Can a Two-State MWC Allosteric Model Explain Hemoglobin Kinetics?

Eric R. Henry,* Colleen M. Jones,‡ James Hofrichter, and William A. Eaton*

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received August 2, 1996; Revised Manuscript Received February 10, 1997[⊗]

ABSTRACT: We have analyzed the nanosecond−millisecond kinetics of ligand binding and conformational changes in hemoglobin. The kinetics were determined from measurements of precise time-resolved optical spectra following nanosecond photodissociation of the heme−carbon monoxide complex. To fit the data, it was necessary to extend the two-state allosteric model of Monod, Wyman, and Changeux (MWC) to include geminate ligand rebinding and nonexponential tertiary relaxation within the R quaternary structure. Considerable simplification of the model is obtained by using a linear free energy relation for the rates of quaternary transitions, and by incorporating concepts from recent studies on the physics of geminate rebinding and conformational changes in myoglobin. The model, described by 85 coupled differential equations, quantitatively explains a demanding set of complex kinetic data. Moreover, with the same set of kinetic parameters it simultaneously fits the equilibrium data on ligand binding and the distribution of ligation states. The present results, together with those from single-crystal oxygen binding studies, indicate that the two-state MWC allosteric model has survived its most critical tests.

Understanding cooperative oxygen binding to hemoglobin has been one of the central problems in biochemistry. Its role as a paradigm for understanding cooperative behavior in other biological systems has, however, diminished somewhat in recent years because of confusion in the literature concerning the basic explanation of the origin of cooperativity. This is unfortunate, since an enormous amount of structural, spectroscopic, thermodynamic, and kinetic information can be synthesized by a relatively simple model, in spite of numerous claims to the contrary. This MWC−PSK model consists of the two-state allosteric model of Monod, Wyman, and Changeux (MWC) (Monod et al., 1965; Shulman et al., 1975; Edelstein, 1975), the stereochemical mechanism of Perutz (P) (Perutz, 1970; Perutz et al., 1987), and the statistical thermodynamic formulation of the Perutz mechanism by Szabo and Karplus (SK) (Szabo & Karplus, 1972; Lee et al., 1988).

The principal idea of MWC was that hemoglobin exists in only two affinity states, corresponding to a low-affinity structure called T and a high-affinity structure called R. Binding to either structure is statistical, i.e., for a given structure of the tetramer the affinity of each subunit is independent of the number of ligands already bound. In this model cooperative oxygen binding arises from a shift in the population from the low-affinity (T) to the high-affinity (R) structure as the oxygen pressure increases. This contrasted with the earlier proposal by Pauling (1935), later elaborated by others (Koshland et al., 1966), that cooperative interactions arose in a sequential fashion, the binding of a ligand changing the conformation of the subunit and thereby directly influencing the affinity of the neighboring subunits. Perutz's X-ray crystallographic studies showed that fully liganded and fully unliganded hemoglobin have different arrangements of the four subunits, identifying the affinity states of the MWC model with the two quaternary structures, R and T.

Perutz further made the novel proposal that salt bridges between subunits stabilized the low-affinity (T) quaternary structure, lowered the affinity of T relative to R, and were responsible for the Bohr effect (the decrease in oxygen affinity with decreasing pH, promoting the release of oxygen in the more acid tissues) (Perutz, 1970). Szabo and Karplus (1972) showed that Perutz's mechanism could explain the existing equilibrium data. They also showed that the mechanism was perfectly consistent with the hypothesis of noncooperative binding within each quaternary structure, the cornerstone of the MWC two-state model. This fact has rarely been mentioned in the literature (Hess & Szabo, 1979; Rivetti et al., 1993), and consequently Perutz's mechanism has often erroneously been used as support for a sequential mechanism. The Perutz mechanism also explained why the T-state affinity is changed by allosteric effectors, which was not envisaged in the original MWC model (Szabo & Karplus, 1972, 1975).

Because hemoglobin in the R quaternary structure has an affinity very close to that of the free chains, the issue has always centered on whether the increase in affinity with increasing oxygen pressure arises entirely from a switch from T to R, as proposed by MWC, or whether there is significant cooperativity in binding to the T quaternary structure. Single-crystal binding experiments have directly addressed this question. They showed that in crystals of hemoglobin, where the quaternary structure is known to remain T by X-ray diffraction (Liddington et al., 1988), binding is non-cooperative with a value for the Hill parameter $n = 1.0$ (Rivetti et al., 1993; Mozzarelli et al., 1991). Non-cooperative binding has also been observed for deoxyhemoglobin encapsulated in a silica gel, presumed to prevent the T to R transition (Shibayama & Saigo, 1995). Dissociation of oxyhemoglobin encapsulated in the gel is noncooperative as well, with a 500-fold increase in affinity as expected for R-state hemoglobin.

* Correspondence should be addressed to either author.
‡ Current address: Department of Chemistry, University of South Alabama, Mobile, AL 36688.
⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1997.

S0006-2960(96)01917-4

A caveat in both the crystal and gel experiments is that noncooperative binding of a tetramer, defined by a Hill *n* of 1.0, can be the result of exact compensation of cooperative interactions by inequivalence in the intrinsic affinity of α and *β* subunits. Rivetti et al. (1993) considered this problem by taking advantage of the small difference in projection of the α and *β* hemes onto the crystal axes of their measurements. From the two apparent binding curves for light linearly polarized parallel to the two axes, Rivetti et al. calculated that the Hill *n* of 1.0 does indeed result from compensation of inequivalence in oxygen binding to the α and *β* subunits (with intrinsic affinity differences of less than a factor of 5) and a small amount of cooperativity within the T quaternary structure. This cooperativity within T can be most simply described by the dimeric cooperon model of Brunori and co-workers (Brunori et al., 1986; Gill et al., 1986) with the parameter $\delta \approx 2$, which means that there is a 2-fold increase in affinity of an α (or *β*) subunit following ligand binding to a *β* (or α) neighbor. This is small compared to a factor of 100−400 increase in affinity upon changing from T to R. Confirmation of this analysis has been obtained recently, using the nickel−iron hybrid hemoglobin, in which only the *β* hemes bind oxygen (Bettati et al., 1996). This study shows that Rivetti et al., if anything, overestimated the α/*β* inequivalence and therefore did not underestimate the cooperativity in binding to the T quaternary structure. The crystal studies also showed that oxygen binding is pH-independent in the crystal, where the salt bridges are known to remain unbroken by X-ray crystallography, in keeping with Perutz's mechanism for the Bohr effect (Mozzarelli et al., 1991; Rivetti et al., 1993; Bettati et al., 1996). The net result of the crystal work, where there is no ambiguity about quaternary structure, is that two critical elements of the equilibrium properties of the MWC−PSK model are experimentally confirmed: oxygen binds noncooperatively in the absence of a T → R quaternary conformational change and there is no Bohr effect when the salt bridges remain intact.

Crystal studies might be criticized because lattice forces may not just prevent the transition from T to R, but may also distort the properties of the T state. In the case of the nickel−iron hybrids, Bettati et al. (1996) showed that this is not the case, finding that the oxygen affinities in crystal and solution are virtually identical. Rivetti et al. (1993) also pointed out that as the Hill *n* in solution studies decreased toward 1.0 by stabilizing the T state with strong allosteric effectors, the affinity approached that of the crystal. Nevertheless, one would like to design experiments on solutions of hemoglobin that provide comparably incisive tests of models. The problem with equilibrium studies in solution is that there are no unambiguous markers of quaternary structure and the system is so cooperative that intermediates are always present at low populations. To circumvent the latter problem metal substitution or oxidation has been used to create stable ligation states as models for intermediate states of oxygenation. Ackers and co-workers have observed a dependence of the tetramer−dimer dissociation constants on the positions of the bound ligands for analogues of doubly-liganded tetramers (Ackers et al., 1992). These observations also require such a position dependence for the R ⇄ T equilibrium constant for doubly-liganded tetramers. That is, the tetramer−dimer and R ⇄ T equilibrium constants are different for a tetramer with both ligands on the same dimer compared to a tetramer with one ligand on each dimer [α1(x)*β*1(x)α2*β*2 and α1(x)*β*1α2*β*2(x)]. This has intuitive appeal, because hemoglobin does not have pseudotetrahedral symmetry (with only one kind of interface between subunits) but has pseudo 222 symmetry (with four kinds) with different α1*β*1 and α1*β*2 interfaces. In terms of binding, the position-dependence of the tetramer−dimer dissociation constants requires cooperative binding to the T state, as pointed out by Brunori and co-workers (Brunori et al., 1986; Gill et al., 1986). The results suggest that binding to an α (or *β*) subunit raises the affinity of the *β* (or α) subunit on the same dimer but not the subunit on the opposite dimer. Brunori and co-workers showed that the pattern of tetramer−dimer dissociation constants could be explained with such a model. On structural grounds one might have expected that the cooperative dimer is not α1*β*1 (and α2*β*2), as suggested from the work of Ackers and co-workers. Instead one would expect that it is α1*β*2 (and α2*β*1), because it is the interface between these subunits, rather than the α1*β*1 (and α2*β*2) interface, that changes with the quaternary change and is directly connected to the heme via the F helix.

Given that there is some cooperativity in binding to the T quaternary structure in solution, the important question is how large is this deviation from perfect two-state behavior. The investigation of tetramer−dimer dissociation equilibria in cyanomet hybrid hemoglobins by Ackers and co-workers (Ackers et al., 1992, Holt & Ackers, 1995) leads to the conclusion that the T quaternary structure can bind ligands (in this case the hypothetical binding of the CN radical to ferrous heme) with high cooperativity.[1] Edelstein (1996) has pointed out, however, that equilibrium oxygen binding curves cannot possibly be explained using cooperativity parameters similar to those obtained from studies of cyanomet derivatives ($\delta = 170$), showing that they are not reliable analogues of intermediate oxygenation states. [*Note Added in Proof.* Shibayama et al. have just reported that they have redetermined *δ* and find that it is more than 10-fold smaller, i.e., $\delta < 17$; Shibayama, N., Morimoto, H., & Saiger, S. (1997) *Biochemistry 36*, 4375−4381.] In the case of the carbon monoxide complex, the data of Perrella and co-workers on the population of ligation states as a function of carbon monoxide saturation determined by low-temperature electrophoresis show a possible deviation from perfect two-state behavior (Perrella et al., 1990, 1995). Their data, which we shall consider in detail in this paper, require a cooperativity parameter *δ* in the model of Brunori and co-workers of 2.9,[2] which is the same within experimental error as the value of ∼2 derived from the crystal oxygen binding curves (Rivetti et al., 1983; Bettati et al., 1996; Mozzarelli et al., 1997).

Kinetic methods provide another means of generating large populations of intermediate states in solution. In general, kinetic experiments provide more incisive tests of mechanism, and have played a central role in the history of

[1] Gill et al. (1986) were able to fit the data on cyanometHb hybrids with the dimeric cooperon model but did not know it at the time because of an incorrect value for the cooperative free energy of species **22**, which was later found to be the same as that of species **23** and **24** (Ackers et al., 1992).

[2] Using the expressions for the cooperative free energy for each ligation state given in Table 1 by Gill et al. (1986), we varied the three parameters of the "dimeric cooperon" model, *L*, *c*, and *δ*, to obtain the best least-squares fit to the cooperative free energies reported in Table 2 of Perrella et al. (1990) for carbonmonoxy Hb. The resulting values are $L = 1.4 \times 10^5$, $c = 0.0021$, $\delta = 2.9$.

hemoglobin. However, the kinetics of hemoglobin are complicated, and in recent years there has been a tendency to deal with only small pieces of the problem. Gibson made the seminal observation that in mixing experiments the bimolecular rate of carbon monoxide binding to deoxyhemoglobin is ~30 times slower than the rate of bimolecular *rebinding* after flash photolysis (Gibson, 1959). This result immediately showed that the rate of ligand binding was not determined by the number of ligands already bound, as required by Pauling's sequential model. Gibson also found that the deoxyhemoglobin produced by photolysis had an altered structure, as judged by the heme spectrum, and he called this the spectrum of a "fast-reacting" hemoglobin, Hb^*. Hopfield et al. (1971) pointed out another important consequence of Gibson's results for models of hemoglobin, namely, that Hb^* should be identified as deoxyhemoglobin in the R quaternary structure. These investigators also showed that a variety of kinetic experiments on both ligand association and dissociation could be qualitatively explained with a simple two-state allosteric model in which the rates for the quaternary transitions were considered to be instantaneous relative to ligand binding. Sawicki and Gibson (1976), photodissociating carbonmonoxyhemoglobin with 0.6 μs laser pulses, found that deoxyheme spectral changes and ligand rebinding occurred simultaneously in the microsecond−millisecond time regime. Although Sawicki and Gibson were able to explain the ligand rebinding kinetics and spectral changes (presumed to arise from quaternary conformational changes) at alkaline pH with an MWC model, they reported that they could not satisfactorily fit the data on spectral changes at neutral pH.

With the advent of time-resolved optical spectroscopy of hemoglobin using picosecond (Greene et al., 1978) and nanosecond lasers (Hofrichter et al., 1983; Henry et al., 1983a,b), the focus shifted toward the submicrosecond regime, where it was possible to observe geminate recombination of the photodissociated ligands: carbon monoxide (Duddell et al., 1979), oxygen (Chernoff et al., 1980), and nitric oxide (Cornelius et al., 1983). Time-resolved resonance Raman spectroscopy first exposed tertiary relaxation kinetics of the protein (Lyons & Friedman, 1982), which were subsequently studied in detail by optical absorption methods (Hofrichter et al., 1983). In subsequent years the connection of the submicrosecond studies to the problem of the cooperative mechanism was considered [see reviews by Murray et al. (1988a) and Goldbeck et al. (1996)] but never fully explored.

The objective of the present investigation has been to bridge the gap between the classic studies of Gibson and co-workers and the submicrosecond investigations. We ask the following question: can the kinetics of hemoglobin be explained by an obvious extension of the two-state allosteric model of MWC to include geminate rebinding and tertiary relaxation? The "kinetics" in this case are represented by the data of Jones et al. (1992) in which the carbon monoxide complex of human hemoglobin was photodissociated and the ligand rebinding and conformational kinetics monitored by measuring very precise time-resolved spectra from ~10 ns to ~70 ms. Jones et al. used the Gibson strategy of changing the distribution of ligation states by varying the degree of photolysis (Sawicki & Gibson, 1976). They showed that both the kinetics of geminate rebinding and the kinetics of conformational relaxation were independent of the degree of photolysis for times less than ~1 μs. In terms of the two-state model, this result indicates that the quaternary structure remains R until ~1 μs. After ~1 μs the kinetics are highly dependent on the degree of photolysis as observed previously (Sawicki & Gibson, 1976; Hofrichter et al., 1983). In the present study we have taken on the more ambitious project of fitting the kinetic data on both conformational changes and ligand rebinding with a two-state allosteric model over the complete time range, from the beginning of geminate rebinding at nanoseconds to the completion of bimolecular rebinding at tens of milliseconds. To ensure that our results are consistent with equilibrium data, we also simultaneously fit the carbon monoxide binding curve and the distribution of ligation states as a function of carbon monoxide saturation determined by low-temperature electrophoresis (Perrella et al., 1990, 1995).

There are two important simplifying features to the model. One is the use of a linear free energy relation between the rates and equilibria for the quaternary conformational changes (Eaton et al., 1991). This markedly reduces the number of parameters required to describe these rates. The second is the use of a stretched exponential relaxation function for the tertiary conformational change. The extended, continuous nature of this relaxation was apparent from the early optical studies, which showed that the time course consisted of at least two overlapping exponential functions (Hofrichter et al., 1983, 1985). The more precise data of Jones et al. (1992) showed that the description of the tertiary relaxation required three exponential functions, with relaxation times of 20, 90, and 820 ns at room temperature. Frauenfelder and co-workers, by analogy to relaxation in glasses, had suggested that conformational relaxation in myoglobin might be described by a stretched exponential function, i.e., $\exp[-(kt)^{\beta}]$, $0 < \beta < 1$ (Frauenfelder et al., 1990, 1991). The initial observations on this relaxation in myoglobin following photodissociation indicated nonexponential behavior (Ansari et al., 1992), but it was not until the work of Anfinrud and co-workers that there was compelling evidence for continuous, extended behavior of the conformational relaxation in myoglobin (Lim et al., 1993; Jackson et al., 1994). There still remained the problem of the physical basis of the nonexponential time course. This question was recently considered by Hagen and Eaton (1996). They showed that two ingredients were necessary to explain the extended time course: (i) the interconversion of conformational substates having different optical frequencies occurring on the same time scale as the overall change from one average conformation to another and (ii) a simplification of the rate equations with a kinetic rule for the interconversion among substates. Their kinetic rule assumes that all transition states have the same energy. As a consequence, the activation energy becomes progressively larger and the apparent rate for substate interconversion slows dramatically as the system moves from the substates of the unrelaxed conformation to those of the relaxed conformation at lower energy. The solution of the kinetic equations describing this model closely approximates a stretched exponential function. The present model is too complex to consider individual conformational substates or substate distributions. However, we can incorporate the substates indirectly by writing the differential equation for the tertiary relaxation in a form that produces a stretched exponential function upon integration, and by considering geminate rebinding to take place with a

time-dependent rate "constant".

Another aspect of work on myoglobin has influenced our thinking about hemoglobin kinetics. Henry et al. (1983b) and Hofrichter et al. (1985) had suggested that the nonexponential geminate rebinding observed at room temperature in hemoglobin could arise either from multiple geminate states, as in the original Frauenfelder model for myoglobin (Austin et al., 1975; Beece et al., 1980), or from a slowing of geminate rebinding due to tertiary conformational relaxation. These models could not be distinguished by conventional kinetic experiments. The former model was less attractive because it implied that a spectral change very similar to that occurring in the R → T conformational change did not influence the rebinding rate. Steinbach et al. (1991) successfully employed the relaxation model, first suggested for myoglobin by Agmon and Hopfield (1983), to ligand-rebinding data over a wide temperature range. Ansari et al. (1994) supported this analysis by studying the kinetics of the conformational change as well as the ligand rebinding as a function of temperature and viscosity, and showed that the data could be simultaneously fit with a model in which the spectral change was produced by a relaxation that slowed geminate rebinding. A striking confirmation of the influence of the tertiary relaxation on geminate rebinding has been demonstrated recently by Hagen et al. (1995, 1996). They placed myoglobin in a glass (viscosity > 10^{15} cP) at room temperature. All of the deoxyheme spectral changes could be explained by kinetic hole burning, indicating that interconversion among substates and tertiary relaxation of the protein following photodissociation is prevented, as predicted from the high viscosity dependence found by Ansari et al. (1992, 1994). In the glass the geminate rebinding of carbon monoxide becomes distributed because the very slow dynamics effectively trap the substates, preventing the averaging that normally produces near-exponential geminate rebinding at room temperature. The substate-averaged rate in the glass is, however, ~2500 times faster than the rate observed for the relaxed protein. These experiments provide the most convincing evidence that conformational relaxation slows the geminate rebinding rate in myoglobin, strongly suggesting that the tertiary relaxation in hemoglobin does so as well.

THE KINETIC MODEL

Two-State Allosteric Kinetic Model. The historical starting point for constructing a kinetic version of the equilibrium two-state allosteric model for hemoglobin is shown in Figure 1. In this model there are two quaternary structures, R and T, and five ligation states for each structure to produce a total of ten states $\{R_i, T_i, i = 0-4\}$. Interconversion between states is accomplished by bimolecular ligand binding (e.g., $R_i \rightarrow R_{i+1}$), unimolecular ligand dissociation (e.g., $R_i \rightarrow R_{i-1}$), and quaternary conformational changes $R_i \rightleftharpoons T_i$. There is a single ligand dissociation rate and a single bimolecular ligand binding rate for each quaternary structure (apart from statistical factors) and ten quaternary rates. Thermodynamic linkage produces the following relation for the equilibrium constants:

$$L_i = Lc^i \qquad (1)$$

where the quaternary equilibrium constant with no ligands bound $L \equiv [T_0]/[R_0]$, and the allosteric parameter for the ratios of the affinities $c \equiv ({}^{T}k_{on}/{}^{T}k_{off})/({}^{R}k_{on}/{}^{R}k_{off})$.[2]

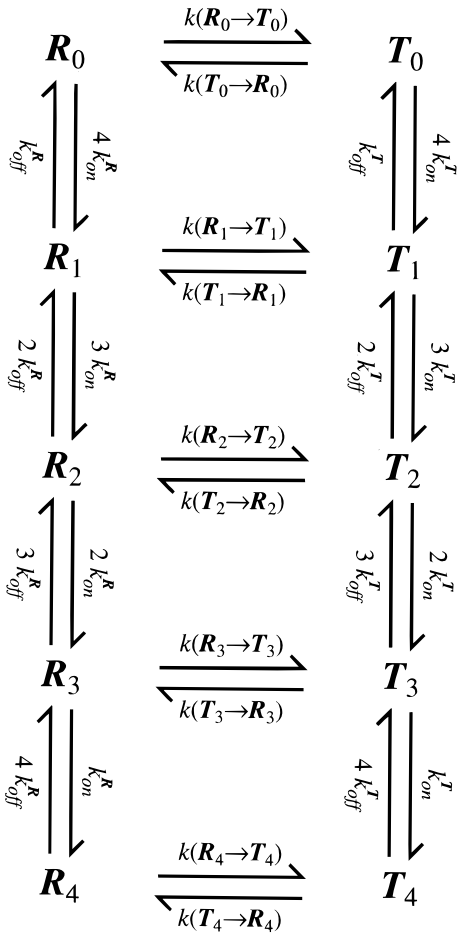

FIGURE 1: Species and rate constants of MWC two-state kinetic model. Possible states of the system are characterized by quaternary structure (R or T) and the number of subunits having ligands bound at the heme (subscript). The rates k_{on} and k_{off} are overall rates of ligation and deligation, respectively, of individual subunits.

Two new kinds of states must be added to the scheme of Figure 1 to account for tertiary relaxation and geminate rebinding (Figure 2). The simplest extension allows protein subunits to exist in two possible tertiary structures, which interconvert on the submicrosecond time scale to account for the spectral changes, and an additional ligand state to account for geminate recombination in which the ligand is not bound to the heme but is still within the protein. We designate the more rapidly geminate-rebinding species produced by ligand dissociation as r*, and the more slowly rebinding but more stable tertiary state as r. Therefore, each tertiary state exists in three possible ligation states:ligand bound (r*x or rx), ligand inside the protein but not bound to the heme (r*·x or r·x), and ligand outside the protein (r* or r). It is assumed that all liganded subunits prior to photolysis are in the r*x state, so that photodissociation initially populates exclusively r*·x.

The T state also exhibits geminate rebinding and tertiary relaxation following CO dissociation (Murray et al., 1988b). However, only the bimolecular ligand rebinding to the T state is observed in the present experiments, so all subunits in T-state tetramers are assumed to be in a single tertiary state, designated t. It is assumed that an R tetramer undergoing a quaternary conformational change to T will have all r and r* subunits undergo a simultaneous structural change to the

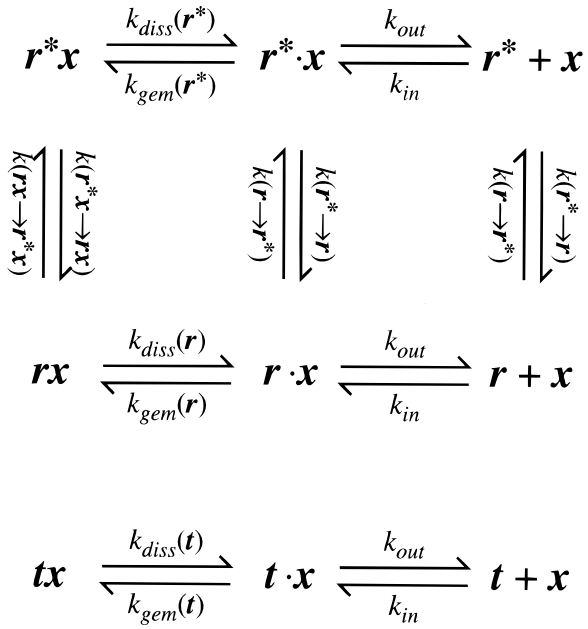

FIGURE 2: Minimal set of tertiary/ligation states of individual subunits required to describe tertiary conformational changes and geminate rebinding. Elementary transitions permitted between states are also shown. The nomenclature for the states and transition rates is described in the text.

T state (without ligand motion). Therefore, three corresponding ligation states (tx, t·x, and t) are also required for the subunits in the T quaternary structure, even though the geminate rebinding step (t·x → tx) is not directly observable in the current data. Since the presence of the unbound ligand in the protein should not influence the spectrum of the deoxyheme, we consider only three deoxyheme optical spectra in the model, corresponding to the tertiary conformations r*, r, and t, with t having the spectrum of equilibrium deoxyhemoglobin.

We initially did not distinguish between α and β subunits, which is justified by the work of Hofrichter et al. (1985). They studied the nanosecond−millisecond kinetics of the α and β subunits separately in cobalt−iron hybrid hemoglobins, in which the iron in either the α or β subunits was substituted with the cobalt; cobalt porphyrins do not bind carbon monoxide. The relaxation times for geminate rebinding, tertiary relaxation, the fastest quaternary relaxation, bimolecular rebinding to the R quaternary structure and bimolecular binding to the T quaternary structure were very similar, and always differed by less than a factor of 2. We shall relax the constraint of equivalence of all four subunits in the last phase of the kinetic fitting, but, as we shall see, it has a small effect.

The scheme in Figure 2 contains 18 rate constants, of which 16 are independent because of the two thermodynamic cycles in describing the kinetics within the R quaternary structure. This number is readily reduced. First, Murray et al. (1988b) studied geminate rebinding in a cobalt−iron hybrid hemoglobin as a function of the fraction of T state tetramers. They found that the geminate yield for the T state was less than 1% compared to 40% for the R state. By also measuring the geminate relaxation time they concluded that the 60-fold lower bimolecular rate for carbon monoxide binding to the T quaternary structure could be almost entirely explained by the decreased rate of geminate rebinding to the T state (t·x → tx), and that the rates of entry of the ligand into the protein and exit from the protein are nearly identical in the two quaternary structures. We have therefore assumed identical rates for exit and entry of ligands for all three subunit conformations, eliminating four rate constants. Two more may be eliminated by making the assumption that the subunit conformation does not depend on the position of the ligand, so that the rates for r*·x → r·x and r* → r are identical, as are the reverse rates r·x → r*·x and r → r*. The ligand dissociation rates are so slow that ligand binding is effectively irreversible, so these rates play no role in fitting the kinetic data. They are, however, determined by the equilibrium data. Finally, we assume that rx → r*x is instantaneous and irreversible. For fitting kinetic data, then, there are only seven independent rate constants: k_{gem}(r*), k_{gem}(r), k_{gem}(t), k_{in}, k_{out}, k(r*→r), and k(r→r*) in addition to the quaternary rates.

Nonexponential Tertiary Conformational Relaxation. As pointed out in the Introduction, the model of Hagen and Eaton (1996) postulates that the nonexponential tertiary relaxation arises because the spectral property monitoring this conformational relaxation is sensitive to the distribution of the substates for each molecular species. In chemical solution kinetics on long time scales at room temperature, there is usually complete thermal averaging over the microstates of the system. Each molecular species can be assigned a single spectral property and its transition to another molecular species can be assigned a single rate constant. However, on short time scales, thermal averaging is incomplete, and when species interconvert on time scales that are comparable to the transitions among the microstates, which are here called conformational substates, the substates must be considered explicitly, as was done by Hagen and Eaton (1996).

Our mechanism is too complex to treat the individual substates. We therefore seek a compromise by altering the differential equations of a conventional chemical kinetics scheme to produce the effects of interconversion of conformational substates occurring simultaneously with the change in the average conformation, labelled r* and r. We accomplish this by introducing time dependence into the rate coefficients describing the interconversion between two average conformations, r* and r. Specifically, we require time-dependent rate coefficients $k_1(t)$ and $k_2(t)$ such that the system of differential equations

$$\mathrm{dr^*/d}t = -k_1(t)\mathrm{r^*} + k_2(t)\mathrm{r} \qquad (2)$$

$$\mathrm{dr/d}t = k_1(t)\mathrm{r^*} - k_2(t)\mathrm{r}$$

has a solution of the form

$$\mathrm{r^*}(t) = \mathrm{r^*}(\infty) + [\mathrm{r^*}(0) - \mathrm{r^*}(\infty)]\,f(t) \qquad (3)$$

$$\mathrm{r}(t) = \mathrm{r}(\infty) + [\mathrm{r}(0) - \mathrm{r}(\infty)]\,f(t)$$

for some specified relaxation function $f(t)$. It is easily shown that for a relaxation function of the general form $f(t) = \exp[-g(t)]$, the desired result is obtained using functions of the form:

$$k_1(t) = \frac{l}{1+l}\,\mathrm{d}g/\mathrm{d}t \qquad (4)$$

$$k_2(t) = \frac{1}{1+l}\,\mathrm{d}g/\mathrm{d}t$$

where l [$= r(\infty)/r^*(\infty)$] is the equilibrium constant between the two states.

For the relaxation function $g(t)$ we choose the stretched exponential form, $f(t) = \exp[-(kt)^\beta]$ with $0 < \beta < 1$. A small technical difficulty in using this specific form arises from the fact that the derivative of $g(t) = (kt)^\beta$ becomes infinite as $t \rightarrow 0$ for $\beta < 1$, which complicates integration of the system of differential equations in eq 2. However, this singularity is readily handled by using a modified $g(t)$ having the form

$$g(t) = \frac{(kt)^\beta}{1 + a(kt)^{\beta-1}} \tag{5}$$

for some small positive constant a. For typical values of k and β encountered in this work, $a = 0.001$ produced a $g(t)$ which was indistinguishable from that appropriate for the simple stretched exponential over the experimental time range $t > 10$ ns. In the full model, the nonexponential tertiary relaxation was simulated by replacing the rate constants $k(r^* \rightarrow r)$ and $k(r \rightarrow r^*)$ by the time-dependent rate expressions (eq 2) in all transitions involving a change in tertiary state, using the modified $g(t)$ of eq 5 with

$$k = k(r^* \rightarrow r) + k(r \rightarrow r^*) \tag{6}$$

The adjustable parameters in the modeling of the tertiary relaxation were then β, $k(r^* \rightarrow r)$, and the tertiary conformational equilibrium constant l.

Quaternary Conformational Relaxation. In analyzing the simpler allosteric mechanism of Figure 1, Sawicki and Gibson (1976) assumed that the $R_i \rightarrow T_i$ rates were slowed by a constant factor d for each ligand bound. This is identical to assuming a linear free energy relation between the quaternary rates and equilibria described by Eaton et al. (1991). That is, the variation in the activation free energy for the quaternary change from the addition of a ligand to the tetramer is proportional to the change in the equilibrium free-energy difference between the quaternary states, with a proportionality constant α, i.e., $\delta(\Delta G^\ddagger) = \alpha\delta(\Delta G)$, or

$$k(R_i \rightarrow T_i) = \gamma(Lc^i)^\alpha \tag{7}$$

$$k(T_i \rightarrow R_i) = \gamma(Lc^i)^{\alpha-1}$$

where γ is a scaling parameter that sets the absolute rates, and Sawicki and Gibson's $d = c^{-\alpha}$. Ten quaternary rates are now described by only four parameters: α and γ and the MWC allosteric parameters L and c.

To include two tertiary conformations for the unliganded subunits within the R quaternary structure, the linkage relation (eq 1) is expanded to

$$L_{ij} \equiv \frac{k_{ij}(R \rightarrow T)}{k_{ij}(T \rightarrow R)} = Lc^i l^j \tag{8}$$

where i is the number of bound ligands and j is the number of unliganded subunits in the r* tertiary conformation. The definitions of the allosteric constants, L and c, now include the two tertiary conformations in the R quaternary structure, i.e.,

$$L \equiv \frac{[T_0]}{[R_{00}]}, \qquad c \equiv \frac{K_{T_i}}{K_{R_{i0}}} \tag{9}$$

The possible (ij) combinations are (00), (01), (02), (03), (04), (10), (11), (12), (13), (20), (21), (22), (30), (31), and (40). The most general description of the quaternary conformational kinetics within this model would require specifying the 15 rates $k_{ij}(R \rightarrow T)$ for the allowed (ij) combinations, plus the parameters L, c, and l, from which the 15 corresponding rates $k_{ij}(T \rightarrow R)$ may be determined using eq 9. However, there is certainly insufficient information in the experimental data to determine such a large number of independent rates, so it is necessary to reduce this number by assuming some relationship among them.

We have extended the linear free energy idea of Eaton et al. (1991) to include the two tertiary conformations of the R quaternary structure with the relation

$$k_{ij}(R \rightarrow T) = \frac{k_{00}(R \rightarrow T)}{d^i f^j} \tag{10}$$

Assuming the same proportionality constant α between activation and equilibrium free energies, the factor $f = l^{-\alpha}$. The complete set of R→T transition rates may then be defined by a single reference rate with all others obtained by scaling this rate up or down by powers of the constants f and d. Because independent information is available on quaternary transition rates in triply-ligated hemoglobin from the work of Ferrone and co-workers (1985) we chose the rate $k_{30}(R \rightarrow T)$ as the reference rate. The complete set of 30 quaternary rates is then determined by only five parameters: $k_{30}(R \rightarrow T)$, the conformational equilibrium constants L and l, c, and the linear free energy parameter α.

Geminate Ligand Rebinding. The nonexponential tertiary relaxation requires that we modify the conventional approach to geminate ligand rebinding kinetics. According to the mechanism of Hagen and Eaton (1996), the system is not at thermal equilibrium with respect to the conformational substates of either r* or r during the conformational relaxation. Consequently, at any given time it cannot be described by two species, r* and r, each with a single geminate rebinding rate. Instead, at any given time there is a *non-equilibrium* distribution among both the substates of r* and r, the extended time course arising mainly from the redistribution among conformational substates of the relaxed conformation (r) to establish the final Boltzmann-weighted occupancies (Hagen & Eaton, 1996). The individual substates have different geminate rebinding rates (Austin et al., 1975; Steinbach et al., 1991; Hagen et al., 1995, 1996), which can be accounted for by considering a property which correlates with the geminate rebinding rate. Steinbach et al. (1991) showed that for myoglobin the logarithm of the geminate rebinding rate is nearly linearly proportional to the frequency of the near-infrared charge transfer band. Hagen et al. (1996) found a similar relation for the frequency shift of the intense porphyrin $\pi \rightarrow \pi^*$ (Soret) band for myoglobin. The spectral change in hemoglobin has components of both a shift in frequency and a change in spectral shape, but can be simply described by the fractional amplitude of a spectrum that must be added to the photoproduct spectrum. It is this

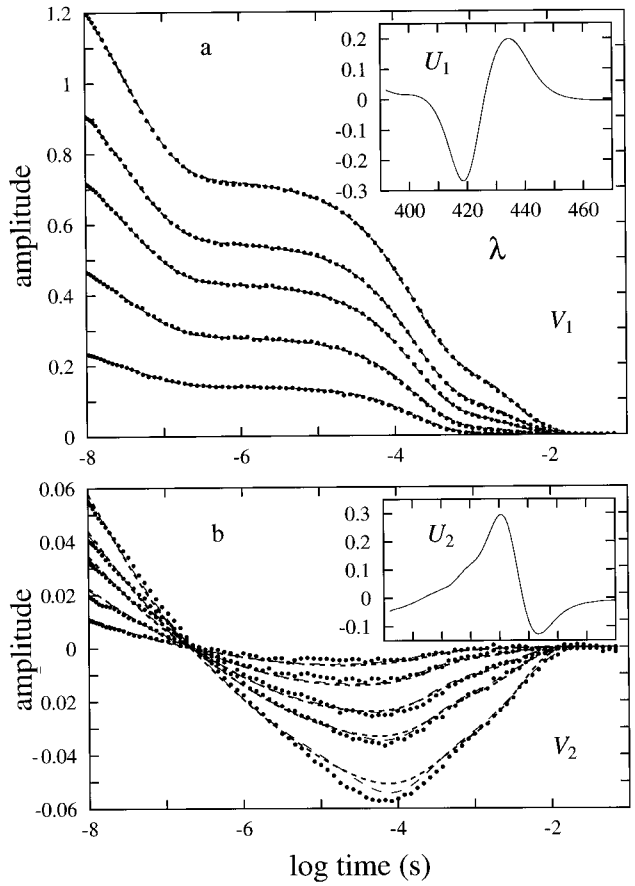

FIGURE 3: Kinetic data and fitted curves. Amplitudes for the experimental (dots) and simulated (dashed curves) data sets were derived by least-squares fitting to the set of basis spectra (columns of **U**) from the SVD of the experimental data set of Jones et al. (1992). (The corresponding most significant basis spectra are shown in insets to each panel.) The dotted curves therefore coincide exactly with the singular-value weighted amplitude vectors from this SVD. The heavy dashed curves show the results from fitting with the 85-state model, and the light dashed curves show the results after the introduction of $\alpha-\beta$ inequivalence. (a) Amplitude profiles corresponding to the most significant basis spectrum (first column of **U**). (b) Amplitude profiles corresponding to the second column of **U**.

amplitude which changes in a highly nonexponential fashion (see Figure 3).

Rather than using the populations of r* and r with their geminate rates, k_{gem}(r*) and k_{gem}(r), we employ the same geminate rebinding rate for both species with a time-dependent geminate rate coefficient k_{gem}(r*/r,t) that logarithmically interpolates between the r* and r rates according to the relation

$$k_{gem}(r^*/r,t) = k_{gem}(r^*)[k_{gem}(r^*)/k_{gem}(r)]^{x(t)-1} \quad (11)$$

where $x(t)$ measures the progress of the tertiary relaxation. $x(t)$ varies from 1 at $t = 0$ [k_{gem}(r*/r,t) = k_{gem}(r*)] to 0 at $t = \infty$ [k_{gem}(r*/r,t) = k_{gem}(r)]. The value of $x(t)$ is obtained from the tertiary relaxation function discussed above:

$$x(t) = \exp\left[-\frac{(kt)^{\beta}}{1 + a(kt)^{(\beta-1)}}\right] \quad (12)$$

In this way the logarithm of the geminate rebinding rate scales with the spectral change, and thereby accounts, at least in a qualitative way, for the differences in geminate rebinding rates among the substates.

METHODS

Simulating the Experimental Data Sets. The molecular species in our kinetic model is of course a tetramer. A tetramer is characterized by its quaternary structure and the tertiary and ligation states of each of the four constituent subunits taken from the set {r*x, rx, r*•x, r•x, r*, r} for the R quaternary structure and from the set {tx, t•x, t} for the T quaternary structure. A detailed description of the implementation of this model may be found in the Supporting Information. The model consists of 70 R + 15 T = 85 kinetic species (ignoring rx). With the connections between tertiary/ligation states of a single subunit shown in Figure 2, the allowed transitions between kinetic species in a single quaternary structure involve a change in state of only one subunit at a time. Transitions between quaternary structures may only connect tetrameric states having the same ligation state. With these restrictions, a total of 600 connections between kinetic species were enumerated for this model.

The data of Jones et al. (1992) consist of difference spectra spanning the wavelength range 390−470 nm measured at logarithmically-spaced time intervals from nanoseconds to about 70 milliseconds following photolysis by a 10-ns laser pulse; sets of such spectra were measured for probe−pulse polarizations both parallel and perpendicular to the polarization of the photolysis pulse, and for 16 different levels of photolysis. The conditions of these experiments were 120 μM hemoglobin (in heme), 0.1 M potassium phosphate, pH 7, 8 mM sodium dithionite, 20 °C. Preliminary analysis of the data set, described in detail by Jones et al. (1992), began with a smoothing operation, the creation of isotropically-averaged data from the measurements for the two probe polarizations to eliminate photoselection effects, and interpolating the spectra measured for different photolysis levels onto a common wavelength grid (consisting of n_λ wavelengths) and a common time grid (consisting of n_t times). At this stage the data set for each photolysis level was in the form of a $n_\lambda \times n_t$ matrix, and a complete data matrix for a set of n_p selected photolysis levels was formed by column-concatenating the data matrices for each photolysis level. The singular value decomposition of the full data matrix formed from all 16 photolysis levels was computed, and the 12 most significant components (singular values, basis spectra, and amplitude vectors) were kept for subsequent analysis. The data matrix re-created from these 12 components provided the experimental data set to be modeled in the present work.

The goal of modeling the $n_\lambda \times (n_p n_t)$ data matrix **D** is, for a specified set of n_s spectroscopically distinct species, to produce a set of time-dependent species populations (which comprise a matrix **P** having n_s rows and $n_p n_t$ columns) and a set of corresponding species spectra (which comprise a matrix **S** having n_λ rows and n_s columns), such that the product **SP** optimally approximates the matrix **D**. In the present case, there is little if any difference between the spectrum of liganded subunits and that of subunits which have been photolyzed and rebound ligand. We therefore assume that all liganded subunits have the same spectrum, so that there are three unliganded species, distinguished by three unliganded-minus-liganded difference spectra: unli-

ganded subunits in the r* tertiary structure, unliganded subunits in the r tertiary structure, and unliganded subunits in the t tertiary structure. For a specified set of elementary kinetic rate constants, the column subset of the matrix **P** corresponding to one level of photolysis is constructed by solving the model system of differential equations over the required set of times using the appropriate initial conditions and computing from the populations of all the tetrameric states at each time the total populations of unliganded subunits in each of the three tertiary structures. The complete **P** matrix is then built up from these populations computed for each of the chosen levels of photolysis.

Given the populations of the species in **P**, the matrix **S** of corresponding species spectra is computed by solving the linear least-squares problem $\mathbf{SP} \approx \mathbf{D}$ using efficient matrix algorithms. Methods for incorporating prior knowledge and assumptions about the shapes of one or more of the species spectra into this operation are described elsewhere (Henry, 1997). For our analysis, we have fixed the shape of the (unliganded-minus-liganded) difference spectrum of the t tertiary structure by assuming that the final decay to zero amplitude of the measured spectra is entirely due to ligand rebinding to subunits in this tertiary structure. All measured difference spectra for time delays greater than 5 ms were extracted from the data sets for all 16 levels of photolysis, and the most significant component of the SVD of the resulting matrix was used to define the shape of the required difference spectrum. However, the amplitude of the difference spectrum per deoxy subunit was not fixed but rather incorporated as an adjustable scale factor that was allowed to vary in subsequent fitting operations. The matrix product of the column vector of this scaled difference spectrum and the row vector of populations of unliganded t subunits from the solutions of the differential equations was subtracted from the data matrix. The (unliganded-minus-liganded) difference spectra for the r and r* tertiary structures were then determined entirely by the linear least-squares fitting involving the residual data matrix and the matrix of populations of unliganded r and r* subunits.

Fitting Procedure, Shapes of Spectra, and Parameter Constraints. The above procedure for simulating the matrix **D** of experimental data first specifies a set of rate constants, then solves the differential equations of the model for several sets of initial conditions to produce the species population matrix **P**, and finally creates a best estimate of the corresponding species spectrum matrix **S** using linear least-squares. The fitting operation adjusts the various elementary rate constants (as well as possibly additional parameters required to fix certain of the species spectra) to minimize the sum of squares of all the elements of the residual matrix $\mathbf{D} - \mathbf{SP}$. Rather than perform the fits using all 16 photolysis levels, we selected the data for five photolysis levels that effectively spanned the range of available values. A Marquardt−Levenberg nonlinear least-squares fitting algorithm, incorporating optional linear constraints on parameter values, was used for all fits.

An optimal simulation of the experimental data is required to produce a residual matrix $\mathbf{D} - \mathbf{SP}$ that is small. Moreover, it must also produce species spectra that are physically reasonable, as judged by the known shapes of such spectra. However, this property is generally not guaranteed for spectra generated by the mathematical linear least-squares operations outlined above. Our attempts to produce physically reasonable species spectra using parametrized analytical functional forms (e.g., Gaussian or Lorentzian shapes) (Jones et al., 1993; Henry, 1997) were unsuccessful, because typical deoxyheme spectra in the Soret region are highly asymmetric. It was, however, possible to obtain undistorted spectra of the unliganded r and r* subunits determined purely from the least-squares conditions by requiring that the time- and photolysis level-dependence of the total population of unrecombined deoxyhemes be well approximated by the first amplitude vector (column of **V**) from the singular value decomposition of the complete data set. This requirement was enforced by introducing a second contribution to the residuals to be minimized in the fitting procedure, given by the norm of the difference matrix between the first amplitude vector and the set of deoxyheme populations computed from the time- and photolysis level-dependent state populations provided by the solutions of the model differential equations. The "strength" of this requirement in the fitting was determined by adjusting the weighting factors applied to the residuals from the kinetic simulation of the experimental data and to the residuals from the deoxyheme population condition.

The adjustable parameters used in the modeling are summarized in Table 1. During the fits linear constraints were imposed on certain sets of parameters in order to ensure that they remained within acceptable bounds. The most important such constraints, reflecting either information from other experimental studies or fundamental assumptions in the kinetic model, are also listed in Table 1.

Simultaneous Fitting of Equilibrium Data. The kinetic model presented here is complete in the sense that rate parameters are included for all interspecies transitions, so that equilibria among all of the states may be established. It is (of course) essential that the kinetic parameters also be consistent with the known equilibrium ligand-binding properties of the system. The partition function which provides the equilibrium populations of all the states in the model may be constructed using the various kinetic parameters introduced above. The partition function is

$$\Theta([\mathrm{CO}]) = (1 + K_{\mathrm{in}}[\mathrm{CO}](1 + K_{\mathrm{R}}))^4 + L\left(\frac{l}{l+1}\right)^4 (1 + K_{\mathrm{in}}[\mathrm{CO}](1 + K_{\mathrm{T}}))^4 \quad (13)$$

$$K_{\mathrm{in}} \equiv (k_{\mathrm{in}}/k_{\mathrm{out}})/[\mathrm{CO}]_k$$

$$K_{\mathrm{R}} \equiv \left(\frac{l}{l+1}\right)\frac{k_{\mathrm{gem}}(\mathrm{r})}{k_{\mathrm{diss}}(\mathrm{r})}$$

$$K_{\mathrm{T}} \equiv \frac{k_{\mathrm{gem}}(\mathrm{t})}{k_{\mathrm{diss}}(\mathrm{t})}$$

where K_{R} and K_{T} are the equilibrium constants for ligand binding to the heme from the r and t geminate states, respectively, and K_{in} is the equilibrium constant for entry/exit from the protein, obtained by scaling the ratio of the ligand entry and exit rates from the kinetic modeling for the ligand concentration $[\mathrm{CO}]_k$ used in the kinetic experiments. (Because the free CO concentration is ~10-fold higher than the heme concentration, the parameter k_{in} determining the rate of ligand entry into the protein is here regarded as a pseudo-first-order rate constant which incorporates the

Table 1: Parameter Definitions

parameter	description	comments
L	quaternary equilibrium constant = $[T_0]/[R_{00}]$	*a*; constrain $0.1 < Lc^3 < 1$[b]
c	ratio of affinities = K_T/K_R	*a*
l	tertiary equilibrium constant = $k(r^*\rightarrow r)/k(r\rightarrow r^*)$	constrain $l > 1$[c]
$k(r^*\rightarrow r)$	tertiary transition rate	
$k_{gem}(r^*)$	geminate rebinding rate to r*	constrain $1.2\ k_{gem}(r) < k_{gem}(r^*) < 100k_{gem}(r)$[c]
$k_{gem}(r)$	geminate rebinding rate to r	
$k_{gem}(t)$	geminate rebinding rate to t	
$k_{diss}(r)$	heme−CO thermal bond breaking rate	generally fixed at "small" value (∼0.1) (see text)
k_{in}	rate of ligand entry into protein	generally pseudo-first-order rate assumed independent of protein state
k_{out}	rate of ligand exit from protein	assumed independent of protein state
$k_{30}(R\rightarrow T)$	quaternary rate, for triply-liganded tetramers containing no r* subunits	*a*; constrain $2.7 < \log(k) < 3.3$[b]
d	scale factor for R→T rates with decreasing ligation	*a*; constrain $k_{30}(R\rightarrow T)d^3 < 10^6\ s^{-1}$[d]
β	"stretching" parameter for tertiary relaxation	constrain $0 < \beta < 1$ (see text)
K_R	equilibrium constants for geminate binding step in R and T quaternary structures	required for fits to O_2 equilibrium binding curve
K_T		
f	scale factor for t-state difference spectrum (see text)	
$\lambda_{Cl}, \gamma_{Cl}$	scale factors relating allosteric L and c respectively, for Cl^- and PO_4^- solvent conditions (see text)	*a*; required only for fits to equilibrium data; constrain $3 < 1/\lambda_{Cl} < 30$ and $0.1 < \lambda_{Cl}\gamma_{Cl}^4 < 10$

[a] Parameters for which multiplicative constraints were required involving other parameters (e.g., $pq^i <$ constant) were replaced by their logarithms as adjustable parameters in the fits, thereby allowing such constraints to be written as linear constraints involving the logarithmic parameters [i.e., $\log p + i \log q < \log(\text{constant})$]. [b] Ferrone et al. (1985) have measured the R → T and T → R quaternary transition rates for triply-liganded Hb under comparable solution conditions to be 1000 and 4000 s^{-1}, respectively. Consistency with these measurements is enforced by the listed constraints. Relaxation of these constraints decreases the sum-of-squares by 20%, compared to a 30% improvement obtained by introducing $\alpha-\beta$ inequivalence which is shown in Figure 3. [c] The kinetic model is symmetric with respect to the interchange of r and r* tertiary states (along with associated kinetic parameters). These constraints force the selection of the r* state as the less stable, more highly reactive tertiary state. [d] Jones et al. (1992) showed that both ligand rebinding (V_1) and conformational (V_2) kinetics are independent of the number of ligands bound at times less than ∼1 μs.

solvent ligand concentration.) It is then possible to compute an equilibrium ligand binding curve from these parameters, by summing the properly weighted populations of all the liganded states. This allows us to construct an equilibrium binding curve for the CO ligand without introducing additional parameters, and thereby obtain a thermal CO dissociation rate, $k_{diss}(r^*x)$, which is too small to be determined by the kinetic data, from fits to the equilibrium data. This description of CO binding permits us simultaneously to fit the available kinetic and equilibrium data on CO binding by varying only the kinetic parameters.

The equilibrium data used in this analysis were taken from Perrella et al. (1990, 1995) and consisted of a CO binding curve (saturation versus CO concentration) and a set of populations of each of the five ligation states as a function of fractional saturation with CO. The solution conditions for these measurements (20 °C, pH 7.0, 0.1 M chloride) were somewhat different from those of Jones et al. (20 °C, pH 7.0, 0.1 M phosphate). To adjust for this difference we made use of two-state MWC model parameters tabulated by Imai (1982) for oxygen binding under these two sets of anionic conditions (pH 7.4, 25 °C) which indicate that for chloride the allosteric parameter L is ∼10-fold lower and c is slightly higher so that Lc^4 is the same as for phosphate. In fitting to the equilibrium data we introduced scale factors λ_{Cl} and γ_{Cl} relating L and c, respectively, for the two anions ($L_{Cl} = \lambda_{Cl}L$ and $c_{Cl} = \gamma_{Cl}c$), as two additional fitting parameters, constrained so that $3 < 1/\lambda_{Cl} < 30$ and $0.1 < \lambda_{Cl}\gamma_{Cl}^4 < 10$. All other parameters required to construct the partition function for CO binding were carried over from the kinetic model.

The partition function for the binding of other ligands may also be constructed from the kinetic parameters for CO binding. To this end, we assume further that the kinetics of the outer ligand entry and exit processes do not depend on the nature of the ligand; all of the ligand dependence of the binding therefore appears in the heme binding and dissociation steps, as found for myoglobin (Henry et al., 1983a). We can then construct binding curves for any other ligand by introducing ligand-specific equilibrium constants K_R and K_T only for these steps, replacing the CO-specific ratios $k_{gem}(r)/k_{diss}(r)$, etc., in the partition function.

The equilibrium binding curve for oxygen has been very well characterized over a wide range of solution conditions [see e.g. Imai (1982)]. Therefore, as an adjunct to the fits to the kinetic model and the equilibrium CO data we also included a simultaneous fit to an oxygen binding curve synthesized over a wide range of pressures using the Adair parameters for the appropriate solution conditions (Imai & Yonetani, 1975). The simulated binding curve was computed from the partition function using the current values of the kinetic parameters and the values of two oxygen-specific equilibrium constants for the geminate step, which were treated as additional adjustable parameters in the fits.

Determination of Best-Fit Parameter Sets. The fitting operation involved selecting initial values for all of the adjustable parameters and initiating an algorithmic search for a set of parameters that minimized the residuals between the simulated and experimental data sets. As is common with most procedures of this type, minimizations starting from different initial parameter values produce different final parameters due to trapping in local minima. Given the number of adjustable parameters, an exhaustive search of parameter space to identify all local minima was not feasible. However, a partial survey was possible using Monte Carlo and simulated annealing techniques. The procedure that we used began by assigning randomly-chosen values to all adjustable parameters (within physically plausible bounds),

which were then varied simultaneously in a Monte Carlo search for a parameter set that minimized the sum-of-squared residuals between the observed and calculated data. The Metropolis criterion for acceptance was employed in which the sum-of-squared residuals, ρ, was treated as the "energy". Steps which decreased ρ were always accepted, while steps involving an increase in ρ, $\delta\rho$, were accepted with a probability $\exp(-\delta\rho/\tau)$, where τ is the simulation "temperature". Also, steps were rejected which entered a region of parameter space that violated one or more of the specified constraints on the relations among parameters. Initially a high τ was employed to allow for an effective sampling of the accessible parameter space, irrespective of the presence of "barriers" between local minima. During the course of a series of such steps, τ was gradually lowered according to a pre-determined schedule; this "cooling" had the effect of drawing the system in a generally downhill direction in ρ, while the finite τ still enabled traversal of barriers.

Several hundred distinct simulations were performed, each involving at least 1000 successful Monte Carlo steps, during which an aggressive annealing schedule [$\tau(\text{step } n) \sim \tau_0 n^{-1/2}$] was followed. For this analysis, contributions to the sum-of-squared residuals from the fits to the limited sets of equilibrium CO binding data were omitted. Therefore, the sum-of-squared residuals consisted of weighted contributions from the comparison of three simulated and experimental data sets: the time-resolved spectra, the total population of unrecombined deoxyhemes obtained from V_1, and the equilibrium oxygen binding curve. The final parameter sets from each annealing simulation were used to initiate Marquardt−Levenberg minimizations. Each such calculation, terminated after a finite number of iterations, yielded an approximate minimum for the residuals in parameter space.

A screening procedure was then applied to the ~300 approximate minima produced in this way, to cluster those minima lying in the same sum-of-squares "wells" in parameter space. The sum-of-squared residuals were evaluated along linear paths connecting all possible pairs of minima. Any two minima connected by a path along which the sum-of-squares did not pass through a maximum—i.e., the two minima were not separated by a barrier—were grouped together.

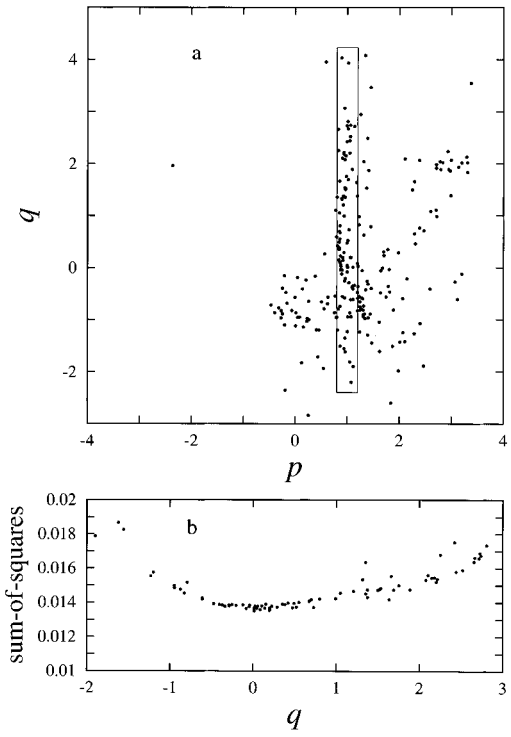

FIGURE 4: Analysis of the principal cluster of minima from the Monte Carlo/simulated annealing study. (a) Least-squares projections of the parameter sets for the 249 minima in the cluster onto a two-dimensional pseudo-parameter space. The logarithms of the 15 parameter values for each minimum [the first 16 parameters listed in Table 1, excluding $k_{diss}(r)$] were computed, except for β and the scale factor f, for which linear variations were kept. The 249 × 249 matrix Δ was computed consisting of all pairwise Euclidean distances between the resulting transformed parameter sets. A pseudo-parameter pair (p_i,q_i) was associated with each minimum, and a least-squares minimization algorithm was used to determine values for the $2(249) = 498$ parameters such that the synthetic data matrix constructed from pairwise distances between all pseudo-parameter pairs optimally matched the full distance matrix Δ (Henry, 1993; Levitt, 1983). The resulting (p_i,q_i) pairs are plotted as points on a $p-q$ plane. (b) Extraction of the set of points of approximately constant p contained within the box in a, showing the actual sum-of-squares value for each minimum vs the q coordinate. The tendency toward a "global" minimum near $q = 0$ may be clearly seen.

RESULTS

Figure 3 shows a subset of the kinetic data of Jones et al. (1992) for photodissociation of the carbon monoxide complex of human hemoglobin at room temperature. The data are presented in the form of a singular value decomposition of the time-resolved spectra between 10 ns and 68 ms. Results from the five levels of photolysis that were used in fitting with the model are shown.

The data can be well represented by two SVD components. The basis spectrum of the dominant component (U_1) corresponds to the average spectrum and is very similar to an HbCO-minus-Hb static difference spectrum. The second basis spectrum (U_2) represents the deviation from the average spectrum, and describes the shift and change in amplitude of the deoxyheme photoproduct. Because there is only a very small feature in U_2 near the sharp peak of the CO complex at 419 nm, U_2 contains very little contribution from ligand rebinding *per se*. The time course of the amplitude of U_1 (V_1) therefore is an excellent approximation to the ligand rebinding curve. The time course of the amplitude of the second basis spectrum (V_2) contains three contributions. One is the kinetics of the spectral changes of the deoxyhemes caused by protein conformational changes, the second is an overall decrease in amplitude due to the depletion of deoxyhemes by ligand rebinding, and the third is spectral changes from kinetic hole burning. There are two types of kinetic hole burning: tertiary and quaternary. Tertiary kinetic hole burning occurs because geminate rebinding in the presence of the two tertiary conformations r* and r occurs on a time scale comparable to or slower than equilibration of these species. Faster geminate rebinding to r* results in an increase in the fractional population of r, producing a deoxyheme spectral change. Similarly, quaternary kinetic hole burning occurs because bimolecular rebinding to R increases the *relative* fraction of T.

Parameters of the Model. The simulated annealing/minimization procedure described above, in which parameter space was first explored by Monte Carlo sampling followed by minimization, identified nine apparently distinct clusters of minima (wells). Only two of these were significantly

Table 2: Parameter Values Resulting from Least-Squares Fits

parameter	range of values[a]	median value[a]	representative set[e]
L	$1.2 \times 10^7 - 1.4 \times 10^8$ [b]	2.2×10^7	2.2×10^7
c	$9 \times 10^{-4} - 2 \times 10^{-3}$ [b]	1.6×10^{-3}	1.6×10^{-3}
l	1.5−22	12	3
k(r*→r)	$2.4 \times 10^6 - 2.1 \times 10^7$ s^{-1}	5.1×10^6 s^{-1}	5.1×10^6 s^{-1}
k_{gem}(r*)	$5.5 \times 10^7 - 5.5 \times 10^8$ s^{-1}	1×10^8 s^{-1}	9.9×10^7 s^{-1}
k_{gem}(r)	$1.7 \times 10^5 - 4 \times 10^5$ s^{-1}	2.6×10^5 s^{-1}	2.6×10^5 s^{-1}
k_{gem}(t)	$5 \times 10^3 - 1.2 \times 10^4$ s^{-1}	7.5×10^3 s^{-1}	7.6×10^3 s^{-1}
k_{diss}(r)	$<10^{-4}$ s^{-1} [c]	$<10^{-5}$ s^{-1}	0.005 s^{-1}
k_{in}	$8.6 \times 10^4 - 2 \times 10^5$ s^{-1}	1.3×10^5 s^{-1}	1.3×10^5 s^{-1}
k_{out}	$5.8 \times 10^6 - 6.6 \times 10^6$ s^{-1}	6.1×10^6 s^{-1}	6.2×10^6 s^{-1}
k_{30}(R→T)	5×10^2 s^{-1} [d]	5×10^2 s^{-1}	5×10^2 s^{-1}
d	4.7−5.6	4.9	4.8
β	0.21−0.36	0.29	0.30
K_R	$2.1 \times 10^5 - 8.2 \times 10^5$	3.2×10^5	3.2×10^5
K_T	$6.7 \times 10^2 - 1.7 \times 10^3$	1.0×10^3	1.0×10^3
λ_{Cl}		0.03	
γ_{Cl}		1.3	

[a] The analysis was performed as follows: the 70 points having the lowest sum-of-squares values (<0.014) were selected from the 249 points in the principal cluster of minima (Figure 4) as being least likely to have parameter values lying well outside the desired distributions. Means and standard deviations were computed for each parameter for these 70 parameter sets [except for k_{30}(R→T), which exhibited no variation] and the 60 parameter sets for which every parameter fell within two standard deviations of its respective mean were retained. The median value and the range of values assumed by each parameter from this set are tabulated. [b] $Lc^3 = 0.1$ (a constrained limit) for each parameter set. [c] This parameter was allowed to vary freely in this analysis but is essentially undetermined by the least-squares criteria used. [d] This parameter assumed the constrained limiting value in all parameter sets. [e] See text.

occupied. The lowest-lying well (minimum sum-of-squares ≈ 0.014) contained by far the largest number of minima. Figure 4 shows the structure of the largest cluster as a two-dimensional projection in a pseudo-parameter space. The spread in the distribution of minima in this projection reflects the variability of the parameter sets that comprise this cluster. Much of this variability arises from the existence of subsets of parameters which are highly correlated in the vicinity of this well—i.e., two or more parameters which may vary by large amounts, but in a correlated fashion, such that the sum of squares changes relatively little. (The existence of such correlated sets of parameters also tends to degrade the convergence properties of typical minimization algorithms.) Two prominently correlated parameters in this cluster are L and c; their high degree of correlation arises from the strong tendency of all the fits to place the product Lc^3 at the constrained limiting value 0.1, while allowing L alone to vary over almost 3 orders of magnitude. Within this cluster there are also highly correlated variations of the three rate parameters k(r*→r), k_{gem}(r*), and k_{gem}(r), as well as an apparent strong coupling among the rate parameters k_{gem}(r), k_{in}, and k_{out}. The correlations are not surprising because these parameters dominate the determination of the largest amplitude decays, namely, geminate ligand rebinding and bimolecular rebinding to the R state.

These strongly correlated variations within certain sets of parameters in this principal cluster make the determination of a unique set of parameters difficult. Given the experimental uncertainties, parameter sets corresponding to all of the lowest-lying points in the cluster must be regarded as nearly equivalent in their ability to reproduce the data. Table 2 summarizes the results as median values and ranges of parameters found in this cluster. k_{diss}(r) was not determined by the least-squares criteria provided in the initial analysis. As discussed above, it was possible to obtain fitted values for this parameter, as well as the equilibrium parameters λ_{Cl} and γ_{Cl}, by expanding the analysis to include fits to the equilibrium CO binding data of Perrella et al. (1990, 1995). A single fitting operation was performed including these additional contributions to the sum-of-squares, using the median parameter values in Table 2 as starting values. The resulting set of parameters is listed in the last column of Table 2. The only substantial change in parameter value brought about by including the CO binding data is a reduction in l, which, like L and c, is presumably condition dependent and therefore different for the kinetic and equilibrium experiments. However, this final value is still within the relatively "soft" range of values of this parameter represented in the principal cluster of minima.

A comparison of the simulated kinetic data set, obtained using this set of parameters, with the measured data set is shown in Figure 3. The simulated and experimental data sets are represented as amplitude vectors obtained by least-squares fitting the complete data set to linear combinations of the basis spectra from the global SVD of the experimental data set. The most significant amplitude vector, which to a good approximation reflects the temporal dependence of spectral changes due to ligand rebinding, is fit very well at all degrees of photolysis. The variations in the second amplitude vector are also fit well but with systematic deviations in the $10^{-5}-10^{-3}$ s time range. [The third amplitude vector (not shown) is only about 10% as large as the second vector and consequently quite noisy, but is nonetheless qualitatively reproduced by the fit.]

We should emphasize that fitting procedures must be constrained in some way to produce physically-plausible spectra, which is generally a difficult computational problem. The only explicit constraint on the tertiary species spectra produced by modeling is that they combine in proportions dictated by the species populations to produce an adequate simulation of the experimental data sets. A useful post facto test of these spectra is provided by the T-minus-R difference spectrum for zero-liganded tetramers, which has been extensively studied (Bellelli & Brunori, 1994). According to our model, the molecule in the unliganded R quaternary state consists of a mixture of r and r* subunits in proportions dictated by the parameter l, and the unliganded T quaternary state consists entirely of t subunits. The difference spectrum

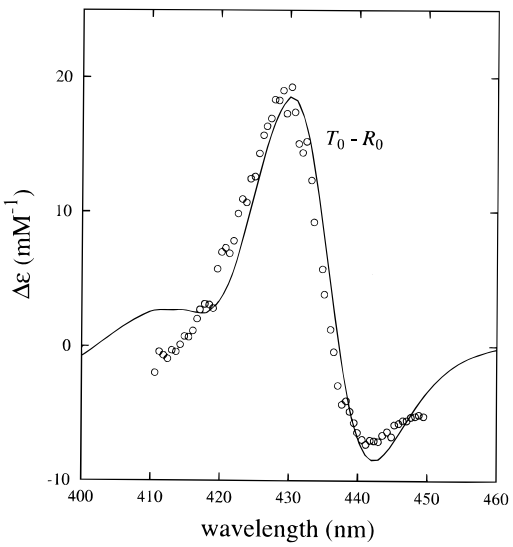

FIGURE 5: Difference spectrum between the T and R quaternary states in the unliganded state synthesized from the tertiary species spectra determined in the kinetic modeling (solid curve), with the corresponding measured difference spectrum provided by Bellelli and Brunori (1994) (circles).

between these two states therefore incorporates contributions from all three tertiary species spectra. A comparison, with no scaling, between a synthetic difference spectrum produced from these species spectra and a measured T-minus-R difference spectrum for the unliganded state is shown in Figure 5. A small (∼1 nm) wavelength shift, as well as some irregularities in the synthetic spectrum in the vicinity of the CO absorption peak (possibly arising from small nonlinearities in the transient spectrometer), are the only differences between these spectra.

The model allows the quaternary transition rate to depend on both the degree of ligation and the number of r* subunits in the molecule. The allosteric parameters L and c are used to relate the R → T and T → R rates for specific ligation/tertiary configurations, and linear free energy relations are used to relate the R → T or T → R rates for different ligation/tertiary configurations. In this way the description of the complex manifold of possible quaternary transition rates is reduced to the specification of a single reference rate and a linear free energy parameter, in addition to the allosteric L and c and the tertiary equilibrium constant l. The reference rate chosen was the R → T transition rate for triply-liganded molecules containing no r* subunits, k_{30}(R→T) (eqs 9 and 10); this rate was constrained to lie within a 2-fold range above and below a value of 10^3 s^{-1}. In almost all cases, the analysis placed the value of this rate firmly against the constrained lower bound of 500 s^{-1}. To enforce consistency with available values for the T → R rate for triply-liganded molecules, the parameter combination Lc^3 was also constrained to lie within a 10-fold range above and below 1. As noted above, the analysis tended to place this combination at the constrained lower bound of 0.1. The combination of these two tendencies, to lower both the R → T and T → R rates for triply-liganded molecules, suggests that the constrained analysis is not allowing the conformational kinetics to become sufficiently slow to fit the measured spectral relaxations at long time delays.

The linear free energy parameter chosen was the scale factor d relating transition rates at successive degrees of ligation (and also by assumption the transition rates for successive numbers of r* subunits in the tetramer, through the derived parameter α). The range of values of this parameter and of the allosteric parameter c produce values of the slope α in the range 0.2−0.25, which is consistent with the value of 0.2 derived by Eaton et al. (1991) from an analysis of the available thermodynamic and kinetic data on quaternary transitions in human hemoglobin. Using $d = 5$ and k_{30}(R→T) = 500 s^{-1} produces a transition rate for zero-liganded molecules of k_{00}(R→T) ≈ 6 × 10^4 s^{-1}. The additional rates k_{01}...k_{04} are successively larger by factors $l^\alpha \approx 10^{0.2} \approx 1.6$, or an increase of ∼6 for four r* subunits in a tetramer. However, the participation of these rates in the observed kinetics is suppressed by the reduced populations of r* subunits (determined by the value of l) toward the end of the tertiary relaxation.

Equilibrium Properties. The evaluation of the kinetic model undertaken here incorporates the requirement that the kinetic parameters also describe the equilibrium binding properties of the molecule. Equilibrium measurements of the binding properties of Hb for both the CO and O_2 ligands were included as constraints on the kinetic parameters at various stages of the analysis. Consistency of the kinetic model with the well-characterized equilibrium binding properties of O_2 is imposed as an auxiliary constraint on certain combinations of model parameters. The equilibrium description of O_2 binding is linked to the kinetic description of CO binding through the allosteric parameter L (=[T_0]/[R_{00}]), the tertiary parameter l, and the kinetic parameters k_{in} and k_{out} for the rates of ligand entry into and exit from the protein, respectively, which are assumed to be independent of ligand type. The additional parameters K_R and K_T are required to describe the ligand-dependent equilibria for the geminate rebinding step in the two quaternary structures. Contact with the more familiar parameters obtained in the MWC analysis of ligand binding curves is easily accomplished using the model parameters listed in Table 1: the allosteric parameter c for O_2 binding may be expressed as $K_T/K_R \approx 0.003$, and the overall association constants for O_2 binding to the R and T states (at the O_2 pressure of 760 mmHg equivalent to the presumed CO pressure in the kinetic experiments) are $K_{R,T}(k_{in}/k_{out}) \approx 7000$ and 200 atm^{-1}, respectively. Figure 6 shows the fit to the O_2 binding curve using the parameter set listed in the last column of Table 2. As required, the fit is excellent, deviating significantly only in a range of pressures for which the original Adair model parametrization of the binding curve is no longer meaningful.

Consistency of the kinetic model with the available equilibrium binding data for the CO ligand is more diagnostic of the success of the model because several more kinetic parameters are required. Here the connection between the kinetic and equilibrium descriptions involves the allosteric parameters L and c as well as the kinetic parameters k_{in} and k_{out} for ligand entry/exit, k_{gem}(r) and k_{diss}(r) for binding/dissociation at the heme, and the tertiary equilibrium parameter l. The additional parameters λ_{Cl} and γ_{Cl} were required to make small adjustments for differences between the solution conditions of the kinetic and equilibrium measurements. Figure 7a shows a comparison of the ligand binding curve and populations of liganded intermediates computed from these parameters to the corresponding measured data of Perrella et al. The ligand binding curve is well fit with the equilibrium parameters $L_{Cl} = \lambda_{Cl}L \approx 0.03 \times 2.2 \times 10^7 = 6.6 \times 10^5$, $c_{Cl} = \gamma_{Cl}c = 1.3 \times 1.6 \times 10^{-3} =$

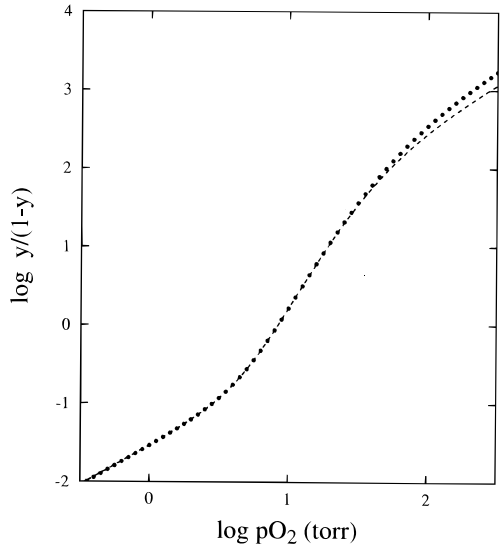

FIGURE 6: Oxygen equilibrium binding curve: (dots) curve obtained from Adair parameters of Imai and Yonetani (1975); (dashed) curve produced from kinetic parameters of the model.

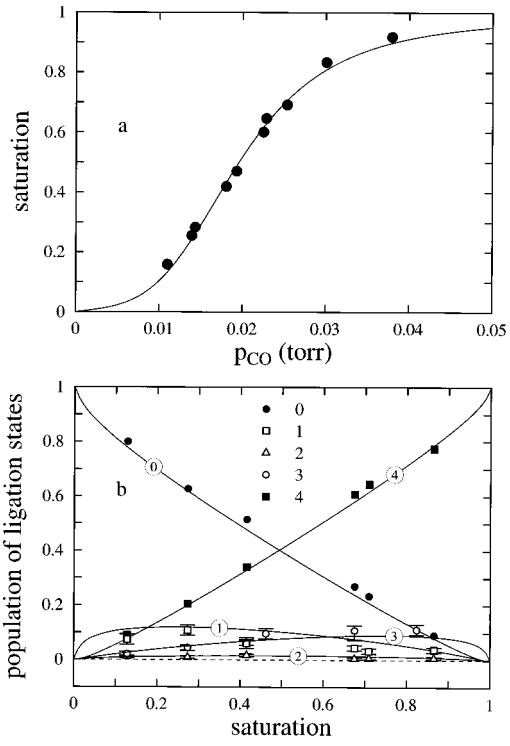

FIGURE 7: Comparison of equilibrium data of Perrella et al. (1990, 1995) with predictions from the kinetic parameters of the model. The conditions of these experiments were 3 mM in heme, 0.1 M KCl, pH 7, 20 °C. (a) CO binding curve, experimental (filled circles) and the synthetic binding curve predicted by the model (continuous curve). (b) Fractional populations of ligation states versus total saturation by CO. Points and error bars represent the measured values: zero-liganded (filled circles), singly-liganded (open squares), doubly-liganded (open triangles), triply-liganded (open circles), and quadruply-liganded (filled squares); continuous curves are synthesized from kinetic parameters of the model.

2.1×10^{-3}, and $K_R K_{in} = [l/(1 + l)][(k_{gem}(r)/k_{diss}(r)](k_{in}/k_{out}) \approx 8 \times 10^4$ mM^{-1} (the latter assuming that the kinetic measurements were performed at 1 atm $\approx$ 1 mM CO). The fit to the populations of the five ligation states as a function of saturation is shown in Figure 7b. There are only very small deviations between the simulated and measured populations, but the most difficult feature to reproduce, the similar maximum populations of the one- and three-liganded species, is clearly being approached. The final fitted value of λ_{Cl} is held at the constrained lower limit; relaxing this constraint, thereby allowing a smaller L_{Cl} than was computed above, improves the reproduction of this feature and the overall quality of the fit.

As noted earlier, the temporal evolution of the second amplitude vector is determined both by the decay in the overall amplitude of the data set due to ligand rebinding (as approximately reflected by the first amplitude vector) and by changes in the spectra of the remaining deoxyhemes with time. To a good approximation, we can extract from this vector the evolution of the deoxyheme spectral changes per remaining deoxyheme by dividing the amplitude at each time by the current deoxyheme population, which is well approximated by the first amplitude vector. Figure 8 shows the evolution of deoxyheme spectral changes computed in this way for data sets with high and low photolysis yields. Also shown is another measure of the progress of the deoxyheme spectral changes using the wavelength of the zero crossing of the HbCO−Hb difference spectrum in the vicinity of 425 nm; in the absence of deoxyheme spectral changes, this crossing point should be a true HbCO−Hb isosbestic point with fixed wavelength. The two measures of deoxyheme spectral changes are seen to be quite consistent. The deoxyheme spectrum evolves continuously over the time range from nanoseconds to milliseconds. In the kinetic modeling we are able to describe this evolution in terms of the temporally extended interconversion between populations of r* and r subunits, which persists until about 10 μs, and the subsequent interconversion between the r* + r population of the R quaternary structure and the t subunits of the T quaternary structure. At high degrees of photolysis these two contributions merge seamlessly to produce a continuous relaxation until the millisecond time scale.

Inequivalence of α *and* β *Subunits.* All of the analysis described to this point has made use of the simplest of the class of two-state allosteric models which incorporate tertiary conformational changes in the R quaternary structure as well as geminate rebinding. This 85-state model, which treats the four subunits in the tetramer as identical, has produced an excellent overall description of the kinetic and equilibrium data presented to it. The clearest, possibly significant discrepancy is the small difference between the measured and predicted amplitudes of the second component vector in the time range 10^{-6}−10^{-3} s (Figure 3b) of the kinetic data. This defect is sufficiently small so that a complete accounting may need to include possible instrumental effects as well as spectroscopic perturbations outside the scope of such modeling (e.g., heme−heme electronic couplings). Possible alternative explanations notwithstanding, it is instructive to ask whether these deviations may be reduced by taking the analysis to the next level.

As discussed earlier, the simplest extension of the basic model introduces inequivalence between α and β subunits, while still requiring that both subunits of a given type behave identically. This extended model requires 225 R + 36 T = 261 states. Moreover, most of the parameters used in the basic model will "split" into distinct possible values for α and β subunits. Because of the ~3-fold larger set of states and ~2-fold larger set of adjustable parameters, the computational effort required in applying this model is considerably greater than that required for the basic model. There

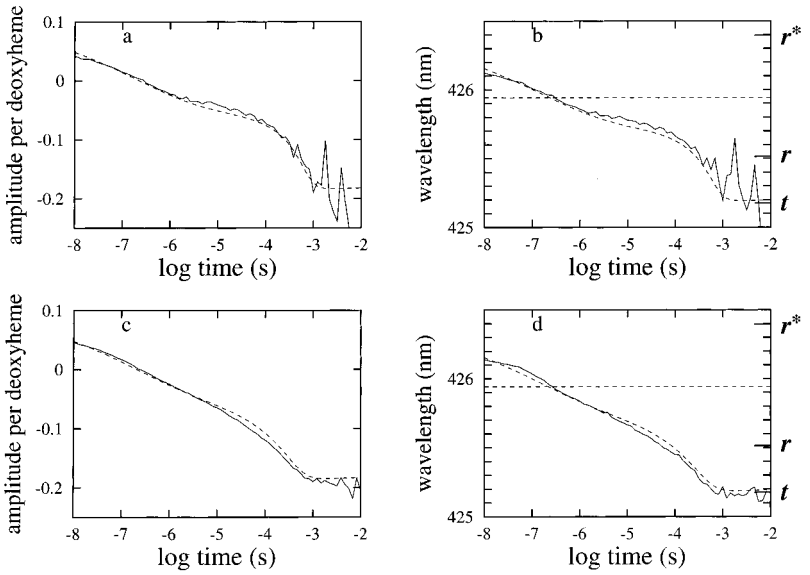

FIGURE 8: Time course of deoxyheme spectral changes from data sets corresponding to 30% photolysis (a,b) and 80% photolysis (c,d). (a,c) Time courses determined by normalizing the second amplitude vector from the global SVD corresponding to each data set using the corresponding first amplitude vector (see text). (b,d) Time course of the wavelength of the zero-crossing of the Hb−HbCO difference spectrum: experimental data set reconstructed from first 12 components of the global SVD of original data (solid), synthetic data set produced by kinetic modeling (dashed). The three labeled wavelengths on the right-hand vertical axis indicate the zero-crossing wavelengths of the deoxy−CO difference spectra of the tertiary species produced by the modeling.

is also a compounding of the number of possible spectroscopically distinct subunit species, because the tertiary spectra may in principle differ between α and β subunits. However, if we assume that the subunit spectra for a given tertiary species differ only by a wavelength shift, then matrix manipulations are possible that enable the least-squares fitting problem to be reduced to an equivalent description in terms of the spectra of a single subunit type plus the introduction of one or more adjustable wavelength-shift parameters (Henry, 1997).

In evaluating this extended model we have chosen to impose limits on the permitted size of the α−β inequivalence (i.e., the differences between corresponding parameters associated with the two subunit types). For most elementary parameters the "splitting" between the values for the two subunit types was constrained to be less than a multiplicative factor of 2, that is, $p_\alpha/2 < p_\beta < 2p_\alpha$. Because of the added complexity of this model, it has not been possible to systematically explore parameter space in search of low-lying minima. However, extended-model minimizations started in the vicinity of the lowest-lying minimum for the basic model (by initially assigning the single optimal value of each parameter from the basic model to both α and β parameters in the extended model) generally succeeded in lowering the sum-of-squares by an additional 30% by adjusting the mean value and/or the α−β inequivalence of some of the parameters. The formulation of this extended model is essentially symmetric with respect to interchange of the identities of the α and β subunits. As a result, only the magnitudes of the individual splittings, and the relative directions of the splittings of different parameters, are meaningful. The only result of fitting with inequivalent α and β subunits which differs significantly from the corresponding result for the basic model is the simulated second amplitude vector (Figure 3b). The simulated curve tracks the experimental vector somewhat more closely at intermediate times.

DISCUSSION

In this work we have carried out exhaustive modeling of a set of time-resolved spectra following varying degrees of photolysis of the carbon monoxide complex of hemoglobin. Because large populations of intermediate ligation states are generated by photolysis, the kinetics make much greater demands on a model than equilibrium studies in which they are sparsely populated. Moreover, different distributions of ligation states are obtained by varying the degree of photolysis. We have fit this data with the simplest possible extension of the two-state allosteric model (Monod et al., 1965), without altering the fundamental postulate of the model that ligand binding within each quaternary structure is non-cooperative. The extension was required to include geminate ligand rebinding and the "modern" kinetic results from time-resolved spectroscopy that there is a structural relaxation following photodissociation before any quaternary structural changes (Friedman & Lyons, 1982; Hofrichter et al., 1983). It is also consistent with the optical polarization (Makinen & Eaton, 1974) and X-ray crystallographic results (Perutz et al., 1987) which show differences in protein conformation in the crystalline state between the fully liganded and fully unliganded hemoglobin, both in the R quaternary structure.

Formally, we postulate two tertiary conformations for subunits in the R quaternary structure, called r and r* (Figure 2). According to the model the r* conformation is favored by ligation, while the r conformation is favored by deligation. Before ligand photodissociation the molecule is in the R quaternary structure, with all four subunits in the r*x state. Photodissociation populates the r*·x state (with the dissociated ligand in the protein and the affected subunit in the r* tertiary conformation) to an extent determined by the photolysis pulse intensity. Subunits in this state may undergo a conformational relaxation (r*·x → r·x), have the ligand geminately rebind (r*·x → r*x) or escape into the solvent (r*·x → r* + x). At later times, a quaternary conformational

change from R to T will occur (to an extent determined by the number and conformation of deoxy subunits remaining after geminate rebinding), converting subunits in the r and r* conformation to the t conformation. Ligands from the solvent may enter any of the ligand-free subunits ($r+x \rightarrow r{\cdot}x$, $r^*+x \rightarrow r{\cdot}x$, $t+x \rightarrow t{\cdot}x$) and either escape back into the solvent or bind to the heme.

Our initial attempts to fit these data ran into difficulties because the tertiary relaxation phase, defined by deoxyheme spectral changes prior to either kinetic or spectroscopic evidence for R → T transitions, was extended in time and required three relaxations when described with exponential functions [relaxation times of 20, 90, and 820 ns (Jones et al., 1992)]. Since there did not appear to be any justification for introducing more than two tertiary conformations, and there was no evidence for clear separation of the three relaxations, we took a different approach motivated by the conformational relaxation kinetics observed in myoglobin (Ansari et al., 1992, 1994; Lim et al., 1993; Jackson et al. 1994). As in myoglobin the extended nature of the time course for the tertiary structural relaxation is well described by a stretched exponential function, $\exp[-(kt)^{\beta}]$. We therefore wrote the differential equations in a form that yielded a stretched-exponential function upon integration (eqs 2−6). Justification for the stretched exponential function was provided by the work of Hagen and Eaton (1996). They showed that this type of extended time course is a consequence of slow equilibration among conformational substates. It also required a different approach to the treatment of geminate ligand rebinding, since geminate rebinding does not occur with only two rates, one corresponding to r* and one corresponding to r. Intermediate rates must be considered which scale with the progress of the conformational relaxation (eqs 11 and 12).

The kinetics of the quaternary conformational change are also influenced by the stretched exponential tertiary relaxation, which is not yet complete at times when the R → T quaternary transition begins. To deal with this complexity, as well as the large number of rates, we introduced a linear free energy relation between the rates and equilibrium constants (eqs 7 and 10), as suggested from the analysis of Eaton et al. (1991). The use of a linear free energy relation produces an enormous simplification in the model, because 30 quaternary transition rates could be described with only five parameters.

In addition to testing the model by its ability to fit the kinetic data, we have also shown that the parameters of the model can simultaneously reproduce the well-known T-minus-R difference spectrum (Figure 5), the equilibrium oxygen binding curve (Figure 6), the equilibrium carbon monoxide binding curve (Figure 7a), and the equilibrium distribution of ligation states (Figure 7b). In the case of the oxygen binding curve we utilized the same allosteric parameter L for the equilibrium between the zero-liganded tetramers as in the kinetic fitting. For the carbon monoxide equilibrium data, we used, in addition to L, the tertiary equilibrium constant l and the kinetic parameters that describe geminate rebinding from the r*, r, and t subunit conformations. The dissociation rate $k_{diss}(r)$ was an adjustable parameter, while the rate $k_{diss}(r^*)$ was determined by the constraint of a thermodynamic cycle.

A total of 17 parameters, all with well-defined physical significance, were varied in the fitting procedure (Table 1). One might argue that fitting data with so many parameters is not a valid test of a model because the problem is underdetermined. Of the 17 parameters, 12 pertain to the kinetic data. A purely empirical description of the kinetic data, in terms of sums of exponentials, requires 12 amplitudes for each of the five different degrees of photolysis, plus seven relaxation times, or a total of 67 parameters. Because the *relative* amplitudes are the same for the first three relaxations at less than 1 μs, and because the lowest degrees of photolysis require only three relaxations for an adequate fit, the total number of parameters required for an empirical description could be as low as ∼30. Nevertheless, the number of empirical parameters still far exceeds the number of model parameters.

The fits to the data are excellent, but not perfect. The model slightly overestimates the amplitude of the deoxyheme spectral changes (V_2, Figure 3b) at low degrees of photolysis and underestimates it at high degrees of photolysis. It is not yet clear whether these small differences have any physical significance. The magnitude of the difference between the simulated and experimental data sets, computed as $\sum_{t,\lambda}|A_{sim}(\lambda,t) - A_{exp}(\lambda,t)|/\sum_{t,\lambda}|A_{exp}(\lambda,t)|$, is 1.3%, decreasing to 1.1% when the inequivalence of the α and β subunits is introduced. One possible explanation for small deviations is the presence of electronic heme-heme interactions, which would result in deoxyheme and liganded heme spectra that are not independent of the number of ligands that are bound. The presence of dimers has not been included in the analysis, but would decrease the photolysis dependence of the amplitude of V_2, which is opposite from the deviation observed between the fitted and measured curves. Furthermore, from tetramer−dimer constants determined for HbCO by flash photolysis under the same solution conditions (Antonini et al., 1967), less than 8% of the hemes are in dimers at 120 μM. To test the model further it would be useful to investigate the concentration-dependence in more detail, and also to measure the effect of CO pressure on the kinetics, which would provide a more detailed description of the competition between quaternary switching and bimolecular rebinding.

It is instructive to comment on each of the parameter values, since another requirement of a successful model is that the parameters not only have physically reasonable values, but also compare sensibly with parameters obtained in other studies. The parameters are defined in Figure 2 and Table 1, and the values obtained from the fitting are listed in Table 2.

The values of the MWC allosteric parameters L and c in the literature have been derived by fitting oxygen equilibrium curves. The value of $L \approx 2 \times 10^7$ obtained by fitting the kinetic data is ∼10-fold higher than usually reported but, as can be seen from the simultaneous fit to the oxygen equilibrium curve in Figure 6, reproduces the data extremely well. On the other hand, c, which describes the ratio of the T and R state affinities is similar for carbon monoxide and oxygen [$c(O_2) \approx 0.003$, $c(CO) \approx 0.002$]. The median value of the linear free energy parameter α [$\equiv -(\log d/\log c)$] that describes the quaternary transition rates is 0.25, close to the value of 0.18 found by Eaton et al. (1991) in a simpler analysis of quaternary transition rates and equilibria measured in different laboratories. This parameter has interesting physical significance. It can only take on values between 0 and 1 (assuming the transition state has thermodynamic

properties intermediate between the R and T structures). A value of 0 indicates that the transition state has the thermodynamic properties of the R state, while $\alpha = 1$ indicates that it has the thermodynamic properties of the T state. In a naive but interesting interpretation, intermediate values measure the fractional "distance" along the reaction coordinate from R to T (Eaton et al., 1991). The value of $\alpha \approx 0.2$ suggests that the transition state is closer to the R state. Surprisingly, this is exactly the value that Eaton et al. (1991) derived from Janin and Wodak's reaction path for the quaternary transition using only buried surface area as a measure of the stabilization free energy (Janin & Wodak, 1985).

The value of the tertiary equilibrium constant, $l \approx 10$, while not precisely determined, indicates that the immediate photoproduct conformation, r*, is mostly a transient, but is significantly populated (~10−20%) in the completely deoxygenated R state at equilibrium. The rate of the r* → r conformational change is 5×10^6 s^{-1}. Because the relaxation is extended, however, with a stretched exponential parameter $\beta \approx 0.3$, it still contributes to the spectral changes at several microseconds. This value of β is higher than the value of 0.1 found for the conformational relaxation in myoglobin by Anfinrud and co-workers (Lim et al., 1993; Jackson et al., 1994; Hagen & Eaton, 1996). However, it is not yet known how much conformational relaxation occurs on a subnanosecond time scale. The 1/e time observed for the conformational relaxation in myoglobin is very much shorter, ~10 ps (Lim et al., 1993; Jackson et al., 1994; Hagen & Eaton, 1996), indicating that not only is the conformational relaxation very much slower in hemoglobin but that the interconversion of conformational substates is also much slower. Studies of the viscosity dependence of the conformational relaxation in myoglobin by Ansari et al. (1992, 1994) showed a simple $1/\eta$ Kramers behavior, indicating a global motion of protein atoms, such as might occur in a slight rearrangement in the packing of helices. The global nature of the conformational change is also suggested by the findings of Hagen et al. (1995, 1996) that conformational relaxation and interconversion of conformational substates is prevented even at room temperature when the protein is embedded in a glass ($\eta > 10^{15}$ cP). Perhaps the intersubunit contacts are responsible for slowing conformational relaxation and interconversion of substates in hemoglobin. Support for this idea comes from nanosecond-resolved crystallographic studies of myoglobin which suggest that intermolecular contacts in the crystal considerably slow the conformational relaxations following CO photolysis (Šrajer et al., 1996; Eaton et al., 1996).

With two tertiary conformations in the R quaternary structure (r* and r), and one in the T quaternary structure (t), there are three geminate rebinding rates: k_{gem}(r*), k_{gem}(r), and $k_{gem}(t)$ (Figure 2, Table 1). The geminate rate slows by almost a factor of 400 upon relaxation from r* to r and by another factor of ~30 in the r → t conformational change that accompanies the R → T quaternary structural change. The conformational relaxation, which is occurring on the same time scale as geminate rebinding, has a very large effect on the geminate yield. Given the rate of escape from the heme pocket of 5×10^6 s^{-1}, the geminate yield ($= k_{gem}/[k_{gem} + k_{out}]$) for the r* conformation is 95%, while it is only 5% for the r conformation and only ~0.1% for the t conformation. In myoglobin Hagen et al. (1995, 1996) have

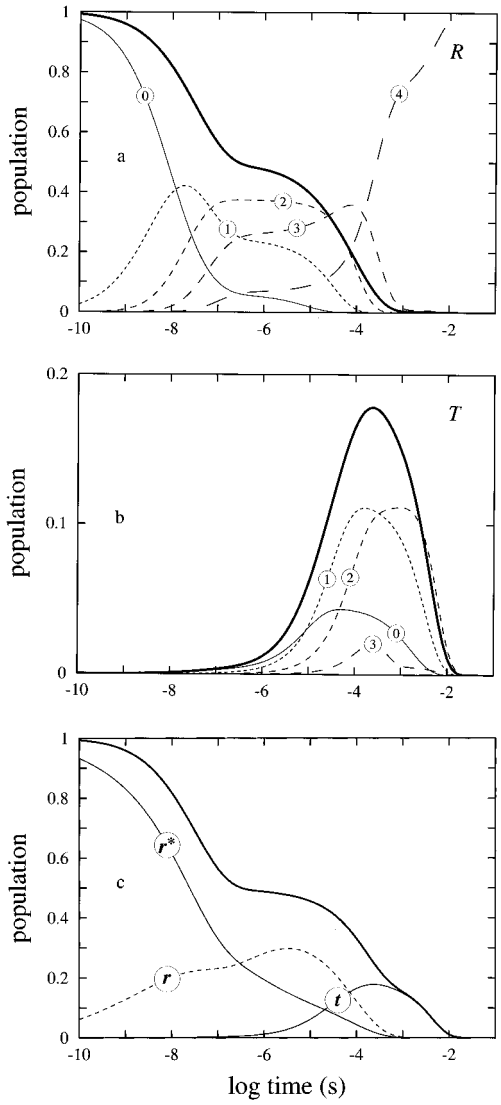

FIGURE 9: Populations of ligation states in the (a) R and (b) T quaternary structures predicted by the model at 100% photolysis. (c) Populations of unliganded tertiary species predicted by the model at 100% photolysis. The thick continuous curve in each panel is the fraction of deoxyhemes. The population ratio of r* and r becomes constant at ~30 μs, indicating that equilibrium has been established. After ~100 μs the t conformation dominates the deoxy subunit populations, because r* and r have been depleted by bimolecular rebinding.

shown that conformational relaxation following photodissociation of CO slows geminate rebinding by a factor of ~2500. The relaxation in myoglobin, however, is much faster so that it is essentially complete before geminate rebinding begins, and there is little competition between conformational relaxation and geminate rebinding (in H_2O). As in myoglobin (Ansari, 1994), the Soret band shifts to the shorter wavelengths as the conformation relaxes to a more slowly rebinding species.

The availability of a complete kinetic and thermodynamic description permits us to calculate the population of each of the 85 molecular species as a function of time at any degree of photolysis. Figure 9 shows two population plots of particular interest at 100% photolysis; one is the fraction of the different ligation states as a function of time, and the other is the population of deoxyhemes in each of the three different conformations, r* and r in the R quaternary structure, and t in the T quaternary structure. The time scales

begin at 100 ps, ~100 times shorter than the beginning of our observations. These plots are useful for designing experiments to test our model, and for also suggesting the magnitude of the amplitude to be observed in the sub-10 ns time scale not resolved in our experiments. For example, the model predicts that after photodissociation with a δ-function light pulse, about 4% geminate recombination is expected at 1 ns. Using excitation with 300 fs pulses and monitoring by infrared spectroscopy Anfinrud et al. (1989) found 4.6% ± 0.7% geminate rebinding at 1 ns, in excellent agreement with the prediction of the model.

CONCLUDING REMARKS

In this work we have critically tested the modern kinetic version of the MWC two-state allosteric model (Monod et al., 1965). We have required that the model fit the kinetics of geminate and bimolecular ligand rebinding, as well as the kinetics of tertiary and quaternary conformational changes, which are contained in a set of very precise time-resolved spectra. These fits simultaneously reproduced three kinds of equilibrium data under very similar solution conditions: the carbon monoxide and oxygen equilibrium binding curves and the distribution of ligation states determined upon electrophoretic separation at partial saturation with carbon monoxide. Our conclusion is that the model satisfactorily explains a demanding set of kinetic and equilibrium data, but the fits are not perfect.

Together with the single crystal oxygen binding curves (Mozzarelli et al., 1991, 1997; Rivetti et al., 1993; Bettati et al., 1996), the kinetic and equilibrium studies in solution described in this work provide strong support for the mechanism of cooperative ligand binding proposed in the original two-state allosteric model of MWC (1965). We should emphasize, however, that we have only dealt with so-called homotropic effects, i.e., interactions between identical ligands, which were the primary concern of Monod et al. (1965). As pointed out earlier, it has been known for a long time that the original MWC model fails to explain so-called heterotropic effects, i.e., the interaction of oxygen (or carbon monoxide) with protons, organic phosphates and other so-called allosteric effectors. In the MWC model allosteric effectors influence only L, the R ⇆ T equilibrium constant, and not the R- or T-state affinities. Szabo and Karplus showed how the MWC model could be expanded to account for the greater complexity of allosteric effectors (heterotropic effects) without altering the mechanism for the origin of cooperative ligand binding (homotropic effects). A next obvious step, in probing structure−function relations of hemoglobin, is to measure the kinetics of ligand rebinding and conformational changes under varying solution conditions, and rigorously model the data as we have done in the present study.

ACKNOWLEDGMENT

We thank Attila Szabo for helpful discussion and a critical reading of the manuscript.

SUPPORTING INFORMATION AVAILABLE

A more detailed description of the implementation of the 85-state kinetic model is available, including two tables and two figures (12 pages). Ordering information is given on any current masthead page.

REFERENCES

Ackers, G., Doyle, M. L., Myers, D., & Daugherty, M. A. (1992) *Science 255*, 54−63.
Agmon, N., & Hopfield, J. J. (1983) *J. Chem. Phys. 79*, 2042−2053.
Anfinrud, P. A., Han, C., & Hochstrasser, R. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8387−8391.
Ansari, A., & Szabo, A. (1993) *Biophys. J. 64*, 838−851.
Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J., & Eaton, W. A. (1992) *Science 256*, 1796−1798.
Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J., & Eaton, W. A. (1993) *Biophys. J. 64*, 852−868.
Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J., & Eaton, W. A. (1994) *Biochemistry 33*, 5128−5145.
Antonini, E., Chiancone, E., & Brunori, M., *J. Biol. Chem. 242*, 4360−4366 (1967).
Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., & Gunsalus, I. C. (1975) *Biochemistry 14*, 5355−5373.
Beece, D., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M. C., Reinisch, L. Reynolds, A. H., Sorensen, L. B., & Yue, K. T. (1980) *Biochemistry 19*, 5147−5157.
Bellelli, A., & Brunori, M. (1994) *Methods Enzymol. 232*, 57−71.
Bettati, S., Mozzarelli, A., Rossi, G. L., Tsuneshige, A., Yonetani, T., Eaton, W. A., & Henry, E. R. (1996) *Proteins 25*, 425−437.
Brunori, M., Coletta, M., & Di Cera, E. (1986) *Biophys. Chem. 23*, 215−222.
Chernoff, D. A., Hochstrasser, R. M., & Steele, A. W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5606−5610.
Cornelius, P. A., Hochstrasser, R. M., & Steele, A. W. (1983) *J. Mol. Biol. 163*, 119−128.
Duddell, D. A., Morris, R. J., & Richards, J. T. (1979) *J. Chem. Soc., Chem. Commun.*, 75−76.
Eaton, W. A., & Hofrichter, J. (1981) *Methods Enzymol. 76*, 175−261.
Eaton, W. A., Henry, E. R., & Hofrichter, J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4472−4475.
Eaton, W. A., Henry, E. R., & Hofrichter, J. (1996) *Science 274*, 1631−1632.
Edelstein, S. J. (1975) *Annu. Rev. Biochem. 44*, 209−232.
Edelstein, S. J. (1996) *J. Mol. Biol. 257*, 737−744.
Ferrone, F. A., Martino, A. J., & Basak, S. (1985) *Biophys. J. 48*, 269−282.
Frauenfelder, H., Alberding, N. A., Ansari, A., Braunstein, D., Cowen, B. R., Hong, M. K., Iben, I. E. T., Johnson, J. B., Luck, S., Marden, M. C., et al. (1990) *J. Phys. Chem. 94*, 1024−1037.
Frauenfelder, H., Sligar, S. G., & Wolynes, P. G. (1991) *Science 254*, 1598−1603.
Gibson, Q. H. (1959) *Biochem. J. 71*, 293−303.
Gill, S. J., Robert, C. H., Coletta, M., Di Cera, E., & Brunori, M. (1986) *Biophys. J. 50*, 747−752.
Goldbeck, S. C., Paquette, S. J., Bjorling, S. C., & Kliger, D. S. (1996) *Biochemistry 35*, 8628−8639.
Greene, B. I., Hochstrasser, R. M., Weisman, R. B., & Eaton, W. A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5255−5259.
Hagen, S. J., & Eaton, W. A. (1996) *J. Chem. Phys. 104*, 3395−3398.
Hagen, S. J., Hofrichter, J., & Eaton, W. A. (1995) *Science 269*, 959−962.
Hagen, S. J., Hofrichter, J., & Eaton, W. A. (1996) *J. Phys. Chem. 100*, 12008−12021.
Henry, E. R. (1997) *Biophys. J. 72*, 652−673.
Henry, E. R., Sommer, J. H., Hofrichter, J., & Eaton, W. A. (1983a) *J. Mol. Biol. 166*, 443−451.
Henry, E. R., Hofrichter, J., Sommer, J. H., & Eaton, W. A. (1983b) in *Hemoglobins: Structure and Function* (Schnek, A. G., & Paul, C., Eds.) pp 193−203, Edition de l'Universite de Bruxelles, Belgium.
Hess, V. L., & Szabo, A. (1979) *J. Chem. Ed. 56*, 289−293.
Hindmarsh, A. C. (1980) *ACM SIGNUM Newsl. 15*, 10−11.
Hofrichter, J., Sommer, J. H., Henry, E. R., & Eaton, W. A. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2235−2239.
Hofrichter, J., Henry, E. R., Sommer, J. H., Deutsch, R., Ikeda-Saito, M., Yonetani, T., & Eaton, W. A. (1985) *Biochemistry 24*, 2667−2679.

Hofrichter, J., Henry, E. R., Szabo, A., Murray, L. P., Ansari, A., Jones, C. M., Coletta, M., Falcioni, G., Brunori, M., & Eaton, W. A. (1991) *Biochemistry 30*, 6583−6598.
Holt, J. M., & Ackers, G. K. (1995) *FASEB J. 9*, 210−218.
Hopfield, J. J., Shulman, R. G., & Ogawa, S. (1971) *J. Mol. Biol. 61*, 425−443.
Imai, K. (1982) *Allosteric Effects in Hemoglobin*, Cambridge University Press, Cambridge.
Imai, K., & Yonetani, T. (1975) *J. Biol. Chem. 250*, 2227−2231.
Jackson, T. A., Lim, M., & Anfinrud, P. A. (1994) *Chem. Phys. 180*, 131−140.
Janin, J., & Wodak, S. (1985) *Biopolymers 24*, 509−526.
Jones, C. M., Ansari, A., Henry, E. R., Hofrichter, J., & Eaton, W. A. (1992) *Biochemistry 31*, 6692−6702.
Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., & Eaton, W. A. (1993). *Proc. Natl. Acad. Sci. U.S.A. 90*, 11860−11864.
Koshland, D. E., Nemethy, G., & Filmer, D. (1966) *Biochemistry 5*, 365−385.
Lee, A., Karplus, M., Poyart, C., & Bursaux, E. (1988) *Biochemistry 27*, 1285−1301.
Levitt, M. (1983). *J. Mol. Biol. 168*, 621−657.
Liddington, R., Derewenda, Z., Dodson, G., & Harris, D. (1988) *Nature 331*, 725−728.
Lim, M., Jackson, T. A., & Anfinrud, P. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5801−5804.
Lyons, K. B., & Friedman, J. M. (1982) in *Hemoglobin and Oxygen Binding* (Ho, C., Eaton, W. A., Collman, J. P., Gibson, Q. H., Leigh, J. S., Margoliash, E., Moffatt, K., & Scheidt, W. R., Eds.) pp 333−338, Elsevier/North-Holland, Amsterdam.
Makinen, M. W., & Eaton, W. A. (1974) *Nature 247*, 62−64.
Marden, M. C., Hazard, E. S., & Gibson, Q. H. (1986) *Biochemistry 25*, 7591−7596.
Monod, J., Wyman, J., & Changeux, J.-P. (1965) *J. Mol. Biol. 12*, 88−118.
Mozzarelli, A., Rivetti, C., Rossi, G. L., Henry, E. R., & Eaton, W. A. (1991) *Nature 351*, 416−419.
Mozzarelli, A., Rivetti, C., Rossi, G. L., Eaton, W. A., & Henry, E. R. (1997) *Protein Sci. 6*, 484−489.
Murray, L. P., Hofrichter, J., Henry, E. R., & Eaton, W. A. (1988a) *Biophys. Chem. 29*, 63−76.
Murray, L. P., Hofrichter, J., Henry, E. R., Ikeda-Saito, M., Kitagishi, K., Yonetani, T., & Eaton, W. A. (1988b) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2151−2155.
Pauling, L. (1935) *Proc. Natl. Acad. Sci. U.S.A. 21*, 186−191.
Perrella, M., & Denisov, I. (1995) *Methods Enzymol. 259*, 468−487.
Perrella, M., Colosimo, A., Benazzi, L., Ripamonti, M., & Rossi-Bernardi, L. (1990) *Biophys. Chem. 37*, 211−223.
Perutz, M. F. (1970) *Nature 228*, 726−739.
Perutz, M. F., Fermi, G., Luisi, B., Shaanan, B., & Liddington, R. C. (1987) *Accts. Chem. Res. 20*, 309−321.
Petzold, L. (1983) *SIAM J. Sci. Stat. Comput. 4*, 136−148.
Rivetti, C., Mozzarelli, A., Rossi, G. L., Henry, E. R., & Eaton, W. A. (1993) *Biochemistry 32*, 2888−2906.
Sawicki, C. A., & Gibson, Q. H. (1976) *J. Biol.Chem. 251*, 1533−1542.
Shibayama, N., & Saigo, S. (1995) *J. Mol. Biol. 251*, 203−209.
Shulman, R. G., Hopfield, J. J., & Ogawa, S. (1975) *Q. Rev. Biophys. 8*, 325−420.
Šrajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M., & Moffat, K. (1996) *Science 274*, 1726−1729.
Steinbach, P. J., Ansari, A., Berendzen, J., Braunstein, D., Chu, K., Cowen, B. R., Ehrenstein, D., Frauenfelder, H., Johnson, J. B., Lamb, D. C., et al. (1991) *Biochemistry 30*, 3988−4001.
Szabo, A., & Karplus, M. (1972) *J. Mol. Biol. 72*, 163−197.
Szabo, A., & Karplus, M. (1975) *Biochemistry 14*, 931−940.

BI9619177