Binding of Carbon Monoxide to Isolated Hemoglobin Chains†

N. Alberding, Shirley S. Chan,‡ L. Eisenstein, H. Frauenfelder,* D. Good, I. C. Gunsalus, T. M. Nordlund,§ M. F. Perutz, A. H. Reynolds, and L. B. Sorensen

ABSTRACT: Binding of carbon monoxide to the separated α and β chains of hemoglobin, with and without bound *p*-mercuribenzoate, has been measured at temperatures from 5 to 340 K for times 2 μs to 1 ks using flash photolysis. All four proteins exhibit three different rebinding processes. The data are interpreted by a model in which the carbon monoxide, moving from the solvent to the binding site at the ferrous heme iron, encounters three barriers. The temperature dependences of the three processes yield activation enthalpies and entropies for the three barriers for all four proteins. Binding at temperatures below about 200 K is nonexponential, implying that the innermost barrier has a distribution of activation enthalpies. The distributions for the four proteins have been determined. At temperatures below 30 K, the CO binding rates approach finite low-temperature limits; binding thus proceeds by quantum-mechanical tunneling. Invoking a simple model, the widths of the innermost barriers are extracted from the measured tunneling rates. The experimental parameters are correlated with structural features of the hemoglobin chains and compared with previously published data on myoglobin and protoheme. A correlation is established between the height of the innermost barrier and the equilibrium CO pressure.

1. Ligand Binding to Heme Proteins

The binding of small ligands to proteins is an important step in many biomolecular reactions. Earlier, we have studied the binding of dioxygen (O_2) and carbon monoxide (CO) to myoglobin (Mb)[1] using flash photolysis (Austin et al., 1975). Two results are essential for the present work: (a) Access of the ligands to the binding site, the heme iron, is governed by multiple barriers. Binding of CO to Mb, for instance, is controlled by four potential barriers in sequence. (b) At low temperatures, rebinding after photodissociation is intramolecular and nonexponential in time. The innermost barrier must be described by an activation enthalpy distribution $g(H)$, where $g(H)\mathrm{d}H$ denotes the fraction of Mb molecules with activation enthalpies between H and $H + \mathrm{d}H$. These features lead to three questions: (i) How are the various barriers produced? (ii) How are they affected by protein structure? (iii) What is the relation between protein structure and the spectral function $g(H)$? We have already answered part of question i in another paper (Alberding et al., 1976b) by comparing the binding of CO to Mb and to protoheme. Heme alone shows two potential barriers; by comparing the properties of these with the four of Mb, we conclude that the innermost barrier is formed by the heme itself, the outermost by the solvent–protein interface. The intermediate barriers must be due to the globin. In the present paper, we obtain more insight into these problems by studying the binding of CO to the separated α and β chains of hemoglobin (αHb and βHb), with and without bound *p*-mercuribenzoate (HgBzO) (Bucci et al., 1965; Rosemeyer and Huehns, 1967; Muirhead et al., 1967). The structures of the separated chains are different from each other and from Mb. In addition, the β chains change structure when HgBzO is bound (Perutz et al., 1974). We show that the potential barriers and the distributions $g(H)$ also exhibit pronounced differences and interpret the observations in structural terms.

2. Experimental Section

We studied the binding of CO to the heme proteins by flash photolysis. The heme protein with bound ligand CO is placed in a cryostat at the desired temperature T and photodissociated by a laser flash. The subsequent rebinding of CO is followed optically with a transient analyzer with a logarithmic time base which records from 2 μs to 1 ks in a single sweep (Austin et al., 1976). The laser pulse, produced by a 0.1 J Coumarin 6 dye laser at 540 nm or a Rhodamine 6G dye laser at 590 nm, has a pulse width of 1 μs. In each experiment, the change in absorbance, $\Delta A(t) \equiv \log [I(t < 0)/I(t)]$ at a suitably chosen wavelength, is measured as a function of time (t) after photodissociation. $I(t)$ is the transmitted intensity of the monitoring light. From $\Delta A(t)$, the fraction $N(t)$ of protein molecules that have not rebound CO at time t is computed. Details of the experimental technique and data evaluation have been reported (Austin et al., 1975, 1976; Alberding et al., 1976b; Beeson, 1975; Frauenfelder, 1977).

Hemoglobin was prepared by standard procedures (Geraci et al., 1969) from washed, packed erythrocytes of freshly drawn human (TMN or AHR) blood. The α and β subunits were dissociated with HgBzO (Rosemeyer and Huehns, 1967) by the procedure of Yip et al (1972) and separated by column chromatography on DEAE-cellulose and CM-cellulose. Sep-

† From the Department of Physics, Biophysics, and Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and The MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, England. *Received July 18, 1977.* This work was supported in part by the U.S. Department of Health, Education, and Welfare under Grants GM 18051 and AM 00562, and the National Science Foundation under Grant No. PCM 74-01366.

‡ Present address: Max Planck Institut für Biophys. Chemie, D-3400 Göttingen-Nikolausberg, Germany.

* To whom correspondence should be addressed at the Department of Physics, University of Illinois.

§ Present address: Biozentrum, Klingelbergstr. 70, CH-4056 Basel, Switzerland.

[1] Abbreviations, symbols, and units used are: Mb, ferrous sperm-whale myoglobin; Hb, ferrous human hemoglobin; αHb and α^{SH}, the separated α chain of human hemoglobin; βHb and β^{SH}, the separated β chain of human hemoglobin; HgBzO, *p*-mercuribenzoate; α^{HgBzO}, the separated α chain of human hemoglobin with HgBzO bound; β^{HgBzO}, the separated β chain of human hemoglobin with HgBzO bound; SMb, sulfmyoglobin; DEAE, diethylaminoethyl; CM, carboxymethyl. Energies are given in kilojoule/mole; 1 kJ/mol = 0.239 kcal/mol = 0.010 eV. Entropies are given in terms of the dimensionless ratio S/k_B, where $k_B = 8.32$ J/mol K is the Boltzmann constant. H_{AB} and A_{AB} denote the activation enthalpy and frequency factor for transitions from well A to well B.

0006-2960/78/0417-0043$01.00/0

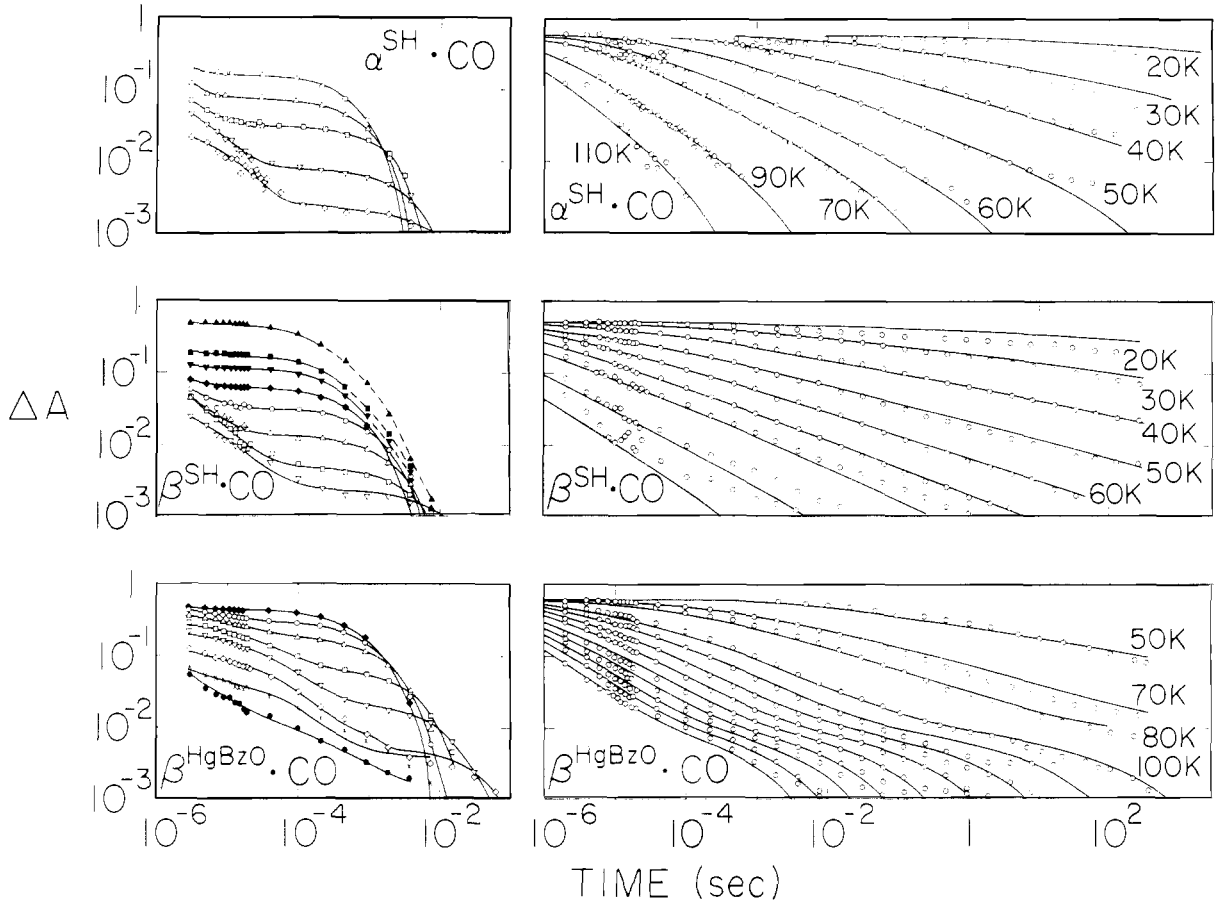

FIGURE 1: Rebinding of CO to α^{SH}, β^{SH}, and β^{HgBzO} chains over the temperature range from 20 to 340 K. For β^{SH} the low temperatures not labeled are 70, 80, 90, and 110 K. For β^{HgBzO} the low temperatures not labeled are 110–180 K, in steps of 10 K. For the high temperatures the symbol key is: (▲) 340 K, (■) 300 K; (▼) 290 K; (♦) 280 K; (○) 270 K; (△) 260 K; (□) 250 K; (▽) 240 K; (◇) 230 K; (×) 220 K; (●) 190 K. The solvent is glycerol-aqueous phosphate buffer, pH 7 (3:1, v/v); final buffer concentration 50 mM; [CO] = 3×10^{-4} M. The observation wavelength is 435 nm. The solid lines in the low-temperature data are fits based on eq 2, using the activation enthalpy distributions given in Figure 3 and Table I. The solid lines in the high-temperature data are fits based on the three-barrier model described in the text, with rate parameters given in Figure 6 and activation enthalpy distributions given in Figure 3 and Table I. The dashed lines are drawn to guide the eye.

aration of the subunits was verified by polyacrylamide disc gel electrophoresis (Gabriel, 1971). The sulfhydryl groups were regenerated by the procedure of Geraci et al. (1969). The number of HgBzO molecules per heme was estimated by optical difference spectra using the value ϵ_{mM}^{230nm} 21 for the extinction coefficient of the HgBzO mercaptide (Boyer, 1954). β^{HgBzO} had 2–3 and β^{SH} less than 0.2 Hg atoms per heme. Removal of HgBzO was also demonstrated by electrophoresis (Gabriel, 1971). The α and β subunits were concentrated by ultrafiltration, dialyzed against 50 mM potassium phosphate buffer (pH 7.0), and converted to the stable carbonmonoxy complex.

The standard solution for low-temperature flash photolysis contained glycerol-aqueous phosphate buffer (pH 7.0) (3:1, v/v) in order to minimize light scattering in the frozen state. (We call this solution glycerol-water, 3:1, v/v.) The final buffer concentration was 50 mM. This solution passes through a glass transition and becomes solid around 200 K at our cooling rates. It was prepared by alternately and repeatedly pumping out and stirring under 1 atm of CP grade carbon monoxide (Matheson) for at least 1 h. Deaerated sodium dithionite solution was then added to 2–3 mM to remove any residual oxygen. Hemoglobin was finally added and the solution stirred slowly under 1 atm of CO at room temperature for another half hour. Heme concentrations were 0.15 to 0.3 mM. The optical path length of the sample cell was 0.5 mm.

3. Results

Binding of photodissociated CO to α^{SH}, β^{SH}, and β^{HgBzO} chains over the temperature range from 20 to 340 K is shown in Figure 1. Data for α^{HgBzO} are omitted because they are virtually identical with those for α^{SH}. Figure 2 gives the normalized absorption change at 10 K for β^{SH}, β^{HgBzO}, protoheme (Alberding et al., 1976b), and myoglobin. The general features of the Hb rebinding curves are like those found in Mb (Austin et al., 1975). At temperatures below about 150 K, a single nonexponential curve is seen which we denote by I. In the low-temperature limit, $T \rightarrow 0$, process I does not disappear; rebinding still takes place with a finite rate. The curves in Figure 2 at 10 K are already close to the low-temperature limits. Between about 200 and 300 K, a second nonexponential process takes place which we call II. At temperatures above about 230 K, a third process becomes observable which, in the limit [CO] $\gg$ [protein], is exponential with rate proportional to the CO concentration. We denote this by process IV. With increasing temperature, IV becomes more dominant and above about 300 K is the only process observed. While these features are similar in Mb and the separated Hb chains, some aspects differ from protein to protein. In particular, Mb shows an additional process, called III, that has no counterpart in the Hb chains. Also, at all temperatures, the rebinding to the Hb chain is faster than to Mb.

Preparations of the same species of Hb subunits show more

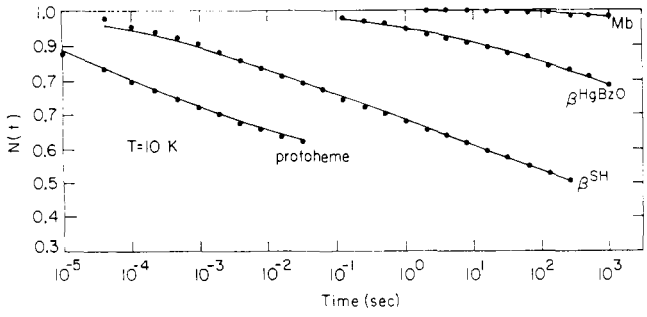

FIGURE 2: Rebinding of CO to Mb, β^{SH}, β^{HgBzO}, and protoheme at 10 K. The solid lines are the fits to a quantum-mechanical molecular tunneling model with $\delta = 1$ (eq 6) and d_0 given in Table I.

sample-to-sample variability than preparations of Mb. In our comparison of α^{SH}, α^{HgBzO}, β^{SH}, and β^{HgBzO} we only discuss features that cannot be attributed to these variations.

4. Data Evaluation

We interpret the data in Figure 1 by assuming that a CO molecule coming from the solvent moves to the binding site at the heme iron by overcoming three sequential barriers. We denote the binding site by A, the solvent by E, and the two intermediate wells (or states) by B and C.[2] Denoting the rate parameter for the transition from B to A by k_{BA}, we postulate the scheme:

$$A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B \underset{k_{CB}}{\overset{k_{BC}}{\rightleftharpoons}} C \underset{k_{EC}}{\overset{k_{CE}}{\rightleftharpoons}} E \qquad (1)$$

A represents carbonmonoxy Hb, where the CO is bound to the heme iron. During photodissociation, the Fe-CO bond is broken and the CO moves to B. At low temperatures, CO cannot leave the protein but returns directly to A, giving rise to process I. At higher temperatures, CO can either return directly, or move to C. From C it can either return to B or move into the solvent, E. Return from the solvent is seen as process IV which is proportional to the CO concentration.

Analysis is performed by solving the coupled linear differential equations describing the motion of a particle in four wells (Austin et al., 1975; Beeson, 1975; Frauenfelder, 1977). When the number of barriers is equal to the number of observed kinetic processes, each rate is uniquely determined. From the temperature dependence of the rates, we deduce activation enthalpies and entropies for the potential barriers.

The assumption that the three barriers are serial requires justification. Mb embedded in a solid matrix supplies evidence for a sequential arrangement of the inner barriers (Austin et al., 1975). Additional support comes from the temperature dependence of process IV. If all barriers were parallel, rates for all processes should follow the Arrhenius law. The rate for process IV deviates from an Arrhenius behavior at high temperatures and this departure is accounted for by serially arranged barriers. Nevertheless, it should be kept in mind that eq 1 is based on the simplest assumption that fits all data well, and that more complex schemes, with competing channels, are not ruled out. Moreover, our numbers of barriers (four in Mb and three in Hb) are lower limits. Experimental refinements may reveal additional barriers and alternate paths.

5. The Innermost Barrier

Figure 1 shows that only process I occurs below about 150 K, and we interpret process I as corresponding to the direct transition B → A. The most conspicuous property of I is its time dependence: rebinding is not exponential with time, but approximately follows a power law, $\Delta A(t) \propto (1 + t/t_0)^{-n}$. Here, t_0 and n are temperature-dependent parameters. A nonexponential time dependence is to be expected if the barriers between B and A are not identical in all Hb molecules, but are described by a distribution $g(H_{BA})$ (Austin et al., 1974, 1975; Alberding et al., 1976a,b; Frauenfelder, 1977). H_{BA} is the activation enthalpy for the step B → A. (In some systems, the activation entropy S_{BA} must also be characterized by a distribution (Alberding et al., 1976b; Frauenfelder, 1977).) Assuming a distributed enthalpy, process I is given by:

$$N(t) = \int_0^\infty dH_{BA} g(H_{BA}) e^{-k_{BA}t} \qquad (2)$$

where $N(t)$, the fraction of Mb molecules that have not rebound a CO at time t after photodissociation, is proportional to $\Delta A(t)$. At temperatures where quantum-mechanical tunneling can be neglected, the rate parameter k_{BA} depends on H_{BA} and S_{BA} through:

$$k_{BA}{}^c = \nu e^{S_{BA}/k_B} e^{-H_{BA}/k_B T} \qquad (3)$$

The superscript c refers to the "classical" Arrhenius motion, $k_B = 8.32$ J mol^{-1} K^{-1} is the Boltzmann constant, and we take the characteristic frequency ν to have the value $\nu = 10^{13}$ s^{-1}. To determine the spectral function $g(H_{BA})$, we approximate the observed rebinding function $N(t)$ by a power law; eq 2 can then be inverted and $g(H_{BA})$ approximated by (Austin et al., 1975; Alberding et al., 1976b; Frauenfelder, 1977):

$$g(H_{BA}) = \text{const} \exp\left[\alpha(H_{BA}{}^p - H_{BA}) - n_f \exp\left\{ \frac{\alpha}{n_f} (H_{BA}{}^p - H_{BA}) \right\} \right] \qquad (4)$$

The normalization constant is determined by the condition $\int_0^\infty dH_{BA} g(H_{BA}) = 1$. $H_{BA}{}^p$, α, and n_f are parameters chosen so that $N(t)$ as calculated from eq 2-4 provides a best fit to the data. $H_{BA}{}^p$ gives the enthalpy of the peak of the distribution, α characterizes the exponential drop-off of $g(H_{BA})$ beyond $H_{BA}{}^p$, and n_f determines the low-enthalpy turn-on of $g(H_{BA})$.

The distribution eq 4 describes process I very well for α^{SH}, α^{HgBzO}, and β^{SH}. β^{HgBzO}, however, shows a small slow phase and cannot be fit by eq 4. A numerically constructed enthalpy spectrum reproduces the observed kinetics accurately, but for further evaluation of the data we use an approximate $g(H_{BA})$ which ignores the slow phase. The parameters for the Hb chains are summarized in Table I; the corresponding distributions are shown in Figure 3. For comparison, the corresponding parameters for Mb and protoheme are listed and shown. The theoretical fits using the enthalpy distributions of Figure 3, shown as solid lines in Figure 1, agree very well with the experimental data.

6. Quantum-Mechanical Molecular Tunnel Effect

The Arrhenius rate parameter k_{BA}, eq 3, vanishes in the low-temperature limit $T \to 0$. Calculations of $N(t)$ with eq 2 and 3 and the parameters given in Table I and Figure 3 imply that binding should be extremely slow below about 30 K. In reality, however, binding after photodissociation occurs even at 2 K. Typical curves at 10 K, where Arrhenius processes should be unobservable, are given in Figure 2. Quantum-mechanical tunneling provides a natural explanation for the finite low-temperature binding rates (Goldanskii, 1976). In two earlier papers, we have evaluated the tunneling effect in

[2] For consistency with earlier work on Mb and protoheme, state D is left out.

TABLE I: Parameters for Process B → A in CO Rebinding to Heme Proteins in Glycerol–Water (3:1, v/v).

Protein	$H_{BA}{}^{p}$ [a] (kJ mol^{-1})	Log A_{BA} [a] (s^{-1})	α (mol kJ^{-1})	n_1	d_0 (nm)
α^{SH}, α^{HgBzO}	4.6	8.6	0.88	0.88	
β^{SH}	4.3	9.4	0.65	1.1	0.097 ± 0.01
β^{HgBzO}	8.5	9.2	0.50	1.0	0.105 ± 0.02
Mb	10.0	8.7	0.31	0.31	0.12 ± 0.03
Protoheme	1.0	9.0	0.48	0.04	0.077 ± 0.02

[a] Errors in $H_{BA}{}^{p}$ are estimated at 10% and those in log A_{BA} at ±0.4, except for protoheme where the errors are larger (20% and ±0.5) due to the speed of recombination.

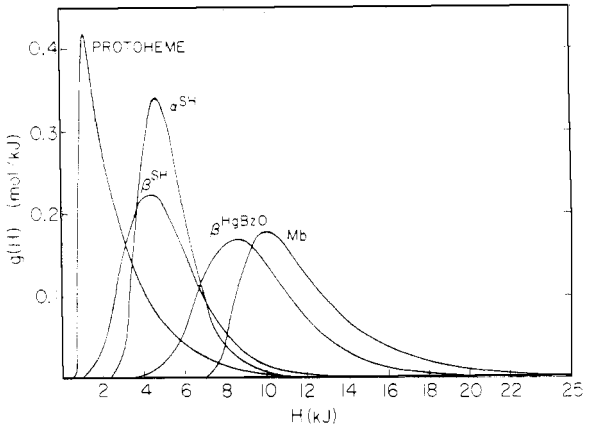

FIGURE 3: Activation enthalpy spectra of the innermost barrier in the binding of carbon monoxide to protoheme, Mb, α^{SH}, β^{SH}, and β^{HgBzO}. Here $g(H_{BA})dH_{BA}$ denotes the probability of finding a molecule with activation ethalpy between H_{BA} and $H_{BA} + dH_{BA}$.

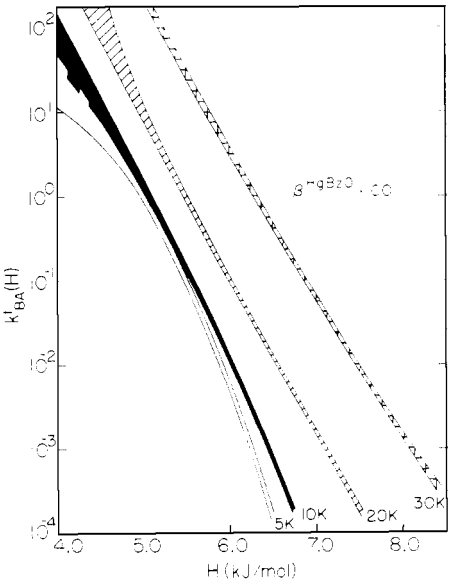

FIGURE 4: The tunneling rate as a function of the activation enthalpy H_{BA} for β^{HgBzO}. The uncertainty in the tunneling rates is indicated by the bands at each temperature.

the binding of CO to β^{SH} and protoheme with a very simple approach (Alberding et al., 1976a,b). Here and in section 7 we introduce a more general method to extract from Figure 2 the information that is meaningful for the discussion of the relation between structure and function. Because of the model change the numerical values of the barrier parameters given here differ from the ones quoted in earlier papers.

Arrhenius and tunneling transitions are but two aspects of the same quantum-mechanical process (Fong, 1975) and consequently are governed by the same activation enthalpy spectrum $g(H_{BA})$. Indeed, the curves shown in Figure 2 are also nonexponential. In order to evaluate the data in the tunneling regime, we write the total transition rate parameter for the step B → A as:

$$k_{BA} = k_{BA}{}^{c} + k_{BA}{}^{t} \qquad (5)$$

where $k_{BA}{}^{c}$ is given by eq 3 and $k_{BA}{}^{t}$ is the tunneling rate parameter. The first term in eq 5 dominates at high, the second at low temperature. Above about 30 K, $g(H_{BA})$ can be found as described in section 5. This distribution, characterized in Table I and Figure 3, is then used in eq 2 to find the total rate parameter k_{BA} from $N(t)$ measured at temperatures below about 30 K. The numerical extraction of k_{BA} as a function of H_{BA} is done by computer. From k_{BA}, the tunneling rate parameter $k_{BA}{}^{t}$ is found by subtracting $k_{BA}{}^{c}$ which can be calculated from eq 3 and Table I. Figure 4 gives log $k_{BA}{}^{t}$ for β^{HgBzO} as a function of H_{BA} for various temperatures. The values at 10 K are extracted from Figure 2; at other temperatures, additional data are used. To discuss the behavior of the tunneling rates, a specific model must be introduced and we will do this in the following section.

7. A Model for Binding at the Heme

Process I, the step B → A in the binding sequence of eq 1, must take place at the heme group. Enough is known about the molecular states involved so that a simple model can be constructed. In state A, CO is bound to the Fe which has spin $S = 0$ and lies in the heme plane. The potential energy as a function of Fe position thus has its minimum in the heme plane. In state B, CO is not bound and the iron has spin $S = 2$ and lies out of the heme plane. The two states B and A and the corresponding potential are shown schematically in Figure 5. In the binding step B → A, CO approaches the heme and Fe moves in concert over the barrier between B and A and changes from $S = 2$ to $S = 0$. We preface the treatment of the transition by some remarks about the meaning of the potentials in Figures 5 and 7. In addition to the concerted motion of Fe and CO, the step B → A involves movement and change in shape of the heme molecule (Gelin and Karplus, 1977; Warshel, 1977). It is difficult to draw a potential that represents this process fully. We therefore select potentials that stress some desired feature. Figure 7, for instance, treats the binding process from the solvent to the heme iron as motion of the CO molecule in a fixed potential produced by the protein, the heme, and the iron. Figure 5, on the other hand, gives the potential of the full system; the particle shown in well B represents the state of the system.

The energy of the system is quantized in wells A and B and we show schematically some of the levels in Figure 5. After

photodissociation, the system is initially in B. In the low-temperature limit, $T \rightarrow 0$, the system occupies the ground state in B. The only way it can then move to A is by quantum-mechanical molecular tunneling through the barrier. At finite temperatures, the system can populate excited states in well B. As long as these states are below the top of the barrier, transition to A can occur only by tunneling. Once levels above the top are reached, the system can move classically to A (Arrhenius transition).

The rate parameter for Arrhenius transitions, eq 3, does *not* depend on barrier width. In contrast, the tunneling rate parameter is a rapidly varying function of the barrier width. It is therefore possible to determine the barrier height at high temperatures, where tunneling can be neglected, and then use the low-temperature data to also find the barrier width. The first part of this program has already been performed in section 5, with results as shown in Table I and Figure 3. The second part, evaluation of tunneling, is complicated[3] and details will be given in a separate publication. The result of interest is the distance d between the center of well B and the "wall" of well A as shown in Figure 5.

The value of the tunneling distance quoted in Table I requires some additional remarks. Figure 5 shows a potential with given values of barrier height, H_{BA}, and width, d. We have noted in section 5, however, that an ensemble of biomolecules at low temperatures must be described by an enthalpy spectrum, $g(H_{BA})$. We interpret this observation by assuming that the ensemble contains biomolecules in different conformational states. In a given conformational state, the barrier B $\rightarrow$ A is specified by H_{BA} and d. In going from one conformational state to another, H_{BA} and d change. In the absence of an understanding of the relation linking structure and barrier height, we assume a simple relationship between H_{BA} and d:

$$d(H_{BA}) = d_0(H_{BA}/H_{BA}{}^{p})^{\delta} \qquad (6)$$

where δ is a fitting parameter and d_0 is the width corresponding to the peak enthalpy $H_{BA}{}^{p}$. The optimum value for δ varies somewhat (± 0.5) from protein to protein, but acceptable fits for all systems can be obtained with $\delta = 1$. The choice $\delta = 1$ implies that the barrier shapes in the various conformers are identical. The values of d_0, for $\delta = 1$, are listed in Table I; the corresponding fits are given by the solid lines in Figure 2.

8. The Outer Barriers

Above about 200 K, Figure 1 shows two processes, denoted by II and IV. II follows a power law, while IV is approximately exponential. We explain the presence of the two processes by the scheme of eq 1. To describe their different shapes we assume that the protein can exist in many different conformational states, with different activation enthalpies. The change from one conformational state to another will require, on the average, a relaxation time τ_r (Austin et al., 1975). Conformational relaxation depends on temperature. At low temperatures, τ_r will be long compared to the rebinding times, each protein will stay in one conformational state, and the barrier B $\rightarrow$ A must be described by the spectral function $g(H_{BA})$. At high temperatures, τ_r becomes faster than the rebinding processes and conformational relaxation takes place. Each protein molecule moves rapidly from one state to another and rebinding becomes exponential with a rate given by $k_{BA}{}^{mean} = \int k_{BA} g(H_{BA}) dH_{BA}$. For the actual calculations, we approximate $k_{BA}{}^{mean}$ by $k_{BA}{}^{p}$, obtained by inserting the peak enthalpy $H_{BA}{}^{p}$ into eq 3. The relaxation time, τ_r, is not well determined by the data; equally good fits are obtained in most cases when relaxation is allowed to occur any time between the end of process II and the beginning of IV.

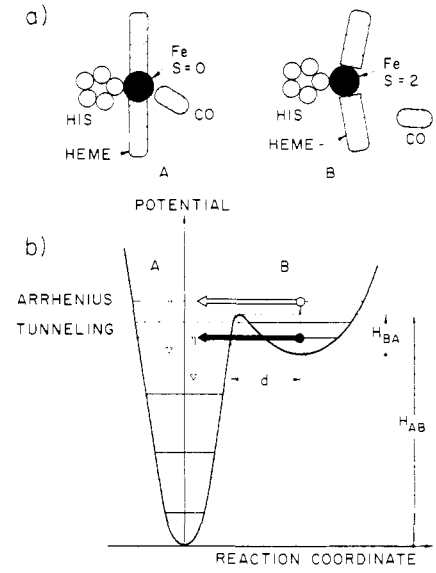

FIGURE 5: (a) Schematic showing positions of the histidine group of the fifth ligand, heme group, Fe, and CO in states A and B. (b) Potential energy surfaces of the system in wells B and A. The Arrhenius and tunneling processes are indicated schematically in the figure.

If only the innermost barrier, B $\rightarrow$ A, is assumed to be distributed, it is impossible to reproduce the experimental data well. In particular the fits to the β^{HgBzO} data would be poor. It is likely that the barrier B $\leftrightarrow$ C is distributed as well (Austin et al., 1975, Section 10.4). Such a possibility is not unexpected. Assume the activation enthalpy to be given by a distribution as in Figure 3, with H_0 and $H_0 + \Delta H$ denoting the minimum and maximum and ΔH the width of the spectrum. The ratio of rate parameters corresponding to H_0 and $H_0 + \Delta H$:

$$\frac{k^{max}}{k^{min}} = \frac{\exp(-H_0/k_BT)}{\exp[-(H_0 + \Delta H)/k_BT]} = \exp(\Delta H/k_BT)$$

is only determined by ΔH. Even a small spread on a high barrier can lead to markedly nonexponential rebinding. Unfortunately, it is difficult to find $g(H_{CB})$ and $g(H_{BC})$ unambiguously because process II can only be observed over relatively small temperature and time ranges. For simplicity we assume barrier C $\leftrightarrow$ B, and thus H_{BC} and H_{CB}, to be distributed in the same way as H_{BA}. The functions $g(H_{BC})$ and $g(H_{CB})$ are taken to have the shape and spread shown for $g(H_{BA})$ in Figure 3, but shifted to have peaks at $H_{BC}{}^{p}$ and $H_{CB}{}^{p}$, respectively.

With the assumptions described above and the procedure sketched in section 4 the various rate parameters are determined at each temperature. The results are plotted vs. $1/T$ in Figure 6. From these plots, the activation enthalpies and frequency factors listed in Table II are derived. Activation entropies are calculated by setting $A = \nu \exp(S/k_B)$ and using $\nu = 10^{13}\ s^{-1}$ for first-order and $\nu = 10^{11}\ M^{-1}\ s^{-1}$ for second-order (E $\rightarrow$ C) rate parameters. The barriers deduced by this procedure are given in Figure 7. The theoretical curves are

[3] To evaluate the data below about 30 K, we approximate the barrier in Figure 5 by a combination of a parabolic well and a square barrier, assume a Boltzmann distribution in well B, take the reduced mass of CO and Fe as tunneling mass, calculate the Gamow factor with the WKB (Wentzel, Kramers, and Brillouin) approximation, use eq 4 and 6 to characterize the barriers in different conformational states, and postulate the tunneling preexponential factor to be independent of the excitation energy in well B.

TABLE II: Activation Enthalpies (H), Frequency Factors (A), and Activation Entropies (S) for the Binding of CO to the Separated Hb Chains (Present Paper), to Mb (Austin et al., 1975), and to Protoheme (Alberding et al., 1976b).[a]

Transition	Protein	H (kJ mol^{-1})	Log A (s^{-1})	Log A' (M^{-1} s^{-1})	S/k_B
Innermost barrier					
B → A	α^{SH}	4.6	8.6 ± 0.4		−10.1
	β^{SH}	4.3	9.4 ± 0.4		−8.3
	β^{HgBzO}	8.5	9.2 ± 0.4		−8.7
	Mb	10.0	8.7 ± 0.4		−9.9
D → A	Protoheme	1.0 ± 0.2	9.0 ± 0.5		−9.2
B → C	α^{SH}	40 ± 10	15.1		4.8
	β^{SH}	23 ± 15	12.1 ± 3.0		−2.1
	β^{HgBzO}	43	16.3		7.6
	Mb[e]	46	16.5		8.1
		51	15.5		5.8
C → B	α^{SH}	19 ± 10	9.4 ± 2.0		−8.3
	β^{SH}	0 ± 20	5.5 ± 4.0		−17.5
	β^{HgBzO}	18 ± 4	8.6		−10.1
	Mb[f]	14	8.3		−10.8
		81	20.9		18.2
Outermost barrier					
C → E	α^{SH}	48	14.8		4.1
	β^{SH}	52	14.9		4.4
	β^{HgBzO}	67	17.7		10.8
D → E	Mb	91	22.7		22.3
	Protoheme	60	20.2		16.5
E → C	α^{SH}	37	10.6[b]	14.1	7.1
	β^{SH}	44	11.5[b]	15.0	9.2
	β^{HgBzO}	54	13.8[b]	17.3	14.5
E → D	Mb	79	18.3[c]	22.8	27.2
	Protoheme	74	17.9[d]	21.3	23.7

[a] Errors in H and log A are about 10%, except as noted. [b] Pseudo-first-order rate, [CO] = 3×10^{-4} M. [c] Pseudo-first-order rate, [CO] = 3×10^{-5} M. [d] Pseudo-first-order rate, [CO] = 4×10^{-4} M. [e] First entry is for associating B → C in Mb with B → C in separated Hb chains; second entry is for associating C → D in Mb with B → C in separated Hb chains. [f] First entry is for associating C → B in Mb with C → B in separated Hb chains; second entry is for associating D → C in Mb with C → B in separated Hb chains.

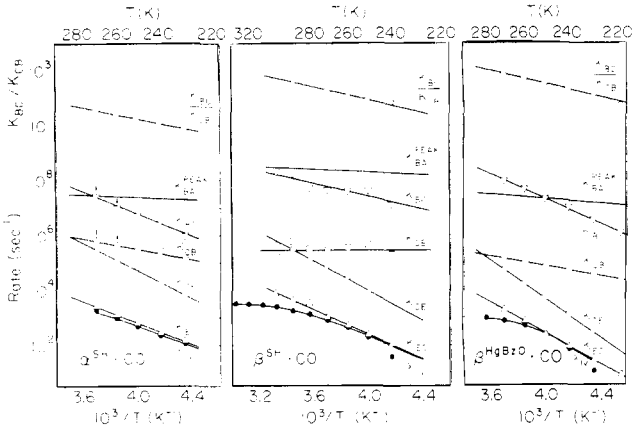

FIGURE 6: Rate parameters for the binding of carbon monoxide to α^{SH}, β^{SH}, and β^{HgBzO}, [CO] = 3×10^{-4} M, solvent glycerol-water (3:1, v/v). Activation enthalpies, entropies, and frequency factors deduced from these Arrhenius plots are given in Table II and Figures 7 and 8. The lines through the solid dots are the theoretical predictions for λ_{IV}.

shown as solid lines in Figure 1. The fits are very good; the complicated behavior of $N(t)$ is reproduced over large temperature and time ranges. The remaining uncertainty lies in k_{BC} and k_{CB}. While the ratio k_{BC}/k_{CB} is well fixed by the experimental data, the individual values have large errors. Extension of the experiments to shorter times should cure the deficiency.

Also given in Figure 6 are the rates λ_{IV} for process IV, which do not obey an Arrhenius law. The pronounced curvature of the rate parameter, λ_{IV}, is explained by our model in which this process corresponds to the motion of the CO from the solvent over all three barriers to the iron site. The excellent agreement between experiment and theory can be seen in Figure 6, where the theoretical predictions for λ_{IV}, based on the measured rate parameters $k_{BA}, \ldots, k_{EC}$, are plotted as solid lines.

In the next section we correlate the distinct kinetic behaviors of the different subunits with their structural differences. Kinetic features distinguishing α^{SH} from β^{SH} and β^{HgBzO} from β^{SH} can be seen distinctly by comparing the rates in Figure 6 and the data in Figure 1.

9. Structure and Function

Despite numerous amino acid replacements, Mb and the Hb chains have similar tertiary structures (Perutz, 1976). The overall dynamic behavior of these proteins mirrors this similarity; all possess similar multiple barriers (Figure 7) and similar enthalpy spectra (Figure 3). Closer inspection, however, reveals significant differences and we will now attempt to correlate the kinetic properties with structural features of the Hb chains, Mb, and the isolated protoheme.

We begin the discussion with the innermost barrier, between wells B and A. In section 7 and Figure 5, we introduced a simple model that accounts for this barrier. In this model the peak barrier height H_{BA}^{P} is determined by the depth and width of well A, and position and width of well B.[4] For fixed well shapes and positions, H_{BA}^{P} decreases with increasing depth H_{AB} (Bronsted relation). H_{BA}^{P} is found from the low-tem-

[4] For the discussion here we do not need to consider the various conformational states and can replace the variable H_{BA} by its value, H_{BA}^{P}, at the peak of the distribution.

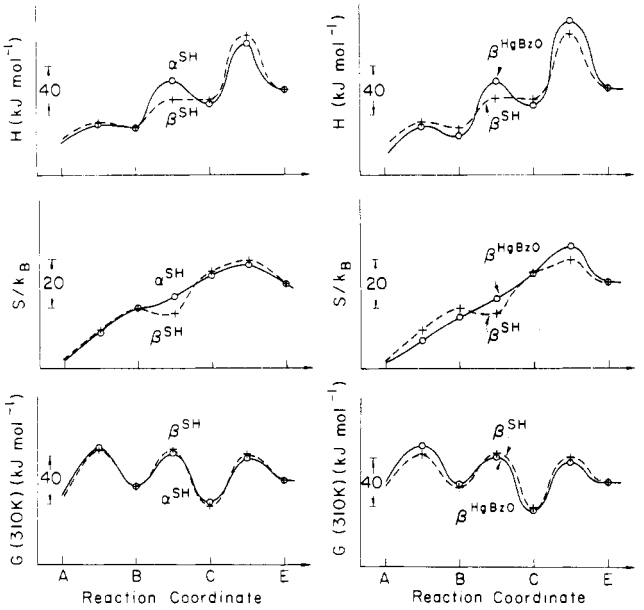

FIGURE 7: Enthalpy, entropy, and Gibbs energy (at 310 K) for isolated Hb chains as a function of reaction coordinate. On the left HgBzO-free β chains (crosses) are compared with HgBzO-free α chains (circles). On the right HgBzO-free β chains (crosses) are compared with HgBzO-bound β chains (circles). S_{EC}/k_B and G_{EC} are calculated for a second-order reaction with unit mole fraction of CO. Thermodynamic quantities in well E have been set equal for the two proteins.

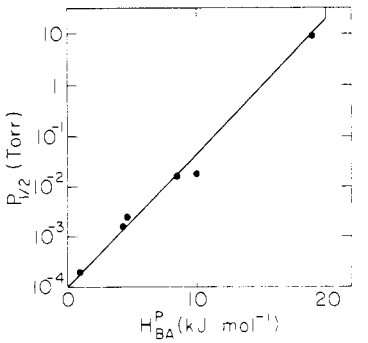

FIGURE 8: Correlation between the peak enthalpy (H_{BA}^{P}) of the innermost barrier and the equilibrium CO pressure ($p_{1/2}$).

TABLE III: Correlation between Peak Enthalpy (H_{BA}^{P}) of the Innermost Barrier, Equilibrium CO Pressure ($p_{1/2}$), and Tunneling Barrier Width (d_0).

System	H_{BA}^{P} [a] (kJ mol^{-1})	$p_{1/2}$ [b] (Torr)	d_0 (nm)
Protoheme CO	1	0.0002	0.077 ± 0.02
β^{SH} CO	4.3	0.0015	0.097 ± 0.01
α^{SH} CO (α^{HgBzO} CO)	4.6	0.0025	
β^{HgBzO} CO	8.5	0.015	0.105 ± 0.02
Mb CO	10	0.018	
SMb CO [c]	5.5		
	19	9	

[a] Errors in H_{BA}^{P} are estimated at 10% for Mb and the Hb chains and at 20% for heme and sulfmyoglobin. [b] The $p_{1/2}$ measurements were performed with solvents different from ours but the general trend should not be affected. The $p_{1/2}$ values for the Hb chains are taken from Antonini and Brunori (1971) and for sulfmyoglobin from Berzofsky et al. (1972). The $p_{1/2}$ value for protoheme is estimated from those obtained for deuterohemes in benzene (Rougee and Brault, 1975). [c] Sulfmyoglobin (SMb) has an activation enthalpy spectrum with two peaks; we tentatively assign the observed value of $p_{1/2}$ to the higher one. Evaluation of the SMbCO high-temperature data supports this assignment.

perature rebinding curves and is given in Table I. The depth of well A is close to the calorimetrically measured heat of binding H_{EA}; it also determines the off-rate parameter k_{AE} and enters the equilibrium constant K and pressure $p_{1/2}$, the pressure of CO gas at which one-half of the heme molecules have CO bound. The available data for $p_{1/2}$ and H_{BA}^{P} are collected in Table III; $p_{1/2}$ is plotted vs. H_{BA}^{P} in Figure 8. It shows the expected correlation: log $p_{1/2}$ increases with increasing H_{BA}^{P}. The model of Figure 5 makes a second prediction: the barrier width d should decrease with increasing depth of well A. The barrier widths, as determined from the tunneling rate in section 7, are also listed in Table III and are not inconsistent with the predicted correlation.

The model in Figure 5 suggests how the off- and the on-rates at the active center could be affected differently by the protein structure. The off-rate is mainly determined by H_{AB} and hence by the strength of the Fe–ligand bond. The bond strength may be controlled through the axial and the planar charge distribution in the heme (Falk et al., 1966; Peisach, 1975; Sono et al., 1976). The on-rate is essentially proportional to $\exp(-H_{BA}/k_BT)$ (Austin et al., 1975); the barrier height H_{BA} can be changed, for instance, by adjusting the position and width of well B through the tension exerted by the proximal histidine, or by changing the width and depth of well A.

So far we have considered the enthalpies at the innermost barriers. The entropy factor A_{BA} also provides information. We first note that all log A_{BA} lie within a narrow range; the structure appears to affect the entropy less than the enthalpy. However, one effect is seen in comparing α^{SH} and β^{SH}; the latter has a larger preexponential than the former. The ratio can be explained by assuming, as a simple model, that the number of states in well B is proportional to the volume V and that A_{BA} is inversely proportional to the number of states and hence to V. The volume ratio of wells B in α^{SH} and β^{SH} then becomes $V^{\alpha}/V^{\beta} \simeq A_{BA}{}^{\beta}/A_{BA}{}^{\alpha} = 6$. It is known from x-ray studies (Perutz, 1970) that the pocket is bigger in the α chain; thus we can tentatively identify well B with the pocket. Work with other heme proteins and with mutants is necessary to conform or invalidate this assumption.

Can the barrier properties be correlated with the protein structures? The innermost barrier should be related to the energy needed to bring the iron atom from its out-of-plane position in deoxy- to its position in the plane of the porphyrin in carbonmonoxyhemoglobin. This movement is linked to changes in the tertiary structure of the polypeptide segments surrounding the hemes. The energy barrier will therefore be influenced by the rigidity of that structure. H_{BA} is higher for myoglobin than for the two Hb chains. Myoglobin also has a two–three times lower oxygen affinity than the free chains (Baldwin, 1975); in derivatives in which there is a thermal equilibrium between two alternate spin states, that equilibrium is biased toward higher spin in myoglobin than in hemoglobin (Iizuka and Kotani, 1969a,b); the displacement of the iron atoms from the plane of the porphyrin is greater in aquometmyoglobin than in aquomethemoglobin. It will be shown below that the heme is more rigidly held in myoglobin than in hemoglobin (Takano, 1977a,b; Ladner et al., 1977). Therefore, ligand affinities, spin equilibrium, and molecular structure are all consistent with the innermost enthalpy barrier to CO binding being higher in myoglobin than in the hemoglobin chains.

Reaction with HgBzO leaves the barriers of the α chain unchanged but raises the heights of the barriers in the β chain.

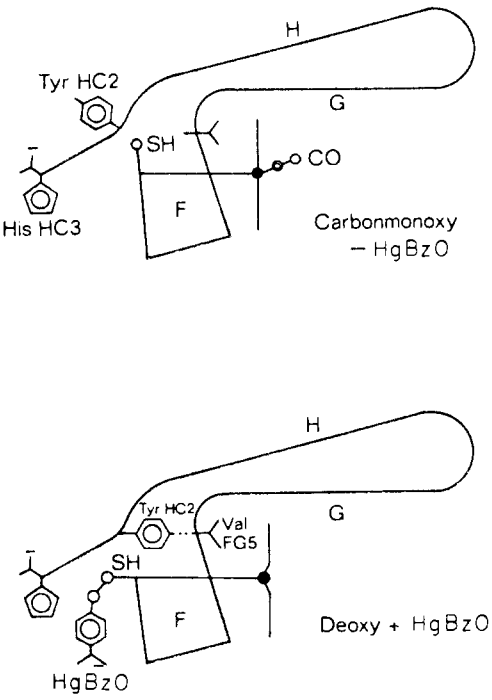

FIGURE 9: Alternative conformations of side chains of cysteine F9(93)β and tyrosine HC2(145)β. (Top) SH in pocket, tyrosine out of the pocket, favors iron being in the plane of the porphyrin. (Bottom) SH held out of the pocket by HgBzO, tyrosine in the pocket, favors iron being out of the plane of the porphyrin. Conversely, the SH reactivity rises with rising spin of the heme (Bucci et al., 1965). Both diagrams picture the quaternary R structure, so that the salt bridge between His HC3 and Asp FG1 is open.

HgBzO reacts with cysteines G11(104)α and G14(112)β which are both far from the heme and have little influence on its behavior; it also reacts with cysteine F9(93)β which is the residue immediately following the heme-linked histidine and does affect the heme. Cysteine F9 can influence the heme because its side chain can take up the two alternative conformations shown in Figure 9. In one conformation the SH group is external and therefore highly reactive; tyrosine HC2(145)β lies in the pocket between helices F and H; its phenolic OH is hydrogen bonded to the CO of valine FG5(97)β. This hydrogen bond stabilizes the tertiary structure of the β chain in the deoxy or high-spin conformation with the iron out of the plane of the porphyrin. In the alternative conformation the SH group is in the pocket between helices F and H, displacing the tyrosine side chain, so that the tertiary structure is loose. This conformation is dominant in carbonmonoxyhemoglobin and other low-spin forms (Heidner et al., 1976). HgBzO fixes the SH group in the external position, thereby "bending" the β chain toward its tertiary deoxy, or high-spin, structure and lowering the oxygen affinity. The raising of H_{BA} by HgBzO is therefore consistent with its other effects.

We now come to the outer barriers. Myoglobin and hemoglobin are so constructed that, in the equilibrium position of all atoms, there is no hole through which ligands such as O_2 or CO can enter or leave the heme pocket. A door must open. The opening mechanism is obscure, but it must involve widespread changes of conformation. The energy barrier opposing these changes must therefore be related to the rigidity of the tertiary structure. Myoglobin is more rigidly constructed than either of the hemoglobin chains, mainly as a result of hydrogen bonds linking each of the propionates of the heme to basic side chains of the globin (Takano, 1977a). This stabilization may be the reason why the outer barriers are markedly higher in myoglobin than in the hemoglobin chains. It is also noteworthy that HgBzO raises all the enthalpy barriers in β chains; this happens probably because it allows the hydrogen bond between tyrosine HC2(145) and valine FG5(97) to be formed. The bond may tighten up the entire molecule and oppose the changes in tertiary structure needed for entry or exit of ligands. Moreover, $C_{\gamma 2}$ of valine E11, one of the two distal residues, is closer to the ligand and offers more steric hindrance to its combination with the iron atom in the β than in the α chains (Ladner et al., 1977).

The present work is clearly only a beginning, and the three questions posed in the first section have not been answered fully. However, the understanding obtained so far shows that low-temperature measurements provide a tool to investigate the connection between structure and function in detail. Studies on other systems, using mutant proteins, other ligands, other heme side groups, and other effectors will very likely unravel many of the relations that are at present still obscure.

Acknowledgments

We thank R. H. Austin, P. G. Debrunner, G. DePasquali, M. Marden, T. Pederson, and G. C. Wagner for support, help, suggestions, and criticism. Four of us (S.S.C., H.F., T.M.N., and M.F.P.) would also like to thank Manfred Eigen, whose stimulating Winter Seminar gave us the possibility of meeting and exchanging ideas.

References

Alberding, N., Austin, R.H., Beeson, K. W., Chan, S. S., Eisenstein, L., Frauenfelder, H., and Nordlund, T. M. (1976a), *Science 192,* 1002.

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

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

Austin, R. H., Beeson, K., Eisenstein, L., Frauenfelder, H., Gunsalus, I. C., and Marshall, V. P. (1974), *Phys. Rev. Lett. 32,* 403.

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

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

Baldwin, J. M. (1975), *Prog. Biophys. Mol. Biol. 29,* 225.

Beeson, K. W. (1975), Ph.D. Thesis, University of Illinois at Urbana-Champaign.

Berzofsky, J., Peisach, J., and Alben, J. O. (1972), *J. Biol. Chem. 247,* 3774.

Boyer, P. D. (1954), *J. Am. Chem. Soc. 76,* 4331.

Bucci, E., Fronticelli, C., Chiancone, E., Wyman, J., Antonini, E., and Rosse-Fanelli, A. (1965), *J. Mol. Biol. 12,* 183.

Falk, J. E., Phillips, J. N., and Magnusson, E. A. (1966), *Nature (London) 212,* 1531.

Fermi, G. (1975), *J. Mol. Biol. 97,* 237.

Fong, F. K. (1975), Theory of Molecular Relaxation, New York, N.Y., Wiley.

Frauenfelder, H. (1977), *Methods Enzymol.* (in press).

Gabriel, O. (1971), *Methods Enzymol. 22,* 565.

Gelin, B. R., and Karplus, M. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 801.

Geraci, G., Parkhurst, L. J., and Gibson, Q. H. (1969), *J. Biol. Chem. 244,* 4664.

Goldanskii, V. I. (1976), *Annu. Rev. Phys. Chem. 27,* 85.

Heidner, E. J., Ladner, R. C., and Perutz, M. F. (1976), *J. Mol. Biol. 104,* 707.
Iizuka, T., and Kotani, M. (1969a), *Biochim. Biophys. Acta 181,* 275.
Iizuka, T., and Kotani, M. (1969b), *Biochim. Biophys. Acta 194,* 351.
Ladner, R. C., Heidner, E. J., and Perutz, M. F. (1977), *J. Mol. Biol.* (in press).
Muirhead, H., Cox, J. M., Mazzarella, L., and Perutz, M. F. (1967), *J. Mol. Biol. 28,* 117.
Peisach, J. (1975), *Ann. N.Y. Acad. Sci. 244,* 187.
Perutz, M. F. (1970), *Nature (London) 228,* 726.
Perutz, M. F. (1976), *Br. Med. Bull. 32,* 195.
Perutz, M. F., Fersht, A. R., Simon, S. R., and Roberts, G. C. K. (1974), *Biochemistry 13,* 2174.
Rosemeyer, M. A., and Huehns, E. R. (1967), *J. Mol. Biol. 25,* 253.
Rougee, M., and Brault, D. (1975), *Biochemistry 14,* 4100.
Sono, M., Smith, P. D., McCray, J. A., and Asakura, T. (1976), *J. Biol. Chem. 251,* 1418.
Takano, T. (1977a), *J. Mol. Biol. 110,* 533.
Takano, T. (1977b), *J. Mol. Biol. 110,* 569.
Warshel, A. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 1789.
Yip, Y. K., Waks, M., and Beychok, S. (1972), *J. Biol. Chem. 247,* 7237.

Analysis of Rat Repetitive DNA Sequences†

William R. Pearson, Jung-Rung Wu, and James Bonner*

ABSTRACT: Parameters of repetitive sequence organization have been measured in the rat genome. Experiments using melting, hydroxylapatite binding, and single strand specific nuclease digestion have been used to measure the number, length, and arrangement of repeated DNA sequences. Renaturation and melting or S1 nuclease digestion of 1.0 kbp DNA fragment show about 20% of rat DNA sequences are 3000-fold repeated. Renatured duplexes from 4.0 kbp DNA fragments display two repetitive size fractions after nuclease digestion. About 60% of the repeated sequences are 0.2–0.4 kbp long while the remainder are longer than 1.5 kbp. The arrangement of the repeated sequences has been measured by hydroxylapatite fractionation of DNA fragments of varying lengths bearing a repeated sequence. Repeated DNA sequences are interspersed among 2.5 kbp long nonrepeated sequences throughout more than 70% of the rat genome. There are approximately 350 different 3000-fold short repeated sequences in the rat interspersed among 600 000 nonrepeated DNA sequences.

Evidence is accumulating that one of the mechanisms for differential gene expression is the sequence specific regulation of RNA transcription. Recent work on a variety of organisms has demonstrated an almost universal highly ordered pattern of repetitive sequence organization in DNA (see Davidson et al., 1975a, for review). The proximity of subsets of repeated sequences to transcribed and translated DNA sequences supports a regulatory function. These structural observations on sequences near functional DNA sequences suggest that repeated sequences may play an important role in DNA sequence recognition during the regulation of transcription.

Repeated DNA sequences display strikingly similar, highly ordered structures across a wide range of organisms from insects to mammals (Davidson et al., 1973a,b; Bonner et al., 1973; Davidson et al., 1975a,b; Graham et al., 1974; Angerer et al., 1975; Chamberlin et al., 1975; Goldberg et al., 1975; Schmid and Deininger, 1975; Efstratiadis et al., 1976). With a few exceptions (Manning et al., 1975; Crain et al., 1976), repeated sequences 0.2–0.4 kbp long are interspersed among 1.0–2.0 kbp nonrepeated single copy sequences in more than 65% of the DNA of all organisms studied. This organization was predicted by Britten and Davidson (1969; also Davidson and Britten, 1973) in a model which suggested that repeated DNA sequences can coordinately control the expression of adjacent single copy genes.

While the similarity of organizational patterns across the evolutionary spectrum is striking evidence for repeated DNA function in gene expression, experiments demonstrating specific subsets of repeated sequences adjacent to transcribed and translated genes provide even stronger support for the regulatory function of repeated sequences. Experiments on large nuclear transcripts in the rat (Holmes and Bonner, 1974b) and sea urchin (Smith et al., 1974) show that most large nuclear RNA is transcribed from DNA containing interspersed repeated sequences. More recent experiments suggest that a select subset of repeated sequences is near transcribed and translated sequences. Gottesfeld has reported (Gottesfeld et al., 1976) that a fraction of DNA from chromatin with increased transcriptional activity contains not only a subset of single copy DNA but also a subset of all repeated sequences. Experiments using sea urchin messenger RNA have shown 80% of the translated sequences are adjacent to repeated DNA sequences (Davidson et al., 1975b) and that these adjacent sequences are a subset of the repetitive sequence population (Klein et al., in preparation). Selected repeated sequences are adjacent to single copy DNA sequences which are transcribed into message and translated.

In this paper we describe experiments which measure the structural parameters of repeated DNA sequence organization in a mammal, the rat. We have made three basic measurements. First, the fraction of the DNA which contains repeated sequences has been determined. Second, we have measured the

† From the Division of Biology, California Institute of Technology, Pasadena, California 91125. *Received July 29, 1977.* This work was supported in part by U.S. Public Health Service Training Grant GM 00086 and in part by U.S. Public Health Service Research Grants GM 13762 and GM 20927.

0006-2960/78/0417-0051$01.00/0