and Carver, T. E. (1992) *J. Biol. Chem.* **267,** 22022–22034
38. Phillips, G. N., Jr., Arduini, R. M., Springer, B. A., and Sligar, S. G. (1990) *Proteins Struct. Funct. Genet.* **7,** 358–365
39. Straub, J. E., and Karplus, M. (1991) *Chem. Phys.* **158,** 221–248
40. Frauenfelder, H., and Wolynes, P. G. (1985) *Science* **229,** 337–345
41. Glumoff, T. (1991) *On the Structure and Properties of Lignin Peroxidase: A biochemical and crystallographic study.* Ph.D. thesis, Swiss Federal Institute of Technology, Zürich
42. Phillips, S. E. V. (1980) *J. Mol. Biol.* **142,** 531–554
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
44. Zhao, X., Vyas, K., Nguyen, B. D., Rajarathnam, K., La Mar, G. N., Li, T., Phillips, G. N., Jr., Eich, R. F., Olson, J. S., Ling, J., and Bocian, D. F. (1995) *J. Biol. Chem.* **270,** 20763–20774