Anatomy and Dynamics of a Ligand-Binding Pathway in Myoglobin: The Roles of Residues 45, 60, 64, and 68

ARTICLE *in* BIOCHEMISTRY · JUNE 1994

Impact Factor: 3.02 · DOI: 10.1021/bi00184a021 · Source: PubMed

4 AUTHORS, INCLUDING:

Steven G Boxer
Stanford University
96 PUBLICATIONS **6,230** CITATIONS

SEE PROFILE

Anatomy and Dynamics of a Ligand-Binding Pathway in Myoglobin: The Roles of Residues 45, 60, 64, and 68[†]

David G. Lambright, Sriram Balasubramanian, Sean M. Decatur, and Steven G. Boxer*

Department of Chemistry, Stanford University, Stanford, California 94305-5080

Received September 13, 1993; Revised Manuscript Received March 3, 1994[⊗]

ABSTRACT: In order for diatomic ligands to enter and exit myoglobin, there must be substantial displacements of amino acid side chains from their positions in the static X-ray structure. One pathway, involving Arg/Lys45, His64, and Val68, has been studied in greatest detail. In an earlier study (Lambright et al., 1989) we reported the surprising result that mutation of the surface residue Lys45 to arginine lowers the inner barrier to CO rebinding. Until then, it had been thought that this barrier primarily involves interior distal pocket residues such as His64 and Val68. In this report, we present a detailed study of the CO rebinding kinetics in aqueous solution of a series of single- and double-site mutants of human myoglobin at positions 64, 68, 45, and 60. On the basis of the observed kinetics, we propose that the effect of surface residue 45 on the inner barrier can be explained by a chain of interactions between surface and pocket residues. Very large, and in some cases unexpected, changes are observed in the kinetics of recombination and in the partitioning between geminate and bimolecular recombination.

The binding of diatomic ligands to myoglobin not only is a simple and well-characterized biological reaction but also is among the best-studied general models to probe the coupling of protein dynamics and function. The crystal structures for deoxyMb[1] (Takano, 1977b), MbO_2 (Phillips, 1980), and MbCO (Kuriyan et al., 1986) show that the ligand-binding site is buried in the tightly packed interior of the protein, so there are no permanently open channels to the surface of the protein large enough to accommodate a diatomic ligand. Thus, the migration of ligands between solvent and the binding site at the heme iron must be facilitated by fluctuations in the positions of protein atoms.

The development of bacterial expression systems for recombinant myoglobin from several species (Varadarajan et al., 1985; Springer and Sligar, 1987; Dodson et al., 1988) has enabled the study of the impact of specific amino acids on the ligand-binding reaction. Much of this effort has focused on four highly conserved residues which are believed to play a key role: His64, Val68, Asp60, and Lys45 (Arg in sperm whale myoglobin). Historically, these residues have been implicated by several lines of evidence. From examination of the X-ray crystal structure of myoglobins, the shortest path from pocket to solvent is between His64 and Val68. The escape barrier through this pathway has been estimated to be greater than 90 kcal/mol if these residues maintained the positions observed in the X-ray structure (Case & Karplus, 1979). However, this barrier can be lowered by the movement of His64 and Val68 toward solvent, displacing residue 45 and the heme propionate side chain. Furthermore, in an X-ray crystal structure of sperm whale Mb with a bulky phenyl group as the sixth ligand, only Arg45, His64, and Val68 were significantly displaced (Ringe et al., 1984). This crystal structure suggests a model for an "open" conformation of the distal heme pocket in which His64 rotates out toward solvent and Arg45 disrupts hydrogen bonding with the heme 6-propionate side chain to open up a channel. A similar open conformation is observed in crystal structure of Mb with imidazole (Bolognesi et al., 1982) and ethyl isocyanide (Johnson et al., 1989) bound to the heme iron. Although these crystal structures do not provide direct evidence that the binding of diatomic ligands is coupled to such amino acid fluctuations, they offer a basis for hypotheses concerning the roles of residues 45, 64, and 68.

In an earlier paper (Lambright et al., 1989) we compared proton-exchange and ligand-rebinding kinetics of human myoglobin mutants K45R and D60E to those of the wild-type protein. While sperm whale myoglobin has an Arg at position 45 which is within hydrogen-bonding distance of the propionate of heme ring IV and the carboxylate of Asp60, the lysine at position 45 in the human protein is too far away from Asp60 for any interaction (Scouloudi & Baker, 1978). It was anticipated that replacing Lys45 with Arg in the human protein would result in slower ligand recombination kinetics, as the added hydrogen bond would be expected to raise the barrier for "opening" the channel. To the contrary, we observed a faster k_{on} for K45R than for wild type (Lambright et al., 1989). This result has been reproduced in sperm whale and porcine myoglobins (Carver et al., 1991). Thus, although the residue at position 45 plays some role in determining the ligand recombination kinetics, our initial findings suggested that a more complex mechanism is at work, calling for a more extensive analysis by site-directed mutagenesis.

In the present paper, we report the CO recombination kinetics in aqueous buffer of a series of single- and double-site mutations of human myoglobin involving residues His64, Val68, Lys45, and Asp60. Our results suggest that the coupling between surface and pocket residues is more subtle than predicted by the simplest model. Although it is likely that this pathway is not the only one for ligand entry and exit (Case & Karplus, 1979; Elber & Karplus, 1990; Huang & Boxer, 1994), we are able to dissect the roles of each residue in this one pathway.

MATERIALS AND METHODS

Sample Preparation. Expression and purification of recombinant human myoglobin mutants has been described

[†] This work was supported in part by a grant from the National Institutes of Health (GM27738). S.M.D. is an NSF Predoctoral Fellow.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1994.

[1] Abbreviations: Mb, myoglobin, MbCO, carbonmonoxymyoglobin; MbO_2, oxymyoglobin; WT, wild type; NMR, nuclear magnetic resonance; SVD, singular value decomposition; MEM, maximum entropy method.

0006-2960/94/0433-5518$04.50/0

elsewhere (Varadarajan et al., 1985; Balasubramanian et al., 1993a). Samples for photolysis were prepared immediately prior to use. Typically, 300 μL of 100 mM potassium phosphate buffer, pH 7.0, was placed in a 1 mm path length quartz cuvette, sealed with a rubber septum, and deoxygenated by bubbling with argon through a 30-gauge needle for 1 h. The buffer was then equilibrated with CO by bubbling through a 30-gauge needle for an additional hour. Sodium dithionite ($Na_2S_2O_4$) from a freshly prepared 1 M stock solution under an argon atmosphere was added to a final concentration of 10 mM. Finally, a few microliters of concentrated met aquo Mb in aqueous solution was added to give an OD_{423} of 0.5.

Measurement of CO Recombination by Flash Photolysis. The setup for collecting transient absorption spectra with nanosecond time resolution following flash photolysis has been described in detail elsewhere (Lambright et al., 1991; Lambright, 1992). Briefly, samples were photolyzed using 8-ns (fwhm) pulses from the 532-nm doubled output of a Q-switched Nd:YAG laser. The change in absorbance, $\Delta A(\lambda,t)$, was measured as a function of time at single wavelengths over the wavelength range 400–460 nm by using CW probe light from a xenon arc lamp passed through a monochromator. Signals were detected with a 200 MHz bandwidth photodiode detector and digitized with a Tektronix DSA602 dual time base digitizing oscilloscope.

Singular Value Decomposition. The task of analyzing the kinetics of absorption changes, $\Delta\mathbf{A}(\lambda,t)$, measured at multiple wavelengths can be greatly simplified by the singular value decomposition

$$\Delta\mathbf{A}(\lambda,t) = \mathbf{USV}^{\mathrm{T}} \tag{1}$$

where **U** is a column orthonormal matrix containing basis spectra which span the column space of $\Delta\mathbf{A}(\lambda,t)$, **V** is a row and column orthonormal matrix with columns which represent the normalized time dependence for each of the basis spectra in **U**, and **S** is a diagonal matrix containing a set of weights, called the singular values, which measure the overall contribution of each basis spectrum to $\Delta\mathbf{A}(\lambda,t)$ (Golub & Reinsch, 1970; Hofrichter et al., 1985; Lambright et al., 1991; Lambright et al., 1993). The first basis spectrum $\mathbf{U}_1$ and its corresponding time course $\mathbf{SV}_1$ together form the best single-component representation of $\Delta\mathbf{A}(\lambda,t)$ in the least squares sense. If the only spectral change is that of the monotonic decay of the Mb–MbCO spectrum due to CO rebinding to photolyzed Mb, then in the limit of infinite signal-to-noise, $\mathbf{U}_1$ would be a normalized Mb–MbCO spectrum, $\mathbf{SV}_1$ would be the time course for ligand binding, and the remaining singular values would all be zero.

Maximum Entropy Rate Distributions. The data for CO rebinding to Mb exhibit a multiexponential decay, $N(t)$, which is presumed to be related to some underlying distribution of rates $p(k)$. The rate and time domains are related by a La Place transform

$$N(t) = \int_0^{k(\max)} p(k)\exp(-kt)\,dk \tag{2}$$

The distribution function $p(k)$ may be continuous, discrete, or both. Given $p(k)$, it is generally straightforward to calculate $N(t)$ numerically. In the present case, however, one begins with a discretely sampled $N(t)$, which is corrupted with measurement noise and seeks to recover an unknown $p(k)$. It is well-known that the numerical inversion of eq 2 is an ill-posed mathematical problem, particularly in the presence of noise. Consequently, there will be many feasible solutions which both satisfy the data and are physically reasonable.

Table 1: Bimolecular Rate Constants for CO Association to Human Myoglobin Mutants at 295 K

myoglobin	$k_{on} \times 10^{-6}$ (M^{-1} s^{-1})	myoglobin	$k_{on} \times 10^{-6}$ (M^{-1} s^{-1})
WT	0.998 ± 0.011	V68A	6.00 ± 0.18
K45R	1.88 ± 0.021	V68E	0.0272 ± 0.002
K45A	1.04 ± 0.021	V68I	0.0547 ± 0.002
K45Q	0.917 ± 0.017	V68L	1.12 ± 0.017
D60E	1.07 ± 0.06	V68N	0.150 ± 0.0056
D60A	0.878 ± 0.028	V68A/K45R	8.90 ± 0.33
H64A	3.75 ± 0.14	D60E/K45R	1.79 ± 0.06
H64L	33.1 ± 0.15	H64A/K45R	4.90 ± 0.50
H64Q	1.71 ± 0.01	sperm whale	0.733 ± 0.04
H64V	7.25 ± 0.35		

It is possible, however, to select a solution from the feasible set which has no correlations in $p(k)$ that are not present in the data. This is done by maximizing the Lagrangian function $Q = S - \lambda\chi^2$, where $\chi^2 = \sum(d_i - m_i)^2/\sigma_i^2$ is the usual misfit statistic and S is the Shannon–Jaynes entropy (Jaynes, 1983)

$$S = \sum p_k \log(p_k/m_k) \tag{3}$$

where m_k is a prior model for the distribution and

$$p_k = \frac{\int_k^{k+\Delta k} p(k)\,dk}{\int_0^{k(\max)} p(k)\,dk} \tag{4}$$

In principle, the value of λ is chosen such that $\chi^2 \simeq N$, where N is the number of data pairs. This solution is defined as the maximum entropy (MEM) solution.

In the present application, the MEM solution for $p(k)$ was found using the general algorithm developed by Skilling and Bryan (1984). The transform variable k was represented by 256 points which were equally spaced on a log scale. An "ignorant" prior distribution was established by setting m_k = constant, which corresponds to $m_k = k$ on a linear scale (Jaynes, 1983; Livesly et al., 1986). The progress of the algorithm was monitored by calculating

$$\mathrm{TEST} = \frac{1}{2}\left|\frac{\nabla S}{|\nabla S|} - \frac{\nabla\chi^2}{|\nabla\chi^2|}\right|^2 \tag{5}$$

Convergence was determined by two criteria, $\chi^2/N = 1.3$ and TEST < 0.01. This choice for the final χ^2/N allows for some uncertainty in the estimate of σ_i and avoids the potential problem of overfitting the data.

RESULTS

Bimolecular Rebinding Process. Values of the second-order rate constants (k_{on}), derived from the observed pseudo-first-order rate constants, are listed in Table 1. The errors are the standard deviations of at least four independent measurements. The bimolecular rate is insensitive to mutations at positions 45 and 60, with the exception of the K45R mutation (Lambright et al., 1989). However, the substitutions at positions 64 and 68 result in a wide range of rates for bimolecular recombination. Substitution of His64 by glutamine, similar to histidine in volume and polarity, leads to a less than 2-fold increase in k_{on}. Much larger effects are seen when His64 is replaced with aliphatic residues Ala, Val, and Leu, with roughly 4-, 7-, and 30-fold increases, respectively. Interestingly, the bimolecular rates for this series of mutants increases with the volume of the aliphatic side chain (Ala < Val < Leu).

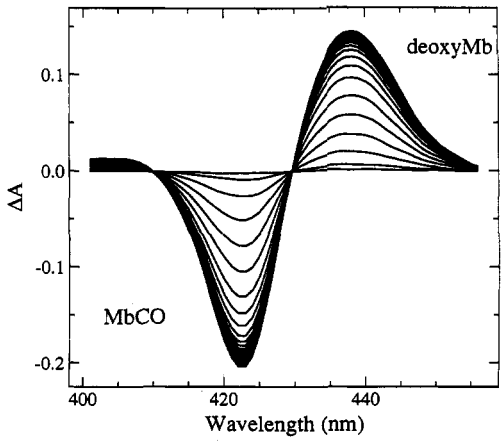

FIGURE 1: Transient ΔA spectra following photolysis of WT MbCO at 295 K in aqueous solution, pH 7.0.

Mutations at position 68 also have a significant effect on bimolecular kinetics. The k_{on} measured for V68L is only slightly greater than that for WT, while the k_{on} for V68I is approximately one-twentieth that of the WT. This remarkable result indicates that the more branched structure of the isoleucine side chain relative to its structural isomer, leucine, has a 20-fold effect on the bimolecular kinetics. Whereas replacement of Val68 with bulky Ile and Asn substantially decreases the bimolecular rate, replacement with Ala results in a 6-fold increase in the rate.

Analysis of Transient Absorption Spectra. In glycerol–water solutions, a relaxation process of deoxymyoglobin occurs on the same time scale as geminate recombination (Lambright et al., 1991; Ansari et al., 1992; Lambright et al., 1993). To correctly determine geminate-rebinding kinetics, this relaxation process must be separated from the ligand-rebinding component of the decays. In order to see if this process contributes to decays measured on aqueous samples, transient absorption spectra have been measured in aqueous buffer. The spectra at all times resemble the equilibrium Mb–MbCO spectrum, indicating that the changes in amplitude primarily result from the rebinding of CO to photolyzed Mb.

The transient spectra (Figure 1) were processed by singular value decomposition (SVD). The first two basis spectra (columns $\mathbf{U}_1$ and $\mathbf{U}_2$) from the SVD of $\Delta A(\lambda,t)$ for WT are shown in Figure 2. As anticipated, $\mathbf{U}_1$ is nearly identical to the equilibrium Mb–MbCO spectrum. $\mathbf{U}_2$, on the other hand, has a large feature near the peak of the deoxy band but relatively little amplitude near the peak of the MbCO band. The time-dependent amplitudes of these basis spectra, $\mathbf{S}_{11}\mathbf{V}_1$ and $\mathbf{S}_{22}\mathbf{V}_2$, are shown in Figure 3. $\mathbf{S}_{11}\mathbf{V}_1$ makes the dominant contribution; $\mathbf{S}_{22}\mathbf{V}_2$ is considerably smaller (1% or less of $\mathbf{S}_{11}\mathbf{V}_1$). Because $\mathbf{S}_{22}\mathbf{V}_2$ is so small, the possibility that it is an artifact cannot be eliminated. Thus, to a very good approximation, $\mathbf{S}_{11}\mathbf{V}_1$ is an experimental measure of the survival probability of ligand recombination. This is in contrast to observations at lower temperature and in 75% glycerol–water, where conformational relaxation makes a significant contribution to the observed decays (Lambright et al., 1991, 1993).

Ligand-rebinding survival probability curves ($\mathbf{S}_{11}\mathbf{V}_1$) for WT at four temperatures between 280 and 310 K are shown in Figure 3. Two well-resolved processes are evident in the data, a nanosecond process due to geminate rebinding and a millisecond process due to bimolecular rebinding. Under the conditions of these experiments ([CO] $\gg$ [Mb]), the bimolecular process is pseudo-first-order. In order to quantitate the kinetics of geminate rebinding and their dependence on

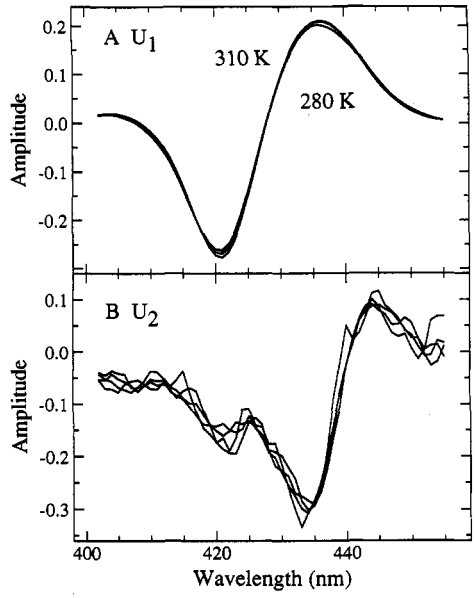

FIGURE 2: Basis spectra (columns of U) from the SVD of $\Delta A(\lambda,t)$ for WT at 280, 290, 300, and 310 K: (A) the basis spectrum with the largest singular value ($\mathbf{U}_1$); (B) the basis spectrum with the second largest singular value ($\mathbf{U}_2$).

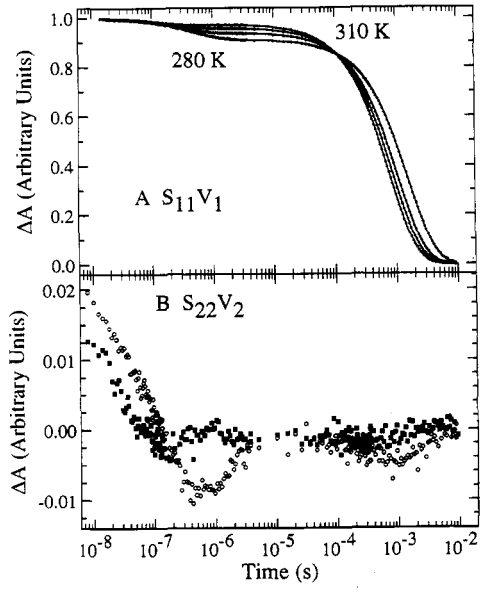

FIGURE 3: Time dependence (columns of S·V) of the basis spectra from the SVD of $\Delta A(\lambda,t)$ for WT: (A) $\mathbf{S}_{11}\mathbf{V}_1$ at 280, 290, 300, and 310 K; (B) $\mathbf{S}_{22}\mathbf{V}_2$ at 280 and 310 K.

temperature, the data at each temperature have been fit to a sum of exponentials

$$N(t) = \sum_{i=1}^{n} a_i \exp(-k_i t) \qquad (6)$$

Fitted models for $n = 2$ are overlaid with the data for the wild-type protein (WT) in Figure 3. Within the signal-to-noise limits of these measurements, the kinetics of the geminate process in WT are fit well by a single-exponential model in this temperature range.

Geminate Recombination Process. Parameters for the geminate kinetics were calculated from fits of survival probability curves (Table 2). For most mutants, the geminate process at 295 K is well-described by a single exponential. For the mutants H64L and V68A, the geminate process is best fit by two exponentials. In four of the mutants (H64V, V68N,

Table 2: Fitted Parameters for CO Geminate Rebinding in Aqueous Solution at 295 K

myoglobin	$k_g \times 10^{-6}$ [a] (s^{-1})	ϕ_g
sperm whale	5.53 ± 0.19	0.032 ± 0.003
WT	5.87 ± 0.24	0.060 ± 0.002
K45R	6.88 ± 0.11	0.123 ± 0.005
K45A	6.45 ± 0.10	0.043 ± 0.002
D60E	6.16 ± 0.07	0.068 ± 0.001
D60E/K45R	8.03 ± 0.63	0.112 ± 0.001
H64Q	7.46 ± 0.12	0.672 ± 0.001
V68L	44.1 ± 0.36	0.049 ± 0.003

	process I[b]		process II[b]	
	$k_{1g} \times 10^{-6}$ (s^{-1})	ϕ_{1g}	$k_{2g} \times 10^{-6}$ (s^{-1})	ϕ_{2g}
H64L	28.1 ± 0.17	0.31 ± 0.006	5.07 ± 0.30	0.068 ± 0.002
V68A	24.7 ± 0.19	0.20 ± 0.008	4.85 ± 0.30	0.14 ± 0.005
V68A/K45R	34.8 ± 0.2	0.23 ± 0.002	5.09 ± 0.06	0.17 ± 0.006

[a] Data for these mutants fit well with one intermediate (eq 7). [b] Data for these mutants require a four-state model (eq 9).

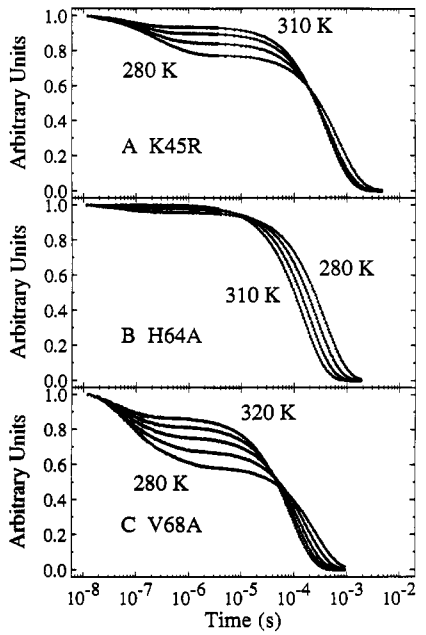

FIGURE 4: Kinetics for CO rebinding to the human Mb mutants (A) K45R, (B) H64A, and (C) V68A. The dots are the data and the solid lines are fits to a sum of two (H64A and K45R) or three exponentials (V68A).

V68I, and V68E), the geminate process is too small to be measured in the temperature range investigated; in the mutant H64A, the geminate process is too small to be accurately measured at 295 K but is larger at lower temperatures. Sample kinetic decay curves of mutants with large (V68A), small (H64A), and intermediate (K45R) geminate yields are presented in Figure 4.

Although the geminate process in H64L and V68A is well fit by a sum of two exponentials, it is possible that the kinetics arise from a distribution of rates rather than discrete exponential processes. In order to test this possibility, the maximum entropy method has been used to find a discretized distribution $p(k_i)$ which is one feasible solution to the inverse of eq 2. MEM distributions for WT, H64L, and V68A at 280 K are compared in Figure 5. The WT MEM distribution consists of two resolved peaks, one corresponding to the nanosecond geminate process and the other to the millisecond bimolecular process. The H64L MEM distribution, on the other hand, shows three peaks, one well-resolved peak for the bimolecular process and two partially overlapping peaks for

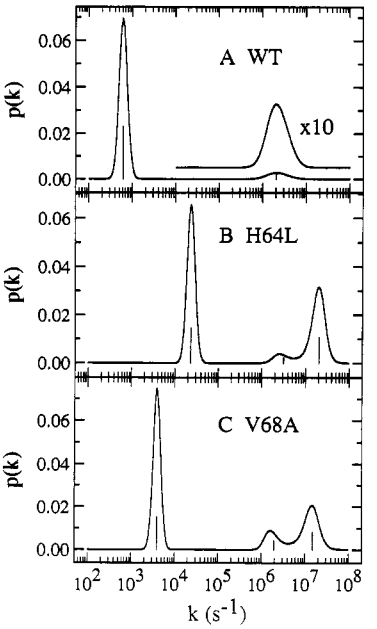

FIGURE 5: MEM rate distributions $p(k)$ for CO rebinding to human Mb mutants at 280 K: (A) WT; (B) H64L; (C) V68A.

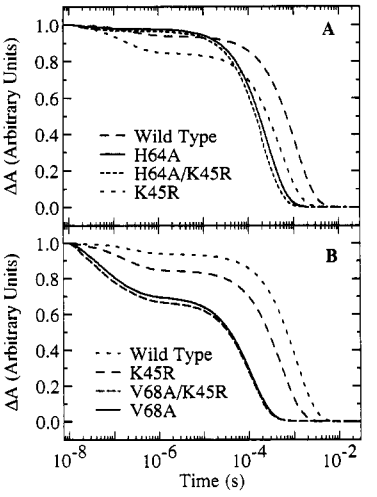

FIGURE 6: Survival probability curves for single and double mutants: (A) H64A, H64A/K45R, K45R, and WT at 295 K; (B) V68A, V68A/K45R, K45R, and WT at 295 K.

the geminate process. Given that the data for H64L and V68A are well-fit by a sum of three exponentials with rates corresponding to the peaks in the MEM distribution, a reasonable conclusion is that the nonexponential geminate kinetics in these mutants reflect two separate geminate processes rather than one distributed process.

Ligand Rebinding in Double-Site Mutants. The ligand-rebinding curves of the double mutants H64A/K45R and V68A/K45R are compared to single-site mutants in Figure 6. Both the geminate and bimolecular kinetics of H64A/K45R and V68A/K45R closely resemble those of H64A and V68A, respectively. Whereas the K45R mutation increases both k_{on} and the geminate yield about 2-fold over that of wild type, the H64A/K45R and V68A/K45R mutants give parameters close to those of H64A and V68A (Tables 1 and 2). The effects of the mutations are clearly not additive.

Kinetic Models. A variety of reaction schemes has been proposed for the binding of diatomic ligands to Mb. These models help to systematize trends in the data; however, the molecular details of the models are difficult to assess. The simplest of these is the three-state sequential scheme

$$\underset{A}{MbCO} \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} \underset{B}{Mb{:}CO} \underset{k_{SB}}{\overset{k_{BS}}{\rightleftharpoons}} \underset{S}{Mb + CO} \quad (7)$$

where MbCO is the liganded state in which CO is bound to the heme iron, Mb:CO is a geminate state in which the iron–CO bond has been photolyzed but CO is still trapped within the protein, and Mb + CO is a state in which the photolyzed CO molecules have either rebound or escaped into bulk solvent. This scheme has been used previously to interpret the binding of CO to sperm whale Mb in aqueous solution (Henry et al., 1983). In the limit where [CO] ≫ [Mb], ligand migration into the protein becomes a pseudo-first-order process with an apparent rate constant $k_{SB} = k_{SB}[CO]$. In all of the proteins considered in this work, k_{AB} is much smaller than the other rate constants, and its contribution will therefore be neglected. Assuming initial populations of $p_B(0) = 1$ and $p_A(0) = p_S(0) = 0$ and using a steady-state approximation for the decay of Mb + CO, the observables k_g (geminate rate constant), ϕ_g (geminate yield), and k_{on} (pseudo-first-order association rate constant) are related to the rate constants for the three-state sequential scheme by

$$k_g = k_{BA} + k_{BS} \quad (8a)$$

$$\phi_g = k_{BA}/(k_{BA} + k_{BS}) \quad (8b)$$

$$k_{on} = k_{SB}k_{BA}/(k_{BA} + k_{BS}) \quad (8c)$$

This simple scheme adequately describes the binding of CO to sperm whale Mb and mutants of human Mb which have only one resolved geminate process. For the mutants which have two resolved processes, a modification is necessary. One possibility is the four-state sequential scheme (Gibson et al., 1986; Ansari et al., 1986)

$$\underset{1}{MbCO} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \underset{2}{Mb{:}CO} \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} \underset{3}{Mb{::}CO} \underset{k_{43}}{\overset{k_{34}}{\rightleftharpoons}} \underset{4}{Mb + CO} \quad (9)$$

where states 1 and 4 correspond to states A and S, respectively, in the three-state scheme. State 2 (Mb:CO) might correspond to a state in which the ligand is weakly bound in the heme pocket, and state 3 (Mb::CO) might correspond to a state in which the ligand has migrated into the protein matrix. Assuming k_{21}, k_{23}, k_{32}, $k_{34} \gg k_{43}[CO] \gg k_{12}$ with initial populations $p_2(0) = 1$ and $p_1(0) = p_3(0) = p_4(0) = 0$ and making a steady-state approximation for ligand binding from state 4, the survival probability $N(t)$ is given by a sum of three exponentials:

$$N(t) = \phi_{g1} \exp[-k_{g1}t] + \phi_{g2} \exp[-k_{g2}t] + (1 - \phi_{g1} - \phi_{g2}) \exp[-k_{on}t] \quad (10)$$

In the limit where k_{21}, $k_{23} \gg k_{32}$, $k_{34} \gg k_{12}$, $k_{43}[CO]$, the relations above reduce to the expressions obtained by making steady-state approximation for rebinding from states 3 and 4:

$$\phi_{g1} \simeq \phi_2 \quad (11a)$$

$$\phi_{g2} \simeq \phi_3(1 - \phi_2) \quad (11b)$$

$$k_{g1} \simeq k_{21} + k_{23} \quad (11c)$$

$$k_{g2} \simeq k_{32}\phi_2 + k_{34} \quad (11d)$$

$$k_{on} \simeq k_{43}\phi_3[CO] \quad (11e)$$

Table 3: Calculated Rate Constants[a] for CO Rebinding to Mb Mutants in Aqueous Solution at 295 K

myoglobin	$k_{BA} \times 10^{-6}$ (s^{-1})	$k_{BS} \times 10^{-6}$ (s^{-1})	$k_{SB} \times 10^{-6}$ (M^{-1} s^{-1})
sperm whale	0.18	5.4	21
WT	0.35	5.5	18
K45R	0.85	6.0	16
K45A	0.28	6.2	27
D60E	0.42	5.7	17
H64A	1.2	49	160
H64Q	0.50	6.2	24
V68L	2.2	42	22

	$k_{21} \times 10^{-6}$ (s^{-1})	$k_{23} \times 10^{-6}$ (s^{-1})	$k_{32} \times 10^{-6}$ (s^{-1})	$k_{31} \times 10^{-6}$ (s^{-1})	$k_{34} \times 10^{-6}$ (s^{-1})	$k_{43} \times 10^{-6}$ (M^{-1} s^{-1})
H64L	9.0	22	1.6	0.36	7.2	650
V68A	5.4	16	2.3	0.58	4.4	61

[a] Rate constants for mutants, except H64L and V68A, were calculated from the kinetic parameters in Tables 1 and 2 using eq 8. With the exception of k_{31}, rate constants for H64L and V68A were obtained directly from the fits of eq 10 to the data. The apparent rate of rebinding from state 3, k_{31}, was calculated from the other rate constants using eq 13.

$$\phi_2 \simeq k_{21}/(k_{21} + k_{23}) \quad (11f)$$

$$\phi_3 \simeq k_{32}\phi_2/(k_{32}\phi_2 + k_{34}) \quad (11g)$$

Furthermore, if $\phi_{g1} < \sim 0.01 < \phi_{g2}$, then rebinding from state 2 will not be observed. In this case, there are still five elementary rate constants but only three observables:

$$\phi_{g2} \simeq \phi_2 \quad (12a)$$

$$k_{g2} \simeq k_{31} + k_{34} \quad (12b)$$

$$k_{on} \simeq k_{43}\phi_3[CO] \quad (12c)$$

These relations allow three rate constants k_{34}, k_{43}, and the apparent rate of rebinding from state 3

$$k_{31} \equiv k_{32}\phi_2 \quad (13)$$

to be determined from the observables. This limit corresponds to an alignment of the three- and four-state kinetic schemes such that state B in the three-state scheme corresponds to state 3 in the four-state scheme.

Fitting Data to Models. The information contained in the ligand-rebinding survival probability curves was parametrized by fitting it to eq 6. Model parameters for geminate rebinding are summarized in Table 2. For most mutants, only one geminate process could be resolved and the kinetics were analyzed in terms of eq 8. As discussed above, a second geminate process was observed for H64L and V68A. Because the two geminate processes partially overlap, the ligand-rebinding curves for these mutants were fit to eq 10, which yielded directly the elementary rate constants for the four-state sequential reaction scheme (except k_{12}, which cannot be determined from these experiments). An estimate of the apparent rebinding rate k_{31} was then calculated from eq 13 for comparison with WT. Rate constants at room temperature are listed in Table 3.

Over the narrow range of temperature in these experiments, the rate constants exhibit a dependence on temperature which obeys an Arrhenius law

$$k = A \exp(-E_a/RT) \quad (14)$$

Estimates of the prefactor A and activation energy E_a, obtained

Table 4: Activation Parameters for CO Rebinding to Human Mb Mutants in Aqueous Solution

myoglobin	rate constant[a]	A (s^{-1})	E_a (kJ mol^{-1})
WT	k_{BA}	1×10^7	9.7
	k_{BS}	2×10^{13}	37
	k_{SB}	8×10^{14}	43
K45R	k_{BA}	3×10^6	3.2
	k_{BS}	8×10^{12}	35
	k_{SB}	2×10^{14}	40
H64Q	k_{BA}	8×10^7	12
	k_{BS}	2×10^{13}	37
	k_{SB}	6×10^{14}	41
H64L	k_{21}	7×10^7	5.2
	k_{23}	7×10^{12}	31
	k_{32}	1×10^{12}	33
	k_{34}	3×10^{14}	43
	k_{43}	9×10^{16}	46
V68A	k_{21}	2×10^7	3.1
	k_{23}	3×10^{11}	24
	k_{32}	3×10^{12}	35
	k_{34}	8×10^{13}	41
	k_{43}	9×10^{14}	41

[a] Rate constants for WT, K45R, and H64Q were calculated from the kinetic parameters in Table 2 using eq 8. Rate constants for H64L and V68A were obtained directly from the fits of eq 10 to the data for these mutants.

from a linear least squares fit to the data, are listed in Table 4. For WT, this analysis yields values of $A = 1.5 \times 10^{13}$ and $E_a = 37$ kJ/mol for the escape rate k_{34}, which are similar to the values reported for escape of O_2 from sperm whale Mb ($A = 4.3 \times 10^{12}$, $E_a = 32.6$ kJ/mol; Chatfield et al., 1990). This observation supports the view that the rate-determining barrier for escape is the same for both diatomic ligands.

DISCUSSION

SVD Analysis. In the analysis of the transient spectra in aqueous solution using SVD, the time-dependent amplitudes $S_{22}V_2$ are very small (1% or less of $S_{11}V_1$). This stands in contrast to our earlier studies in 75% glycerol–water, where a significant second component was observed and was attributed to conformational relaxation prior to geminate recombination (Lambright et al., 1991, 1993). This conformational relaxation has been shown to be dependent on both temperature and viscosity, and in aqueous solution at room temperature, conformational relaxation is rapid compared with the time scale of geminate rebinding (Ansari et al., 1992; Lambright et al., 1993). Thus, in the following discussion of the kinetics in aqueous solution, we will not consider conformational relaxation.

Ligand Binding Pathways. In general, the diffusion of diatomic ligands through the protein matrix is a complex dynamical process involving a large number of different pathways between the distal heme pocket and solvent. For instance, internal cavities resulting from packing defects might provide weak binding sites for small ligands. Some internal cavities exist on a time-averaged basis while others, created by fluctuations of the protein structure, exist only transiently. Crystallographic studies of sperm whale met Mb equilibrated with 7 atm of Xe gas have found four different Xe sites in the protein matrix with occupancies ranging from 0.45 to 1.0 (Tilton et al., 1984). Molecular dynamics simulations which include protein atoms but not solvent molecules have identified at least five possible pathways for ligand escape from the heme pocket (Case & Karplus, 1979), and random mutagenesis has revealed several more (Huang & Boxer, 1994). One of these pathways is the His64/Val68/Arg45 pathway investigated here. While the study of individual site-directed mutants cannot give a complete and satisfactory resolution to the larger issues of the number and relative importance of one pathway over others, a systematic study of residues involved in one specific path can at least eluciate the structural and dynamic operation of that path.

Effect of Residue 45 on Rebinding Kinetics. A simple model for ligand entry into the heme pocket is the distal pocket gating hypothesis, which asserts that ligand access to the heme pocket is gated by fluctuations between open and closed conformations of the distal heme pocket (Nobbs, 1966; Case & Karplus, 1979; Ringe et al., 1984; Kuriyan et al., 1986; Kottalam & Case, 1988). Arg45 is implicated in stabilizing a "closed" pocket conformation through hydrogen-bonding interactions with adjacent polar groups (Takano, 1977a,b). Most myoglobins, including the human protein, have a Lys at position 45. Our initial hypothesis was that substitution of Arg for Lys45 in human myoglobin should stabilize the "closed" pocket and therefore have slower rebinding kinetics. In fact, the K45R mutant does have a more stable closed pocket conformation than the wild-type protein, as evidenced by a 10-fold decrease in the base-catalyzed exchange rate of the His64 N^{ϵ} proton (Lambright et al., 1989). However, this substitution results in an increase in the apparent rebinding rate k_{BA} but no change within experimental error in either the rate of CO entry or escape. Substitution of Arg45 for Ala produces a small decrease in the apparent rebinding rate, little change in the escape rate, and a 1.5-fold increase in the rate for ligand entry into the protein. The results for O_2 rebinding to Arg45 mutants of SW and porcine Mbs are similar (Carver et al., 1991).

Interestingly, the surface mutation K45R affects the rate of rebinding rather than the rate of escape. In addition, we have reported earlier that the rate k_{off} for the mutant K45R is slower by a factor of 1.45; thus the apparent equilibrium binding constant for CO is increased relative to the wild-type protein (Lambright et al., 1989). In a separate study, we have shown that this mutation also affects the picosecond kinetics of NO geminate rebinding (Petrich et al., 1994), which have been thought to reflect primarily the iron–ligand bond formation step (Jongeward et al., 1988; Petrich et al., 1988). One possible mechanism by which the K45R substitution might perturb the environment of the heme pocket involves the interaction between Arg45 and His64. In such a model, the K45R substitution perturbs the ligand-binding step via a specific interaction with distal pocket residues, such as the imidazole side chain of His64. Thus, in the mutant K45R, this added interaction between Arg45 and His64 may serve to draw the imidazole ring away from the Fe atom, thus facilitating the rotation of Val68 and lowering the inner barrier (Case & Karplus, 1979) as observed. This interaction may involve two or more bridging water molecules which have been observed in the crystal structure of K45R (Hubbard et al., 1990) and in a neutron diffraction structure of sperm whale myoglobin (Cheng & Schoenburn, 1990).

Rebinding Kinetics in Double-Site Mutants. To test this hypothesis, the double mutants V68A/K45R and H64A/K45R were prepared. The absence of the imidazole ring of His64 should diminish the effect of the residue at position 45 on the inner barrier. Indeed, as seen in Figure 6A, the geminate processes for H64A and H64A/K45R are nearly identical. In the case of V68A, the removal of the forked valine residue at position 68 significantly reduces the inner barrier as compared to WT (Table 4). Further, in spite of the presence of a histidine residue at position 64, the added interaction of Arg45 in the V68A/K45R double mutant no longer affects the barrier

height to rebinding of CO (Figure 6B). Thus, the magnitude of effect of the surface K45R substitution depends on the nature of interior residues at positions 64 and 68. The similarity in rebinding kinetics between the single mutants H64A and V68A and their double-mutant counterparts H6A/K45R and V68A/K45R provides strong evidence that the effect of the K45R surface mutation on the inner barrier is due to its interactions with His64 and Val68. Mutation of either of these residues to Ala attenuates this chain of interaction and reduces the impact of this mutation on the observed rebinding kinetics. The ability of a surface residue to affect the inner barrier to CO rebinding can be explained by this simple model. The observation of coupled motion between surface and inner residues which has been systematically explored here may prove to be a quite general phenomenon.

Effect of Residues 64 and 68 on Rebinding Kinetics. The mutants at positions 64 and 68 exhibit variations over a large range in both geminate and bimolecular kinetics and the amplitude of the geminate process. One of the more striking differences is the geminate process: wild type and the majority of mutants have a single-exponential geminate process, some mutants show no measurable geminate phase under these conditions (H64A, V68N, V68I, and V68E), and two (H64L and V68A) have a geminate process best fit by two exponentials. Nonexponential kinetics might reflect the presence of discrete kinetic intermediates or, alternatively, a distribution of rebinding rates. The MEM distributions for H64L and V68A indicate that the signal-to-noise of the data is adequate to resolve the geminate rebinding into two partially overlapping components, both of which are well-separated in time from the bimolecular process. Furthermore, the analysis of transient spectra here and in previous work suggests that conformational relaxation is complete before the ligand rebinding process begins in aqueous solution at room temperature (Ansari et al., 1992; Lambright et al., 1993). When the relaxation rate is fast compared to the rate of rebinding, the distributed function can be collapsed to a single exponential (Balasubramanian et al., 1993a). Consequently, the nonexponential kinetics are interpreted here in terms of discrete kinetic intermediates rather than a distribution of rebinding rates due to slow time-scale averaging over conformational substrates, as has been proposed for CO rebinding to Mb in 75% glycerol–water solutions below 270 K (Austin et al., 1975; Steinbach et al., 1991; Balasubramanian et al., 1993a).

The observation of two discrete geminate processes requires a reaction scheme with a minimum of two intermediates. There are many possible schemes, including sequential, parallel, and more complicated mixed schemes. At the present time there is no *a priori* basis for selecting a single best scheme. As a working hypothesis, the data were analyzed in terms of the four-state sequential reaction scheme (eq 8), which is the simplest reversible scheme consistent with the data. There are two geminate intermediates in this scheme; state 2 (Mb:CO) represents an intermediate in which photolyzed ligand molecules are trapped within the heme pocket, while in state 3 (Mb::CO) ligands have diffused into the protein matrix (Ansari et al., 1986). Ligands in state 3 may either escape from the protein or rebind via state 2. For most mutants, only a single geminate process is observed, which could be due to rebinding from either state 2 or state 3.

Replacing His64 with Leu results in a dramatic increase in the association rate (Table 1) which is primarily due to a 30-fold increase in the rate of ligand entry into the protein (Table 3). The H64A change also results in a large increase in ligand entry rate k_{SB} (Table 3); however, this is matched by a similar decrease in the ligand escape rate as well. Thus, the faster bimolecular association rate in this mutant is mostly due to a 3-fold increase in the geminate rebinding rate k_{BA}. The Val substitution at His64 results in a 7-fold increase in the overall association rate. The small geminate yield in this mutant indicates that the ratio of the escape rate to the rebinding rate has increased at least 6-fold relative to WT. Although it is not possible to determine the rate for entry into the pocket, it is possible to establish a lower limit. Assuming $\phi_g \leq 0.01$ yields $k_{SB} \geq 7.3 \times 10^8$, which is an increase of at least 30-fold compared to WT. Finally, substitution of Gln for His64 leads to a 1.7-fold increase in the overall association rate which appears to be the result of both an increase in the rate of rebinding, k_{BA}, and an increase of similar magnitude in the rate of ligand entry, k_{SB}. Carver et al. (1990) report similar results for O_2 rebinding to sperm whale Mb mutants. These observations are consistent with the expectation that His64 presents a barrier to ligand access to the distal heme pocket; however, it does not indicate the extent to which ligands escape via this pathway in the native protein.

Val68 is a buried hydrophobic residue which is thought to influence ligand binding through steric interactions in the transition state of the iron–ligand bond formation step as well as by contributing to the barrier for ligand entry and escape from the distal heme pocket (Case & Karplus, 1979; Kottalam & Case, 1988). In fact, a 6-fold increase in the association rate constant for the V68A substitution is observed, arising from a 3-fold increase in the rate of entry into the protein (k_{43}) combined with a 2-fold increase in the apparent rate of rebinding from state 3. There is also a greatly increased geminate yield for V68A, which is consistent with a large rate of rebinding from state 2. In contrast, the V68L substitution results in a 10-fold increase in both the rate of rebinding from state B and the rate of escape but no significant change in the rate of entry into the protein. Replacing Val68 with Ile, Asn, or Glu results in overall association rates which decrease by 6- to 30-fold and geminate yields which decrease by at least 6-fold. The decreased geminate yields appear to account for a large fraction, if not all, of the decrease in the association rates. However, it is not clear whether the decrease in geminate yield is due to a decrease in the rate of rebinding or an increase in the rate of escape or both.

Finally, it is observed from Table 4 that the barrier height for the final rebinding step is much decreased in H64L and V68A as compared to WT. While this is also true of the mutant K45R, the Arrhenius prefactor in the case of this mutant is also reduced. In H64L and V68A, however, this prefactor is larger than in WT, indicating a smaller volume change on activation and possibly also a smaller volume change on going from the geminate to the bound state (Szabo, 1978).

Conclusions. Substitutions at His64 and Val68 result in 1–2 order of magnitude changes in the rates for CO rebinding, escape, and entry into the protein. Substitutions at Lys45 and Asp60, on the other hand, have a smaller effect on the kinetics; the most significant change is a factor of 2 increase in the apparent rate of rebinding when Lys45 is replaced with Arg. As reported earlier (Lambright et al., 1989), this effect is curious because a surface mutation results in a change in the inner barrier to CO rebinding. From the analysis of double mutants, it was demonstrated that the effect of the surface mutation K45R on the inner barrier to rebinding is likely due to a chain of interactions between this residue and those at positions 64 and 68.

The pathway involving residues His64/Val68/Lys45 is likely only one of several pathways for ligand rebinding (Case & Karplus, 1979; Elber & Karplus, 1990; Huang & Boxer, 1994). While it is difficult to make general conclusions about the relative importance of a given pathway by even the most detailed site-specific mutagenesis of the residues involved, this technique makes it possible to elucidate potential mechanisms by which the structures of the residues changed affect the kinetics. Furthermore, the results presented here demonstrate that multiple changes of adjacent residues can be a useful tool for dissecting the roles of residues which might not be expected to have an effect on ligand-rebinding kinetics.

REFERENCES

Ansari, A., DiIorio, E. E., Dlott, D. D., Frauenfelder, H., Iben, I. E. T., Langer, P., Roder, H., Sauke, T. B., & Shyamsunder, E. (1986) *Biochemistry 25*, 3139–3146.

Ansari, A., Jones, C., Henry, E. R., Hofrichter, J., & Eaton, W. A. (1992) *Science 256*, 1796–1798.

Balasubramanian, S., Lambright, D., Marden, M. C., & Boxer, S. G. (1993a) *Biochemistry 32*, 2202–2212.

Balasubramanian, S., Lambright, D., & Boxer, S. G. (1993b) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4718–4722.

Bolognesi, M., Cannillo, E., Ascenzi, P., Giacometti, G. M., Merli, A., & Brunori, M. (1982) *J. Mol. Biol. 158*, 305–315.

Carver, T. E., Rohlfs, R. J., Olson, J. S., Gibson, Q. H., Blackmore, R. S., Springer, B. A., & Sligar, S. G. (1990) *J. Biol. Chem. 265*, 20007–20020.

Carver, T. E., Olson, J. S., Smerdon, S. J., Krzywda, S., Wilkinson, A. J., Gibson, Q. H., Blackmore, R. S., Dez Ropp, J., & Sligar, S. G. (1991) *Biochemistry 30*, 4697–4705.

Case, D. A., & Karplus, M. (1979) *J. Mol. Biol. 132*, 344–368.

Cheng, X., & Schoenborn, B. (1991) *J. Mol. Biol. 220*, 381–399.

Elber, R., & Karplus, M. (1990) *J. Am. Chem. Soc. 112*, 9161–9175.

Gibson, Q. H., Olson, J. S., McKinnie, R. E., & Rohlfs, R. J. (1986) *J. Biol. Chem. 261*, 10228–10239.

Golub, G. H., & Reinsch, C. (1970) *Numer. Math. 14*, 403–420.

Hofrichter, J., Henry, E. R., Sommer, J. H., Deutsch, R., Ikeda-Saito, M., Yonetani, T., & Eaton, W. A. (1985) *Biochemistry 24*, 2667–2679.

Huang, X., & Boxer, S. G. (1994) *Nat. Struct. Biol. 1*, 226–229.

Hubbard, S. W., Hendrickson, W. A., Lambright, D. G., & Boxer, S. G. (1990) *J. Mol. Biol. 213*, 215–218.

Jaynes, E. T. (1986) in *Maximum Entropy and Bayesian Methods in Applied Physics* (Justice, J. H., Ed.) pp 26–58, Cambridge University Press, Cambridge, UK.

Johnson, K. A., Olson, J. S., & Phillips, G. N., Jr. (1989) *J. Mol. Biol. 207*, 459–463.

Jongeward, K. A., Magde, D., Taube, D. J., Marsters, J. C., Traylor, T. G., & Sharma, V. S. (1988) *J. Am. Chem. Soc. 110*, 380–387.

Kottalam, J., & Case, D. A. (1988) *J. Am. Chem. Soc. 110*, 7690–7697.

Kuriyan, J., Wilz, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133–154.

Lambright, D. G. (1992) Ph.D. Thesis, Stanford University, Palo Alto, CA.

Lambright, D. G., Balasubramanian, S., & Boxer, S. G. (1989) *J. Mol. Biol. 207*, 289–299.

Lambright, D. G., Balasubramanian, S., & Boxer, S. G. (1991) *Chem. Phys. 158*, 249–260.

Lambright, D. G., Balasubramanian, S., & Boxer, S. G. (1993) *Biochemistry 32*, 10116–10124.

Lecomte, J. T. J., & La Mar, G. N. (1985) *Biochemistry 24*, 7388–7395.

Livesey, A. K., Licinio, P., & Delaye, M. (1986) *J. Chem. Phys. 84*, 5102–5107.

Nobbs, C. L. (1966) in *Hemes and Hemoproteins* (Chance, B., Eastbrook, R. W., & Yonetani, T., Eds.) pp 143–147, Academic Press, New York.

Petrich, J. W., Lambry, J. C., Kuczera, K., Karplus, M., Poyart, C., & Martin, J.-L. (1991) *Biochemistry 30*, 3975–3987.

Petrich, J. W., Lambry, J. C., Balasubramanian, S., Lambright, D. G., Boxer, S. G., & Martin, J. L. (1994) *J. Mol. Biol. 238*, 437–444.

Phillips, S. V. E. (1980) *J. Mol. Biol. 142*, 531–554.

Ringe, D., Petsko, G. A., Kerr, D. E., & Ortiz de Montellano, P. R. (1984) *Biochemistry 23*, 2–4.

Scouloudi, H., & Baker, E. N. (1978) *J. Mol. Biol. 126*, 637–660.

Springer, B. A., & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961–8965.

Szabo, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2108–2111.

Takano, T. (1977a) *J. Mol. Biol. 110*, 537–568.

Takano, T. (1977b) *J. Mol. Biol. 110*, 569–584.

Tilton, R. F., Jr., Kuntz, I. D., Jr., & Petsko, G. A. (1984) *Biochemistry 23*, 2849–2857.

Varadarajan, R., Szabo, A., & Boxer, S. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5681–5684.