Biochemistry

 Volume 21, Number 20 September 28, 1982

Control and pH Dependence of Ligand Binding to Heme Proteins†

W. Doster, D. Beece, S. F. Bowne, E. E. DiIorio, L. Eisenstein, H. Frauenfelder,* L. Reinisch, E. Shyamsunder, K. H. Winterhalter, and K. T. Yue

ABSTRACT: The recombination after flash photolysis of dioxygen and carbon monoxide with sperm whale myoglobin (Mb), and separated β chains of human hemoglobin (β^A) and hemoglobin Zürich (β^{ZH}), has been studied as a function of pH and temperature from 300 to 60 K. At physiological temperatures, a preequilibrium is established between the ligand molecules in the solvent and in the heme pocket. The ligand in the pocket binds to the heme iron by overcoming a barrier at the heme. The association rate is controlled by this final binding step. The association rate of CO to Mb and β^A is modulated by a single titratable group with a pK at 300 K of 5.7. The binding of CO to β^{ZH}, in which the distal histidine is replaced by arginine, does not depend on pH. Oxygen recombination is independent of pH in all three proteins. Comparison of the binding of CO at 300 K and at low temperatures shows that pH does not affect the preequilibrium but changes the barrier height at the heme. The pH dependence and the difference between O_2 and CO binding can be explained by a charge–dipole interaction between the distal histidine and CO.

Control is a crucial feature of biological reactions. One goal of the investigation of biomolecular processes is an understanding of control on a molecular basis. To this end, a suitable process must be selected and explored in detail. The Bohr effect, the dependence of the oxygen affinity of hemoglobin (Hb)[1] on pH, is the prototype of a biologically important control. Studies of the Bohr effect are complicated, however, by the fact that Hb is a tetramer and tertiary and quaternary structural changes are involved. The binding of small ligands, dioxygen or carbon monoxide, to monomeric proteins, such as myoglobin, or separated chains of hemoglobin[1] is a simpler test case [see, e.g., Szabo (1978)].

The major steps in the binding of a small ligand to a monomeric heme protein are sketched in Figure 1 which represents a schematic cross section through Mb (Antonini & Brunori, 1971; Austin et al., 1975; Case & Karplus, 1979; Beece et al., 1980). A small molecule, coming from the solvent S, migrates through the globin matrix M to the heme pocket B. From B, it either moves back into the globin or binds covalently to the heme iron (state A). If the bond between the iron and the ligand is broken by a light pulse, the ligand returns to state B. Rebinding from B depends on temperature. Below about 180 K, the ligand remains in the pocket and rebinds directly. Above about 180 K, some ligands move into the globin matrix before returning to the pocket and binding. At around 300 K, nearly all ligands pass through the globin matrix into the solvent; all ligands in the solvent then compete for the free binding site. We will show that the association rate is controlled by the properties of the heme pocket and by the transition B → A. The transition B → A involves the protein, the heme, the iron atom, and the ligand molecule. In B, the heme is domed toward the proximal histidine, the iron has spin 2 and lies out of the mean heme plane, and the ligand is somewhere in the heme pocket. In A, the heme is nearly planar, the iron has spin 0 and lies close to the mean heme plane, and the ligand is covalently bound to the iron. In the step B → A, the ligand moves toward the iron, the heme flattens, the iron changes spin and moves into the heme plane, and the protein may undergo some structural rearrangements (Warshel, 1977). The ligand passes close to the distal histidine, His E7, while approaching the iron. The rate B → A can consequently be influenced by the ease with which the protein rearranges, the heme flattens, and the iron moves into the plane and by the hindrance encountered by the ligand, for instance, from His E7.

Evidence that His E7 is implicated in the binding of O_2 and CO to Hb and Mb comes from evolutionary observations, from neutron and X-ray diffraction, from IR and NMR spectra, and from EPR data on cobalt-substituted Hb. His E7 is highly conserved (Dayhoff, 1976). When His E7 is replaced by another residue, ligand binding is strongly affected. Hb

† From the Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and the Laboratorium für Biochemie, Biochemie I, Eidgenössische Technische Hochschule, 8092 Zürich, Switzerland. *Received April 21, 1982.* The work was supported in part by Grant PHS GM 18051 from the Department of Health and Human Services, by Grant PCM 79-05072 from the National Science Foundation, and by a fellowship of the Deutsche Forschungsgemeinschaft to W.D.

* Correspondence should be addressed to this author at the Department of Physics, University of Illinois at Urbana-Champaign.

[1] Abbreviations: HbA, adult human hemoglobin; Mb, ferrous sperm whale myoglobin; β^A, normal β chains from adult human hemoglobin; β^{ZH}, abnormal β chains from hemoglobin Zürich; β^{HgBzO}, normal β chains with p-mercuribenzoate; α, normal α chain.

0006-2960/82/0421-4831$01.25/0

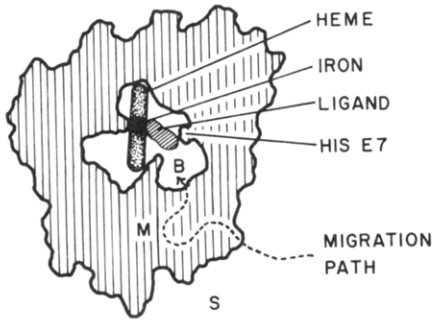

FIGURE 1: Schematic cross section through the active center of myoglobin with the ligand bound to the heme iron and touching the His E7 (state A). The heme pocket and the position of His E7 are indicated. After photodissociation to state B, the ligand moves through the matrix (M) into the solvent (S). The figure also shows schematically the pathway of rebinding.

Zürich, where His E7 in the β chain is replaced by Arg, has a much higher association rate for CO and an abnormally high O_2 affinity (Giacometti et al., 1977, 1980; Winterhalter et al., 1969). In contrast, opossum Hb, where His E7 is replaced by Gln, has a low O_2 affinity (Imai et al., 1980). Neutron and X-ray diffraction (Norvell et al., 1975; Heidner et al., 1976; Baldwin, 1980; Phillips, 1980; Phillips & Schoenborn, 1981) and EPR data (Ikeda-Saito et al., 1977) show that His E7 "touches" the bound ligand, as sketched in Figure 1. Interaction between the distal histidine and bound CO is seen in the CO stretching vibration of HbA (Caughey et al., 1969). The stretching vibration and the optical spectrum are pH dependent in MbCO and HbCO (Fuchsman & Appleby, 1979; Caughey et al., 1981). Optical and NMR measurements of Mb (Hayashi et al., 1976; LaMar et al., 1978) reveal a heme-linked protonation group with a pK of 5.7. Taken together, these features imply that His E7 can influence the ligand through steric hindrance (Collman et al., 1979) and through electromagnetic interactions (Satterlee et al., 1978): hydrogen bonding (Stryer et al., 1964), dipole–dipole force (Barlow et al., 1976), and donor–acceptor interaction (Caughey, 1970). In the present paper, we will show that flash photolysis experiments with Mb, β^A, and β^{ZH} performed as a function of pH and over a wide range of temperatures provide further evidence for the controlling role of His E7 and permit an estimate of the various contributions to the barrier between the pocket and the covalent binding site. Moreover, the experiments imply that the concept of pH still has validity at low temperatures.

Experimental Procedures

We have studied the binding of CO and O_2 to the heme proteins with flash photolysis. The solution containing the protein with bound ligand was placed in a cryostat and photodissociated with a flash from a pulsed laser. A 1-μs dye laser (540 nm; 150 mJ) or a frequency-doubled 50-ns Q-switched Nd^+ glass laser (530 nm; 300 mJ) was used. Both lasers had sufficient energy to completely dissociate CO samples, but some variance in signal (<5%) occurred for O_2 samples. Ligand rebinding after photodissociation was monitored from 10 ns to 100 s by observing the absorbance change at 436 nm with a beam produced by a tungsten lamp and interference filters. The photomultiplier signal was digitized with a Tektronix 7912 transient digitizer from 10 ns to 5 μs and with a logarithmic time-base digitizer from 2 μs to 100 s (Austin et al., 1975, 1976).

Sperm whale myoglobin type II was purchased from Sigma Chemical Co. (St. Louis, MO). The protein was dissolved in water, centrifuged, and filtered through a Millipore filter with

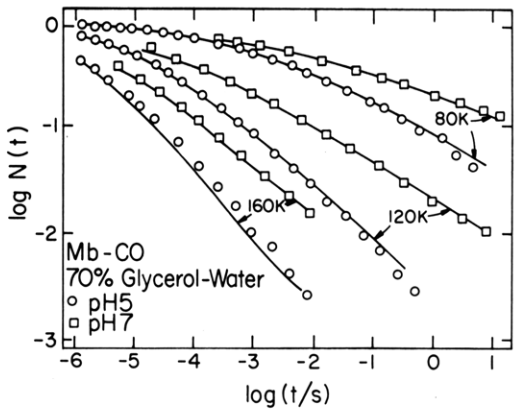

FIGURE 2: Recombination of CO with Mb at pH 7 and 5 over the temperature range 80–160 K. The solid lines are fits based on eq 1 by using the enthalpy distributions given in Figure 5. Observation wavelength, 436 nm; solvent, 70% glycerol–buffer mixture (by weight).

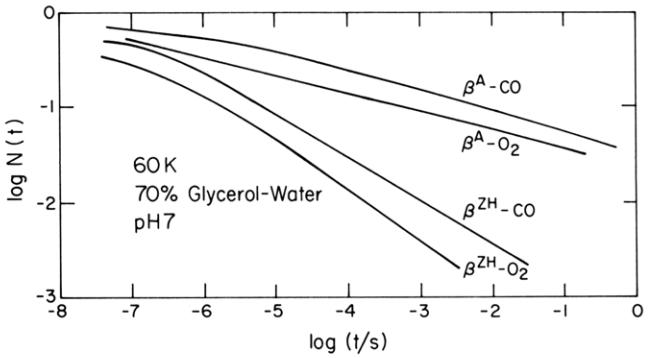

FIGURE 3: Comparison of the rebinding of CO and O_2 to β^A and β^{ZH} at pH 7 and 60 K. Observation wavelength, 436 nm; solvent, 70% glycerol–buffer (by weight).

a pore size of 0.4 μm (Millipore SA-67, Molsheim, France). The stock solution was diluted with 70% (by weight) glycerol in 0.1 M phosphate buffer or, alternatively, with the phosphate alone if measurements were performed only at 300 K. MbCO was obtained by adding to the diluted solution a 5-fold excess of dithionite, under 1 atm of CO. The oxy derivative was prepared by reducing a solution under an argon atmosphere with a 2-fold excess of dithionite and then equilibrating with air. The pH of aqueous samples and the apparent pH of the glycerol–water mixtures were determined by a Horizon 5996 pH meter at 298 K.

Hemoglobin samples were prepared from freshly drawn blood according to the procedures of Geraci et al. (1969). The α and β subunits were dissociated with p-mercuribenzoate (HgBzO) (Bucci & Fronticelli, 1965) by the procedure of Yip et al. (1972) and separated by column chromatography on DEAE-cellulose and CM-cellulose. The sulfhydryl groups were regenerated and tested by the method of Geraci et al. (1969). β^A–CO samples were obtained by gas exchange from β^A–O_2. β^{ZH} chains were prepared in the CO form as described elsewhere (Giacometti et al., 1980). Immediately after the regeneration of the SH groups, the protein was concentrated to ca. 19 mM, by using an ultrafiltration chamber, under 1 atm of CO. The material was then frozen and kept in liquid nitrogen until immediately before use. β^{ZH}–O_2 samples were obtained from β^{ZH}–CO by gas exchange under 1 atm of O_2. The stock solutions were diluted with 0.1 M phosphate buffer or a glycerol–buffer mixture. The CO samples were equilibrated under 1 atm of CO, and the residual O_2 was removed with dithionite.

Typical experimental data are given in Figures 2–4. In each, log $N(t)$ is plotted vs. log t. $N(t)$ is the fraction of protein

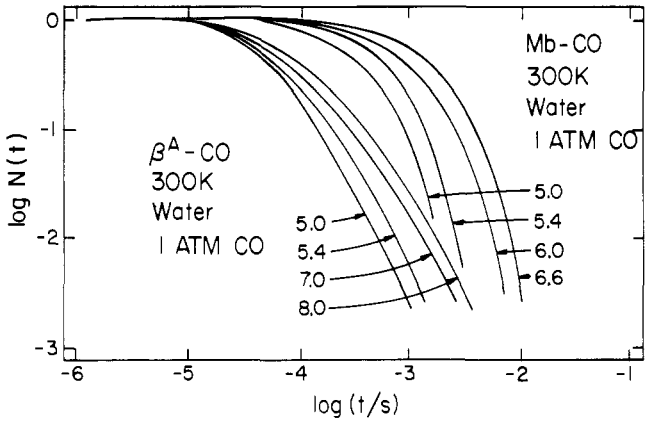

FIGURE 4: Recombination kinetics observed at 436 nm for CO with Mb and with β^A as a function of pH at 300 K. The water solvent was equilibrated with 1 atm of CO at 300 K.

molecules that have not rebound a ligand at time t after the laser flash. With only two optically different states involved, $N(t)$ is related to the observed absorbance change, $\Delta A(t)$, by $N(t) = \Delta A(t)/[(\epsilon' - \epsilon)\phi l]$ where ϕ is the protein concentration, l is the path length, and ϵ and ϵ' are the extinction coefficients for the bound and the free state, respectively.

Results and Evaluation

In earlier experiments (Austin et al., 1975; Beece et al., 1980), we have shown that the mechanism of ligand binding at physiological temperatures can only be unraveled if the binding at low temperatures is understood. The reason for this dependence is evident from Figure 1: Below about 180 K, the transition B → A from the pocket to the covalent binding site can be investigated and characterized. At physiological temperatures, binding occurs via a series of steps terminating with the transition B → A. For separation of the various steps, the extrapolated properties of the step B → A are required. We consequently begin the data evaluation here with the low-temperature phenomena.

Ligand Binding at Low Temperature (60–180 K). At temperatures below about 180 K, the ligand remains in the pocket after photodissociation and a single rebinding process, I, corresponding to the transition B → A, is observed. Figures 2 and 3 demonstrate that $N(t)$ for process I is not exponential in time but approximately follows a power law. The nonexponential time dependence can be understood in terms of conformational substates (Austin et al., 1975; Frauenfelder et al., 1979). A given primary sequence of amino acids does not lead to a unique tertiary structure of the protein but to a large number of related, but in detail different, structures. Below about 180 K, each protein molecule remains in a particular substate, with a corresponding barrier height H_{BA} for the transition B → A. If we denote with $g(H_{BA})dH_{BA}$ the probability of finding a barrier with a height between H_{BA} and $H_{BA} + dH_{BA}$, $N(t)$ is given by

$$N(t) = \int_0^\infty dH_{BA}\, g(H_{BA}) \exp(-k_{BA}t) \qquad (1)$$

where k_{BA} is the rate coefficient for the transition B → A over a barrier with height H_{BA}. In Mb, tunneling becomes important for k_{BA} below about 60 K (Alberding et al., 1976a; Alben et al., 1980). Above about 60 K, k_{BA} is given by an Arrhenius equation, $k_{BA} = A_{BA} \exp[-H_{BA}/(RT)]$. Above about 180 K, some ligands move into the globin matrix before rebinding, and processes other than the direct rebinding B → A contribute to $N(t)$. In the temperature interval from 60 to 180 K, eq 1 and the Arrhenius relation are both valid, and the

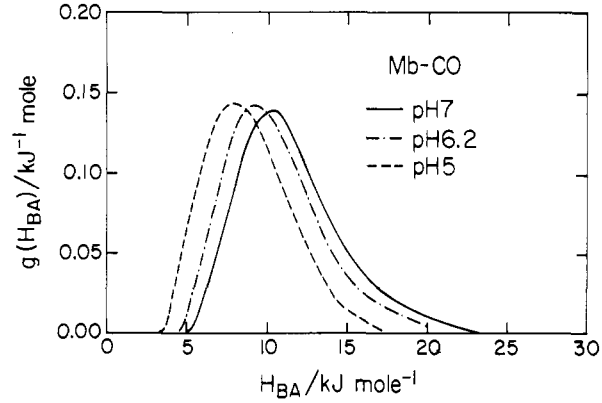

FIGURE 5: Activation enthalpy spectra of the innermost barrier (B → A) of MbCO at different pH values. $g(H_{BA})dH_{BA}$ denotes the probability of finding a molecule with an activation enthalpy between H_{BA} and $H_{BA} + dH_{BA}$. See the text for a discussion of the errors in the distribution.

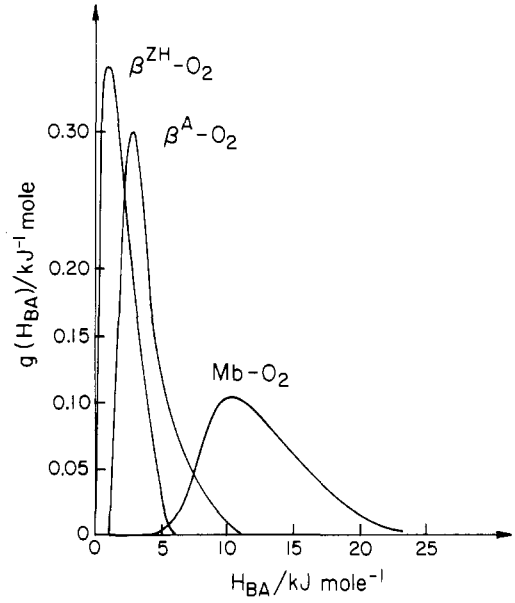

FIGURE 6: Activation enthalpy spectra of the innermost barrier (B → A) for MbO_2, β^AO_2 and $\beta^{ZH}O_2$.

probability distribution $g(H_{BA})$ and the preexponential A_{BA} can be computed from the measured function $N(t)$ by numerically inverting the integral equation (eq 1).

The inversion of eq 1 is similar to an inverse Laplace transform and suffers the instability inherent in such calculations (Bellman et al., 1966). We control the instability (Provencher & Dovi, 1979) by using a combination of constrained least-squares fitting and smoothing. The smoothing damps out oscillations which are mainly due to statistical fluctuations in the data. Probability distributions, $g(H_{BA})$, that have negative parts or exhibit excessive structure are excluded. Correlations in the parameters of the distribution function provide a limit on the validity of the results. We believe that we determine A_{BA} to within a factor of 2 and the peak of the distribution to within 1 kJ/mol. For internal consistency, we assign equal A_{BA} to systems where the evaluation gives closely similar values.

Probability densities, $g(H_{BA})$, determined by the procedure described above, are shown in Figures 5 and 6. The peak values of the enthalpy distribution and the preexponentials, A_{BA}, are listed in Table I.

Ligand Binding at High Temperatures (~300 K). Between about 180 and 250 K, many ligands still rebind directly, but others move into the protein matrix and either rebind from

Table I: Parameters Characterizing the Barrier between the Heme Pocket (B) and the Iron Binding Site (A) [a]

protein	ligand	pH	log A_{BA} (s^{-1})	H_{BA}^{peak} (kJ mol^{-1})	remarks
Mb	CO	5	8.7	7.8	
Mb	CO	7	8.7	10.2	
Mb	CO	9	8.7	10.9	
α	CO	7	8.6	4.5	
β^A	CO	5	8.7	3.1	
β^A	CO	7	8.7	4.4	
β^{ZH}	CO	7	8.6	2.0	b
β^{HgBzO}	CO	7	9.2	8.5	
protoheme	CO		11	3.6	
Mb	O_2	7	9.3	10	b
β^A	O_2	7	9.3	3.2	b
β^{ZH}	O_2	7	9.3	1.0	b

[a] The values given here are in some cases different from the ones given in earlier publications (Austin et al., 1975; Alberding et al., 1976a, 1978). The changes are due to reevaluation with improved techniques. Moreover, where the preexponentials A_{BA} turned out to be equal within errors, they were set to be equal. The values for protoheme have been obtained from measurements over considerably wider ranges in time and temperature. [b] No pH dependence.

Table II: Parameters Characterizing Rebinding after Photodissociation [a]

protein	ligand	pH	T (K)	N^{out}	λ_{IV} (10^3 s^{-1}) ± 10%	$\bar{k}_{BA}$ (10^8 s^{-1}) ± 50%	$P_B(c,T)$ (10^{-4}) ± 50%
Mb	CO	5	300	1.0	2.9	0.19	1.5
Mb	CO	9	300	1.0	0.72	0.066	1.1
Mb	O_2	7	300	0.8	28	0.41	8.0
α	CO	7	270	0.28	1.7	0.5	1.2
β^A	CO	5	300	1.0	17.0	1.1	1.5
β^A	CO	7	300	1.0	8.8	0.66	1.3
β^A	O_2	7	300	0.5	41.0	5.4	1.5
β^{HgBzO}	CO	7	280	0.55	1.1	0.16	1.2
β^{ZH}	CO	7	300	0.8	19.0	1.6	1.5
β^{ZH}	O_2	7	300	0.3	45.0	12.0	1.2
protoheme	CO		300	0.82	8.0	260	0.004

[a] All symbols are defined in the text. Data are from the present work and from our earlier papers (Austin et al., 1975; Alberding et al., 1976a, 1978). The data refer to a ligand pressure of 1 atm at 300 K.

there (processes II and III) or move into the solvent (Austin et al., 1975). Above about 270 K, nearly all photodissociated ligand molecules migrate into the solvent. All ligand molecules in the solvent compete for the vacant binding site, and we call the resulting rebinding process "IV". As we will show, the binding process at physiological temperatures can be understood if only processes I and IV are considered. The "matrix processes" II and III, which provide information about the protein matrix, will therefore not be treated here.

Figure 4 gives process IV as a function of pH for the binding of CO to Mb and to β^A at 300 K. The curves in Figure 4 can be approximated by

$$N_{IV}(t) = N^{out} \exp(-\lambda_{IV} t) \quad (2)$$

N^{out} is the fraction of photodissociated protein molecules whose ligands have migrated into the solvent; λ_{IV} is a pseudo-first-order rate coefficient that is proportional to the ligand concentration. The exponential time dependence at high temperatures is the result of conformational relaxation: The protein molecules are not frozen into particular substates but move rapidly from one substate to another (Austin et al., 1975). The rate coefficient k_{BA} is no longer given by a distribution but characterized by a single average value, $\bar{k}_{BA}$. We calculate $\bar{k}_{BA}(T)$ at temperature T from the probability density $g(H_{BA})$:

$$\bar{k}_{BA}(T) = A_{BA} \int_0^\infty dH_{BA}\, g(H_{BA}) \exp[-H_{BA}/(RT)] \quad (3)$$

For the analysis of process IV, we assume the sequential model based on Figure 1 (Austin et al., 1975; Beece et al., 1980), with $\bar{k}_{AB} \ll \bar{k}_{BA}$:

$$S \underset{k_{MS}}{\overset{k_{SM}}{\rightleftharpoons}} M \underset{k_{BM}}{\overset{k_{MB}}{\rightleftharpoons}} B \underset{\bar{k}_{AB}}{\overset{\bar{k}_{BA}}{\rightleftharpoons}} A \quad (4)$$

S, M, B, and A are protein states with the ligand in the solvent, in the protein matrix, in the heme pocket, and covalently bound to the iron, respectively. In the following, we will introduce a relation connecting the rate coefficient λ_{IV} (association rate) to $\bar{k}_{BA}$ and properties of the heme pocket. Since this relation makes the treatment of the ligand binding at physiological temperatures more transparent than our earlier approach, we justify its introduction in detail.

Consider first a hypothetical protein in which the transition B → A is blocked ($\bar{k}_{BA} = 0$) but all other properties are unchanged. An *equilibrium* will then be established between the ligands in the solvent (S) and in the heme pocket (B). To characterize the equilibrium at ligand concentration c and temperature T, we introduce a *pocket occupation factor*, $P_B(c,T)$. $P_B(c,T)$ is the equilibrium coefficient between the heme pocket and solvent in the limit $k_{BA} \to 0$. In terms of the model eq 4, $P_B(c,T)$ is

$$P_B(c,T) = \frac{k_{SM} k_{MB}}{k_{MS} k_{BM}} \quad (5)$$

At concentrations c that are easily obtainable, $P_B(c,T)$ is much smaller than unity and can therefore also be looked at as the probability of finding a ligand molecule in the heme pocket of one deoxy protein molecule if $\bar{k}_{BA} = 0$.

We next consider the case, realized for MbCO at 300 K, where the direct rebinding B → A after photodissociation is very small. The situation occurs if $N^{out} \approx 1$; after photodissociation, essentially all ligand molecules migrate into the solvent. A *preequilibrium* is established between the ligands in the solvent and the heme pocket.[2] The fraction of protein molecules in state B, N_B, is essentially identical with the equilibrium value, $P_B(c,T)$. The binding rate from the solvent, λ_{IV}, is the pocket occupation factor times the rate for the final step:

$$\lambda_{IV}(c,T) = \bar{k}_{BA} P_B(c,T)$$

If N^{out} is measurably smaller than 1, a fraction $(1 - N^{out})$ of proteins will rebind the ligand before it can escape into the solvent. The fraction N^{out} of protein molecules where the ligand has migrated into the solvent approaches a *steady-state* situation in which the ligands return slowly from the solvent and ultimately rebind. The association rate then becomes

$$\lambda_{IV}(c,T) = \bar{k}_{BA}(T)\, P_B(c,T)\, N^{out}(T) \quad (6)$$

Equation 6 can be derived with the approach described in Austin et al. (1975), as discussed in the Appendix. Equation 6 is a very good approximation if binding from the solvent is much slower than all internal rebinding processes and if the photodissociating pulse is so short that each protein molecule

[2] We already discussed the fact that a preequilibrium is established in the limit $N^{out} \to 1$ in section 6.2 of Austin et al. (1975); the same point has also been treated by Reisberg & Olson (1980).

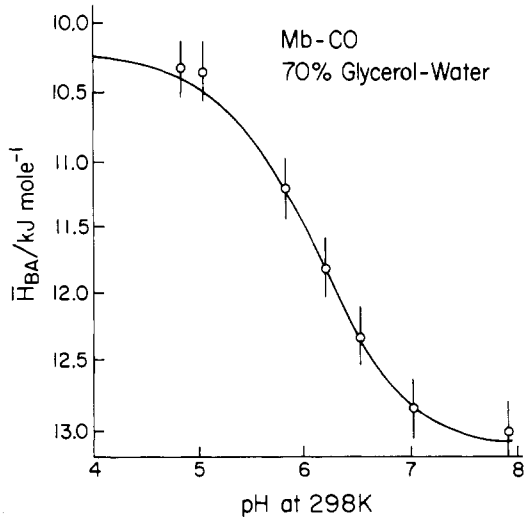

FIGURE 7: Average activation enthalpy of MbCO for the B → A transition as a function of pH. The solid line is the best fit to a Henderson–Hasselbalch equation with pK = 6.1.

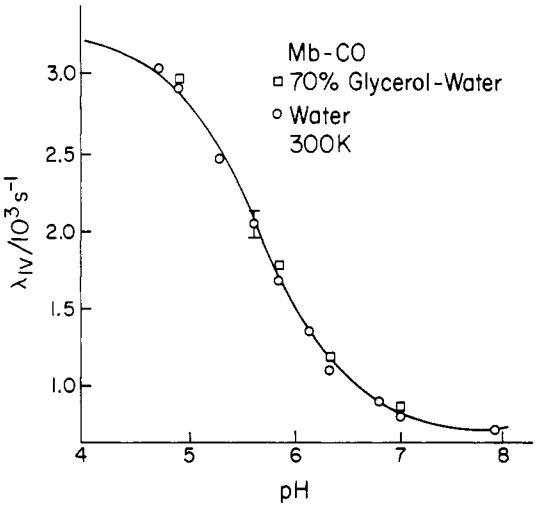

FIGURE 8: Rebinding rate from the solvent (λ_{IV}) of MbCO as a function of pH in water (○) and glycerol–water (□) at 300 K under 1 atm of CO. The solid line is the best fit with pK = 5.7.

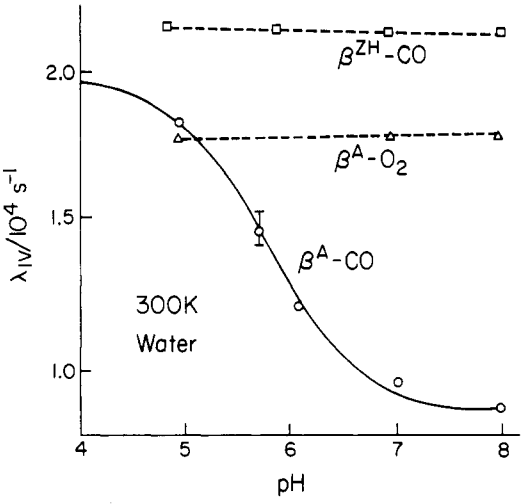

FIGURE 9: Rebinding rates from the solvent (λ_{IV}) of β^A–CO, β^A–O_2, and β^{ZH}–CO as a function of pH at 300 K in water. The CO samples were under 1 atm of CO, and the O_2 sample was equilibrated with air. The solid line is the best fit with pK = 5.8. The dashed lines are drawn to guide the eye.

is photodissociated at most once. Both conditions are satisfied in our experiments, and we have evaluated the experimental data with eq 6; the results are given in Table II.

The advantage of using eq 6 and introducing the pocket occupation factor $P_B(c,T)$ is 2-fold: $P_B(c,T)$ is an equilibrium property and permits comparison of systems that otherwise are much more difficult to collate. As we will show below, $P_B(c,T)$ can be interpreted and yields an intuitive understanding of some pocket properties.

pH Dependence of Binding at Low Temperature. The binding of CO to Mb at temperatures below 160 K is shown for pH 5 and pH 7 in Figure 2. Rebinding is considerably faster at the lower pH value. The rebinding curves can be described by activation enthalpy spectra, $g(H_{BA})$; the spectra for pH 5, 6.2, and 7 are given in Figure 5. The effect of the change in pH is a shift in the peak energy of the distribution by 3 kJ/mol. Shape and preexponential A_{BA} remain essentially constant. To discuss the shift, we assume that binding is influenced by an ionizable residue with two charge states, neutral and protonated. At temperatures below 200 K, the exchange between neutral and protonated states is most likely very slow; a given protein will therefore remain in one state during rebinding, and $N(t)$ can be calculated from eq 1 with $g(H_{BA})$ given by

$$g(H_{BA}) = Yg^+(H_{BA}) + (1 - Y)g^0(H_{BA}) \quad (7)$$

where $g^+(H_{BA})$ is the distribution for the protonated state (pH <5) and $g^0(H_{BA})$ is the one for the neutral state (pH >7). Y is the fraction of protein molecules with the residue protonated. Multiplication of eq 7 by H_{BA} and integrating over H_{BA} yield the average activation enthalpy, $\bar{H}_{BA}$(pH):

$$\bar{H}_{BA}(\text{pH}) = Y\overline{H_{BA}{}^+} + (1 - Y)\overline{H_{BA}{}^0} \quad (8)$$

$\overline{H_{BA}{}^+}$ and $\overline{H_{BA}{}^0}$ are the average enthalpies of the protonated and neutral states, respectively. $\bar{H}_{BA}$(pH) is plotted in Figure 7 as a function of pH. The experimental data can be fit by a Henderson–Hasselbalch equation:

$$pK_{app} - \text{pH} = \log [Y/(1 - Y)] \quad (9)$$

The fit, shown as a solid line in Figure 7, yields pK_{app} = 6.1 ± 0.2, $\overline{H_{BA}{}^+}$ = 10.2 ± 0.5 kJ/mol, and $\overline{H_{BA}{}^0}$ = 13.1 ± 0.5 kJ/mol. This result suggests that the pH dependence of the binding of CO to Mb is due to a single protonation site with an apparent pK of 6.1. CO binding to β^A is also faster at low pH; the speedup can be explained by a shift of the activation enthalpy spectrum by 1.5 kJ/mol in going from pH 7 to pH 5. In contrast, the binding rates of O_2 to Mb and of CO to β^{ZH} are independent of pH.

pH Dependence of the Binding at 300 K. Figure 4 shows that the recombination of CO to Mb is exponential at 300 K. Thus, conformational relaxation and the interchange between protonated and neutral states are rapid compared to binding. The recombination of CO to β^A is only approximately exponential, and it is possible that either (or both) of the relaxation processes is incomplete. We will, however, neglect the slight deviation from exponentiality here and use eq 2 for the determination of λ_{IV} and N^{out}.

The coefficients λ_{IV} for MbCO and β^A–CO in water and in a glycerol–water solvent fit a Henderson–Hasselbalch equation, eq 9, for a single protonation site with pKs of 5.7 ± 0.1 and 5.8 ± 0.2, respectively (Figures 8 and 9). In contrast, λ_{IV} values for the binding of CO to β^{ZH} and of O_2 to Mb, β^A, and β^{ZH} are independent of pH. Thus, low-temperature and high-temperature results agree qualitatively. Furthermore, the variation of $\bar{k}_{BA}$ with pH, observed at low temperature, accounts for the change of λ_{IV} with pH at 300 K quantitatively. We calculate $\bar{k}_{BA}$ with eq 3 by assuming

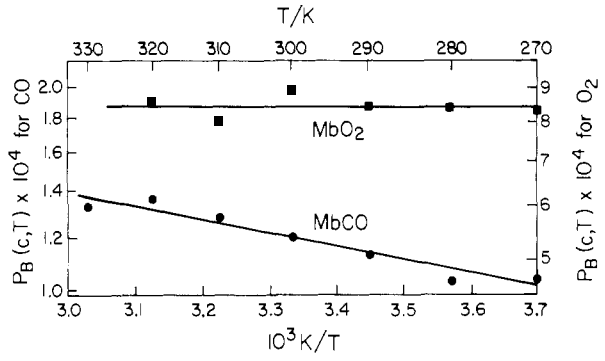

FIGURE 10: Pocket occupation factor, $P_B(c,T)$, at constant ligand concentration (1 atm of the ligand equilibrated in water at 273.1 K) for MbCO and MbO_2 as a function of inverse temperature. See the text for a discussion of the errors associated with the value of $P_B(c,T)$.

that the enthalpy distribution is temperature independent. Inserting the measured values of λ_{IV} and N^{out} and the calculated $\bar{k}_{BA}$ into eq 6 gives the values of $P_B(c, 300\ K)$ listed in Table II. $P_B(c,T)$ is independent of pH; the entire pH dependence of λ_{IV} comes from $\bar{k}_{BA}$. The titratable group is most likely close to the heme.

The pK value of 5.7 for λ_{IV} at 300 K, together with the observation that ligand binding rates of β^{ZH} display no pH effect, suggests that the distal histidine, His E7, is the protonation site. The shift of 0.4 pK unit between high- and low-temperature data is most likely due to the temperature dependence of pH and pK. Both the pH of a glycerol–water mixture (phosphate buffer) and the pK of histidine in glycerol–water increase with decreasing temperature by a similar amount (Douzou, 1977). Consequently, a measurement of the state of protonation at low temperature as a function of pH at room temperature may lead to a small shift in pK_{app} relative to measurements at room temperature. The shift in pK_{app} can also be caused by a dependence of pK on conformation; the protein may remain in the carboxy state after photodissociation at low T while at high T relaxation may be fast. Optical measurements at 300 K (Hayashi et al., 1976) indicate, however, only a small shift in pK between the deoxy- and carboxymyoglobin.

Pocket Occupation Factor, $P_B(c,T)$. Through eq 5, we have defined $P_B(c,T)$ as an equilibrium coefficient between solvent and heme pocket. If the solvent is in equilibrium with the ligand gas, we can redefine $P_B(c,T)$ as the coefficient of equilibrium between the ligand molecules in the gas phase and in state B; c is the ligand concentration in the gas phase (Ackerman & Berger, 1963; Collman et al., 1979; McKinnie & Olson, 1981). For solvents that do not affect the free energy of state B, $P_B(c,T)$ can be defined and studied independent of solvent properties. For convenience, we perform the experiments at constant ligand gas pressure and get the pocket occupation factor at constant concentration from

$$P_B(c = \text{const}, T) = (T/T_0)\, P_B(p = \text{const}, T) \quad (10)$$

where $T_0 = 273.1$ K is the standard temperature. For the further evaluation of $P_B(c,T)$, we write

$$P_B(c = \text{const}, T) = P_0(c) \exp[-H_B/(RT)] \quad (11)$$

We have measured $\lambda_{IV}(p,T)$ and $N^{out}(T)$ for the binding of CO and O_2 to Mb in aqueous solutions at constant gas pressure, 1 atm $\sim 10^5$ Pa. With $\bar{k}_{BA}$ calculated from eq 3, $P_B(c = \text{const}, T)$ was determined with eq 6 and 10; log $P_B(c = \text{const}, T)$ is plotted vs. $1/T$ in Figure 10. A fit of eq 11 to the data yields the following values for Mb–CO (eq 12a) and Mb–O_2 (eq 12b):

$$H_B = +2 \pm 1\ \text{kJ/mol} \qquad P_0 = (5 \pm 3) \times 10^{-4} \quad (12a)$$

$$H_B = 0 \pm 1\ \text{kJ/mol} \qquad P_0 = (8 \pm 4) \times 10^{-4} \quad (12b)$$

Discussion and Conclusions

The data in the previous sections provide insight into the mechanisms of binding of CO and O_2 to heme proteins at physiological temperatures. At the same time, the concepts introduced permit a concise and simplified discussion of the binding process. Since the concepts can be applied to other systems, we summarize here first the essential relations. The pseudo-first-order rate coefficient for binding (association coefficient), defined by eq 2, is given by eq 6, where $\bar{k}_{BA}$ is given by eq 3, and c is the concentration of the ligand in the gas phase. The dissociation rate coefficient, λ_{dis}, is mainly determined by the strength of the covalent bond at the iron. In terms of the model eq 4, λ_{dis} is given by

$$\lambda_{dis} = \bar{k}_{AB} N^{out} \quad (13)$$

so that the equilibrium coefficient K_{eq} becomes

$$K_{eq} = \lambda_{IV}/\lambda_{dis} = \frac{\bar{k}_{BA}}{\bar{k}_{AB}} P_B(c,T) \quad (14)$$

Equation 6 tells us that the details of the ligand migration through the protein matrix are not important for binding at physiological temperatures. The matrix processes only determine how fast preequilibrium is reached, but the influence of pocket and matrix on λ_{IV} can be expressed by the one factor $P_B(c,T)$. Moreover, solubility of the ligand in the solvent is not relevant; the data can be discussed and understood in terms of the concentration c in the gas phase. Finally, we observe that in the systems studied here N^{out} is close to 1 at physiological temperatures; the dissociation rate then is determined by $\bar{k}_{AB}$ and the association rate by $\bar{k}_{BA}$ and $P_B(c,T)$. Table II shows that $P_B(c,T)$ is nearly the same for all protein–ligand systems studied except for MbO_2 (and protoheme). In contrast, $\bar{k}_{BA}$ ranges from $6.6 \times 10^6\ s^{-1}$ to $1.2 \times 10^9\ s^{-1}$, about a factor of 200. The barrier between B and A consequently dominates the control of the association rate. We pointed out this control function of the heme barrier in an earlier paper (Alberding et al., 1978), and it has also been noted by Morris & Gibson (1980). In the following, we discuss the control factors $\bar{k}_{BA}$ and $P_B(c,T)$ in more detail.

Control through the Barrier between the Pocket and Iron. The properties of the barrier between states B and A are given in Figures 5 and 6 and Table I. The parameters in Table I show some remarkable regularities: The preexponentials A_{BA} are, within errors, the same for all heme proteins studied and depend only on the ligand. For the binding of CO, $A_{BA} = 5 \times 10^8\ s^{-1}$; for the binding of O_2, $A_{BA} = 2 \times 10^9\ s^{-1}$. The difference between CO and O_2 is in the expected direction: CO can only bind with one end and O_2 with both. Symmetry thus predicts, other things being equal, $A_{BA}(O_2) = 2A_{BA}(CO)$.

Since the preexponential factor A_{BA} is essentially the same in all systems, the protein adjusts the rate coefficient k_{BA} by changing the barrier height H_{BA} between B and A. Table I implies that the adjustment is done not through just one property but by using the protein structure in different ways; estimates of the various contributions, obtained in the next paragraph by comparing various proteins, are given in Table III.

We first consider the contributions to the control of CO binding. With no distal residue present, the barrier can be caused by motion of the iron into the heme plane, by flattening of the heme, and by repulsion of the CO by the four nitrogen

Table III: Approximate Contributions to the Average Activation Enthalpy in the Transition B → A

source of contribution	contribution to H_{BA} (kJ mol^{-1}) CO	O_2
heme group	2	1–2
steric control at distal His	2	2
charge on distal His	2–4	0
protein structure	3–6	3–6

atoms surrounding the iron. The small value of H_{BA} for β^{ZH} indicates that the total contribution of these factors is of the order of 2 kJ/mol. Control at the distal histidine is apparent from a comparison of β^{A}, where His E7 is present, with β^{ZH}, where the distal residue is removed from the pocket; His E7 adds about 2 kJ/mol to the barrier height. The pH dependence implies that the charge on His E7 can change the activation barrier by 2–4 kJ/mol. The globin may affect the barrier by restricting the motion of the heme group or by "pulling" on the imidazole–iron bond (Sharma et al., 1978; Traylor, 1981). Comparison of β^{A} with β^{HgBzO} and with Mb indicates that interaction of the globin with the iron and the porphyrin through the covalent iron–histidine bond and through hydrogen bonds can contribute as much as 6 kJ/mol to the barrier.

In the discussion in the previous paragraph, we have implicitly assumed that β^{A}, β^{ZH}, and β^{HgBzO} have essentially the same tertiary structure and that the dominant modification occurs near the heme group and affects only the last binding step. This assumption is supported by the X-ray diffraction studies of Phillips et al. (1981), who show that the main changes between β^{ZH} and β^{A} occur near the heme. Strong indirect evidence also comes from the data in Table II: While λ_{IV}(300 K) varies from 1.1×10^3 s^{-1} for β^{HgBzO} to 19×10^3 s^{-1} for β^{ZH}, $P_B(c,T)$ remains unchanged within errors; the entire variation in λ_{IV} is due to the binding process at the heme.

Comparison of the data in Tables I and II for CO and O_2 shows that H_{BA}^{peak} is generally smaller for O_2 than for CO, and A_{BA} is a factor of 2–4 larger. Both properties together are responsible for the faster binding of O_2. Even though O_2 prefers a "bent" geometry and state A should not be affected by the distal His, comparison of β^{A} and β^{ZH} implies a steric hindrance for the transition B → A also for O_2. O_2 binding is not pH dependent. A sizable contribution comes, as in the case of CO binding, from the globin.

The values in Table III acquire more meaning by noting that the effect of a change in barrier height by an amount ΔH at temperature T results in a change of the rate by a factor $\exp[\Delta H/(RT)]$. At 300 K, ΔH = 1.7 kJ/mol produces a rate change by a factor of 2.

Possible Mechanism for the pH Effect. Our data display the remarkable fact that CO binding to Mb and β^{A} depends on pH while O_2 binding does not. Moreover, the data prove that pH affects only the innermost barrier. A conformational change induced by protonation should alter the kinetics of both ligands. In order to explain the observations, we postulate a charge–dipole interaction between the CO dipole and the positive charge of the imidazole His E7. O_2 has no dipole moment and is not affected. The energy (W) of a dipole (**p**) in the field (**E**) of a charge (e) is given by

$$W = -\mathbf{p}\cdot\mathbf{E} = -|\mathbf{p}|\frac{e}{\epsilon r^2}\cos\theta \qquad (15)$$

ϵ is the medium dielectric coefficient, r the distance between the dipole and the point charge, and θ the angle between the field and the dipole. To obtain an upper limit for r, we choose extreme values; $\cos\theta = \epsilon = 1$. For a 3.5 kJ/mol shift in energy of a CO dipole (0.112 D) in the field of a proton, a maximum distance of 0.3 nm is required. This estimate shows that the observed shift in activation energy is consistent with a charge–dipole interaction. More generally, the change in activation enthalpy (ΔH) is given by

$$\Delta H = E_m \langle\cos\theta\rangle \qquad E_m = -|\mathbf{p}|[e/(\epsilon r^2)] \qquad (16)$$

where $\langle\cos\theta\rangle$ denotes a suitable average over the angle θ between the field and the dipole. The magnitude and the temperature dependence of ΔH depend on the allowed values of $\langle\cos\theta\rangle$. Assume first that CO can assume any direction; $\langle\cos\theta\rangle$ then is given by

$$\langle\cos\theta\rangle = \coth\alpha - \frac{1}{\alpha} \qquad \alpha = E_m/(RT) \qquad (17)$$

Equations 16 and 17 give the temperature dependence of the shift ΔH if θ can assume any value. Numerically, these relations show that the 3.0 kJ/mol shift seen with MbCO at low temperatures should decrease by almost a factor of 2 at 300 K. In reality, we observe ΔH to be constant to within 0.5 kJ/mol. To explain the absence of a large temperature dependence of ΔH, we postulate that the angle θ is restricted by steric constraints in the state where the interaction between the CO dipole moment and the charged residue occurs. Such an interaction can take place either at the bottom of state B or at the top of the barrier between B and A. An interaction in state B would not affect λ_{IV} because an increase in k_{BA} will be balanced by a decrease in $P_B(c,T)$. More likely is the alignment in the transition state B → A, possibly between His E7 and Val E11 (Case & Karplus, 1979) or His E7 and Phe CD1.

A heme-linked protonation group with a pK_{app} near 5.7 has been found with a number of techniques. Optical studies of the heme absorption yielded the following pK values: FeMbCO, 5.7 (Hayashi et al., 1976); $FeMbO_2$, 5.8 (Fuchsman & Appleby, 1979); $CoMbO_2$, 5.6 (Ikeda-Saito et al., 1977). The chemical shifts of the heme methyl protons and the NH proton of His F8 of deoxymyoglobin are modulated by a single titratable group with a pK of 5.7 (LaMar et al., 1978). The authors suggest His FG2 alternatively to His E7 as a possible source of the effect. The ESR spectrum of $CoMbO_2$ has a pH dependence with a similar pK which is not present in Glycera hemoglobin where the distal histidine and His FG2 are replaced by nontitrating groups (Ikeda-Saito et al., 1977). The oxygen affinity of $CoMbO_2$ decreases with decreasing pH parallel to the changes of the optical and ESR spectra (Ikeda-Saito et al., 1977). The authors explain this effect by protonation of the δ nitrogen of His E7 which is in contact with the solvent. Protonation of the δ nitrogen may influence the hydrogen bond between the ϵ nitrogen and the bound ligand. This hydrogen bond has been observed in MbO_2 by neutron diffraction (Phillips & Schoenborn, 1981). If we integrate all these observations, we conclude that His E7 is responsible for much of the steric hindrance between states B and A, that the charge influences CO, but not O_2, binding, that the charge–dipole interaction takes place in the transition state between B and A, and that protonation speeds up binding.

Pocket Occupation Factor, $P_B(c,T)$. The values in Table II show that $P_B(c,T)$ is within error the same for all ligand–protein systems investigated here except for MbO_2. For these storage and transport systems, $P_B(c,T)$ therefore does not seem to be used for control. In protoheme, where the binding site is at the heme–solvent interface, $P_B(c,T)$ is about 300 times smaller than in the heme proteins.

Table IV: Ligand Concentration (c), Pocket Occupation Factor [$P_B(c,T)$], and Pocket Volume (V_B)[a]

system	c (Å^{-3})	$P_B(c, 300\ K)$	P_0	V_B (Å^3)
gas (CO, O_2)	2.7×10^{-5}			
water-O_2	8×10^{-7}			
glycerol-water (75% v/v)-CO	3×10^{-7}			
Mb-CO		1.3×10^{-4}	5×10^{-4}	20
Mb-O_2		8×10^{-4}	8×10^{-4}	30

[a] c is the average number of ligand molecules in the gas phase at standard conditions, 1 atm and 273.15 K, in 1 Å^3.

We can gain some insight into the nature of state B by assuming that B is a pocket with volume (V_B) close to the heme iron and that the concentration of the ligand in the pocket is the same as in a gas. We obtain an approximate value for the pocket volume from

$$P_0 = V_B c \tag{18}$$

where c is the average number of ligand molecules in 1 Å^3 at standard conditions, 1 atm at 273.1 K. Table IV gives c for the ligand gas and for the ligand dissolved in water and glycerol–water and also lists P_B(c, 300 K) and P_0, obtained from eq 11. Values for V_B, listed in Table IV, can be compared to the cavity sizes deduced from X-ray diffraction (Connolly et al., 1982). The X-ray data show a single heme pocket with volume of 135 Å^3 for β^A, but two smaller cavities of volumes 48 and 84 Å^3 for Mb and 55 and 38 Å^3 for α. The observed and the calculated values agree surprisingly well, and at least three factors can account for the fact that the values of V_B calculated from the binding experiments are smaller than those observed in X-ray diffraction: (1) Only a part of the cavity may permit access to the binding site A. (2) The ligand may not move freely in state B; hindrance decreases the number of rotational degrees of freedom. Low-temperature infrared data provide evidence that the photodissociated CO in Mb is rotationally constrained (Alben et al., 1982). (3) The ligand in the pocket induces a conformational change of the protein. Additional experiments are required to distinguish among these possibilities.

Acknowledgments

We thank M. C. Marden and G. Wagner for help and advice, J. S. Olson, K. S. Suslick, and M. B. Weissman for illuminating remarks, and I. D. Kuntz for essential information concerning Figure 1.

Appendix

Equation 6 is easily verified for the three-well case with $k_{BS} \ll k_{SB}$ (Alberding et al., 1976b):

$$A \xleftarrow{k_{BA}} B \underset{k_{BS}}{\overset{k_{SB}}{\rightleftharpoons}} S$$

where A, B, and S are protein states with the ligand covalently bound to the iron, in the heme pocket, and in the solvent, respectively. After photodissociation, all ligands are in state B. From B, a fraction N^{out} will go into the solvent, with

$$N^{out} = \frac{k_{BS}}{k_{BA} + k_{BS}}$$

The rate coefficient λ_{IV} for recombination from the solvent is given by

$$\lambda_{IV} = k_{BA}\frac{k_{SB}}{k_{BA} + k_{BS}}$$

Combining the two equations and defining $P_B(c,T)$ as an equilibrium coefficient, $P_B = k_{SB}/k_{BS}$, we get

$$\lambda_{IV} = k_{BA} N^{out} P_B(c,T) \tag{6}$$

Equation 6 can be more generally verified as follows: Denote with $t_g = 1/\lambda_g$ the time after photodissociation at which all geminate (internal) rebinding processes are complete and assume $\lambda_{IV} \ll \lambda_g$. Call $i \rightarrow k$ the effective (integrated) rate coefficient for the transition $i \rightarrow k$ during the time t_g and N_B and N_S the populations of states B and S, with $\lim_{(B\rightarrow A)=0} N_B = P_B(c,T)$ (i). At the time t_g, a fraction $1 - N^{out}$ of photodissociated molecules will have returned to A, a fraction N^{out} will be in S, and $(1 - N^{out})/N^{out} = (B \rightarrow A)/(B \rightarrow S)$ (ii). A steady state will then exist, governed by $N_B[(B \rightarrow S) + (B \rightarrow A)] = N_S(S \rightarrow B)$ (iii), where N_B refers only to the protein molecules that have not undergone the fast geminate rebinding. Equations i–iii give $N_B = P_B(c,T)N^{out}$, and eq 6 is verified.

References

Ackerman, E., & Berger, R. L. (1963) *Biophys. J. 3*, 493–505.

Alben, J. O., Beece, D., Bowne, S. F., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M. C., Moh, P. P., Reinisch, L., Reynolds, H. H., & Yue, K. T. (1980) *Phys. Rev. Lett. 44*, 1157–1160.

Alben, J. O., Beece, D., Bowne, S. F., Doster, W., Eisenstein, L., Frauenfelder, H., Good, D., McDonald, J. D., Marden, M. C., Moh, P. P., Reinisch, L., Reynolds, A. H., Shyamsunder, E., & Yue, K. T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3744–3748.

Alberding, N., Austin, R. H., Beeson, K. W., Chan, S. S., Eisenstein, L., Frauenfelder, H., & Nordlund, T. M. (1976a) *Science (Washington, D.C.) 192*, 1002–1004.

Alberding, N., Austin, R. H., Chan, S. S., Eisenstein, L., Frauenfelder, H., Gunsalus, I. C., & Nordlund, T. M. (1976b) *J. Chem. Phys. 65*, 4701–4711.

Alberding, N., Chan, S. S., Eisenstein, L., Frauenfelder, H., Good, D., Gunsalus, I. C., Nordlund, T. M., Perutz, M. F., Reynolds, A. H., & Sorensen, L. B. (1978) *Biochemistry 17*, 43–51.

Antonini, E., & Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands*, North-Holland Publishing Co., Amsterdam.

Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., & Gunsalus, I. C. (1975) *Biochemistry 14*, 5355–5373.

Austin, R. H., Beeson, K. W., Chan, S. S., Debrunner, P. G., Downing, R., Eisenstein, L., Frauenfelder, H., & Nordlund, T. M. (1976) *Rev. Sci. Instrum. 47*, 445–447.

Baldwin, J. M. (1980) *J. Mol. Biol. 136*, 103–128.

Barlow, C. H., Ohlsson, P. I., & Paul, K. G. (1976) *Biochemistry 15*, 2225–2229.

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.

Bellman, R., Kaluba, R., & Lackert, J. A. (1966) *Numerical Inversion of the Laplace Transform: Applications to Biology, Economics, Engineering, and Physics*, American Elsevier, New York.

Bucci, E., & Fronticelli, C. (1965) *J. Biol. Chem. 240*, 551–553.

Case, D. A., & Karplus, M. (1979) *J. Mol. Biol. 132*, 343–368.

Caughey, W. S. (1970) *Ann. N.Y. Acad. Sci. 174*, 148–153.

Caughey, W. S., Alben, J. O., McCoy, S., Boyer, S., Charache, S., & Hathaway, P. (1969) *Biochemistry 8*, 59–62.

Caughey, W. S., Shimada, H., Choc, M. G., & Tucker, M. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2903–2907.

Collman, J. P., Brauman, J., & Doxsee, K. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6035–6039.

Connolly, M. L., Kuntz, I. D., Ferrin, T. E., & Langridge, R. (1982) *J. Mol. Biol.* (in press).
Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure*, Vol. 5, National Biomedical Research Foundation, Silver Spring, MD.
Douzou, P. (1977) *Cryobiochemistry*, pp 109–115, Academic Press, London.
Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. (1979) *Nature (London) 280*, 558–563.
Fuchsman, W. H., & Appleby, C. A. (1979) *Biochemistry 18*, 1309–1321.
Geraci, G., Parkhurst, L., & Gibson, Q. H. (1969) *J. Biol. Chem. 244*, 4664–4667.
Giacometti, G. M., DiIorio, E. E., Antonini, E., Brunori, M., & Winterhalter, K. H. (1977) *Eur. J. Biochem. 75*, 267–273.
Giacometti, G. M., Brunori, M., Antonini, E., DiIorio, E. E., & Winterhalter, K. H. (1980) *J. Biol. Chem. 255*, 6160–6165.
Hayashi, Y., Yamada, H., & Yamazaki, I. (1976) *Biochim. Biophys. Acta 427*, 608–616.
Heidner, E. J., Ladner, R. C., & Perutz, M. F. (1976) *J. Mol. Biol. 104*, 707–722.
Ikeda-Saito, M., Iizuka, T., Yamamoto, H., Kayne, I. J., & Yonetani, T. (1977) *J. Biol. Chem. 14*, 4882–4887.
Imai, K., Ikeda-Saito, M., & Yonetani, T. (1980) *J. Mol. Biol. 144*, 551–565.
LaMar, G. N., Budd, D. L., Sick, H., & Gersonde, K. (1978) *Biochim. Biophys. Acta 537*, 270–283.
McKinnie, R. E., & Olson, J. S. (1981) *J. Biol. Chem. 256*, 8928–8932.
Morris, R. J., & Gibson, Q. H. (1980) *J. Biol. Chem. 255*, 8050–8053.
Norvell, J. C., Nunes, A. C., & Schoenborn, B. P. (1975) *Science (Washington, D.C.) 190*, 568–569.
Phillips, S. E. V. (1980) *J. Mol. Biol. 142*, 531.
Phillips, S. E. V., & Schoenborn, B. (1981) *Nature (London) 292*, 81–82.
Phillips, S. E. V., Hall, D., & Perutz, M. F. (1981) *J. Mol. Biol. 150*, 137–141.
Provencher, S. W., & Dovi, V. G. (1979) *J. Biochem. Biophys. Methods 1*, 313–318.
Reisberg, P. I., & Olson, J. S. (1980) *J. Biol. Chem. 255*, 4151–4158.
Satterlee, J. D., Teintze, M., & Richards, J. (1978) *Biochemistry 17*, 1456–1462.
Sharma, V. S., Geibel, J. F., & Ranney, H. M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3747–3750.
Stryer, C., Kendrew, J. C., & Watson, H. C. (1964) *J. Mol. Biol. 8*, 96–104.
Szabo, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2108–2111.
Traylor, T. G. (1981) *Acc. Chem. Res. 14*, 102–107.
Warshel, A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1789–1793.
Winterhalter, K. H., Anderson, N. M., Amiconi, G., Antonini, E., & Brunori, M. (1969) *Eur. J. Biochem. 11*, 435–440.
Yip, Y. K., Waks, M., & Beychok, S. (1972) *J. Biol. Chem. 247*, 7237–7244.

Biochemical and Crystallographic Data for Phaseolin, the Storage Protein from *Phaseolus vulgaris*†

Stephen Johnson, Guy Grayson, Lindsay Robinson, Rima Chahade, and Alexander McPherson*

ABSTRACT: We have isolated and biochemically characterized a major seed protein from *Phaseolus vulgaris* (common garden green bean) that appears to be identical with a form of the storage protein of that plant seed known as phaseolin. We have further shown that it appears very similar in most properties to the storage protein from *Canavalia ensiformis* (jack bean) which we have solved to 3.0-Å resolution using X-ray diffraction techniques. [McPherson, A., & Rich, A. (1973) *J. Biochem. (Tokyo) 74*, 155–160; McPherson, A., & Spencer, R. (1975) *Arch. Biochem. Biophys. 169*, 650–661; McPherson, A. (1980) *J. Biol. Chem. 255*, 10472–10480]. We have crystallized phaseolin and conducted a preliminary X-ray diffraction analysis on it as well. The data show the crystals to be of pseudo cubic space group *P*432, although the symmetry can be, strictly speaking, only *P*1. The unit cell has one trimeric molecule of 150 000 daltons in a unit cell of dimensions $a = b = c = 66.6$ and $\alpha = \beta = \gamma = 90°$. The crystals apparently possess some form of disorder that makes reconciliation of the unit cell contents with the observed crystallographic properties difficult, although they do diffract strongly to better than 2.8-Å resolution. No evidence of twinning has been observed.

As an extension of our X-ray structural studies on canavalin, from jack beans (*Canavalia ensiformis*), we undertook a search for similar crystalline proteins from other plant species. We have isolated such a protein from *Phaseolus vulgaris* and have obtained it in crystalline form. The protein has those properties, so far as we can determine, that identifies it as phaseolin, one of the major storage proteins of the seed. This protein has recently gained some additional attention by virtue of its being the first plant protein cloned into bacteria (Sun et al., 1980). The protein demonstrates a number of interesting biochemical and physical properties which indicate that it is closely similar in a structural sense to the protein canavalin which is a major storage protein in jack beans. We have previously studied canavalin by X-ray diffraction and other biochemical techniques and have recently reported the solution

† From the Department of Biochemistry, University of California, Riverside, California 92521. *Received October 30, 1981; revised manuscript received June 15, 1982.* This research was supported by Grant GM 21398 from the National Institutes of Health.

0006-2960/82/0421-4839$01.25/0