Determination of rate distributions from kinetic experiments

P. J. Steinbach, K. Chu, H. Frauenfelder, J. B. Johnson, D. C. Lamb, G. U. Nienhaus, T. B. Sauke, and R. D. Young
Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 USA

ABSTRACT Rate processes in proteins are often not adequately described by simple exponential kinetics. Instead of modeling the kinetics in the time domain, it can be advantageous to perform a numerical inversion leading to a rate distribution function $f(\lambda)$. The features observed in $f(\lambda)$ (number, positions, and shapes of peaks) can then be interpreted. We discuss different numerical techniques for obtaining rate distribution functions, with special emphasis on the maximum entropy method. Examples are given for the application of these techniques to flash photolysis data of heme proteins.

INTRODUCTION

The elucidation of the relationship between the structure and function of proteins requires a detailed analysis of the kinetic features, especially as functions of external parameters such as temperature, pressure, and viscosity. Since even the smallest protein molecules are complex structures, rate processes (e.g., chemical reactions) are often complicated. Therefore, discrete exponentials do not, in general, adequately describe the observed kinetics.

In Fig. 1 we give as an example the kinetics of the binding of carbon monoxide to the separated β-chains of the mutant hemoglobin Zürich after photodissociation at 300 K (1). The fraction $N(t)$ of molecules that has not yet rebound a ligand within time t after the photolysis flash is plotted in two different ways. In Fig. 1 *a* we use a logarithmic time scale. Two different features stand out: the smooth decay in the region 10^{-9} s $< t <$ 10^{-6} s represents nonexponential rebinding kinetics. For $t >$ 10^{-6} s there is a more nearly exponential drop-off. Fig. 1 *b* displays the same data on a linear time scale. Obviously, plots that use linear time scales span at best two decades in time and are not very useful for reactions that proceed over many decades.

Any rate process may be described with a spectrum of rate coefficients λ,

$$N(t) = \int d\lambda\, h(\lambda) e^{-\lambda t}. \tag{1}$$

Here $N(t)$ is a normalized decay function, and $h(\lambda)d\lambda$ gives the probability that the process occurs with rate coefficients between λ and $\lambda + d\lambda$. As is evident from Fig. 1, rate distributions $h(\lambda)$ of nonexponential processes frequently spread over several decades in rate space. Therefore Eq. 1 is inconvenient, and it is better to introduce a rate distribution function $f(\lambda)$ on a logarithmic λ-scale. We set $h(\lambda)d\lambda = f(\lambda)d \log \lambda$ and obtain[1]

$$N(t) = \int d \log \lambda\, f(\lambda) e^{-\lambda t}. \tag{2}$$

The kinetic process is fully characterized by the time function $N(t)$ or by the rate distribution function $f(\lambda)$. To obtain a quantitative description of the kinetics, the modeling of the data can be performed either in the time domain or in the rate domain.

A parameterization in the time domain can be obtained by fitting empirical model functions to the data. This approach leads to a compression of the data to a few parameters at an early stage of the data evaluation, but details in the kinetic curve may be sacrificed. The power law,

$$N(t) = (1 + \lambda_0 t)^{-n}, \tag{3}$$

has been used to model the nonexponential rebinding of ligands to heme proteins (2). The corresponding rate distribution function is given by

$$f(\lambda) = \ln 10 \frac{(\lambda/\lambda_0)^n}{\Gamma(n)} e^{-\lambda/\lambda_0}. \tag{4}$$

Address correspondence to Dr. Nienhaus.

P. J. Steinbach's present address is National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland 20892.

J. B. Johnson's present address is Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235.

T. B. Sauke's present address is NASA Ames Research Center, Moffet Field, California 94035.

R. D. Young's permanent address is Department of Physics, Illinois State University, Normal, Illinois 61761.

[1] We measure time in (s) and (first order) rate coefficients in (s^{-1}). Thus $h(\lambda)$ has the dimension (s), and $f(\lambda)$ is dimensionless. We shall omit the units from the argument of the logarithm.

0006-3495/92/01/235/11 $2.00

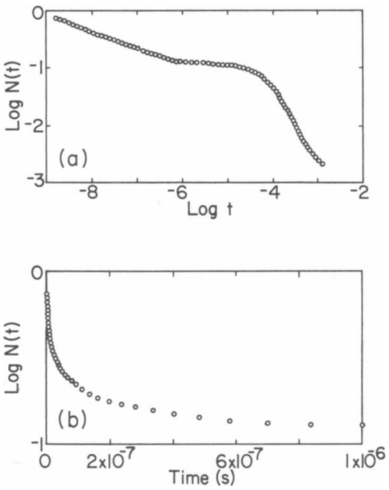

FIGURE 1 Flash photolysis kinetics of CO rebinding to separated β-chains from the mutant hemoglobin Zürich at 300 K plotted on (*a*) a logarithmic, and (*b*) a linear time axis.

The Kohlrausch law (3, 4) or stretched exponential,

$$N(t) = \exp\,[-(\lambda_0 t)^{\beta}], \tag{5}$$

is another example of an empirical model function that has been widely employed to parameterize nonexponential processes in the time domain (5–7). In contrast to the power law, this function does not have a simple analytical expression for the corresponding $f(\lambda)$ distribution (8, 9).

Instead of using model functions in the time domain, Eq. 2 can be inverted numerically to obtain the rate distribution function $f(\lambda)$. The features of $f(\lambda)$ can then be analyzed in detail. This approach has clear advantages: theoretical models usually deal with rates; it is, therefore, convenient to compare theory and experiment on the basis of the rate distribution function $f(\lambda)$. Furthermore, a good method of inversion will preserve details of the experimental data $N(t)$ in $f(\lambda)$. Unfortunately, the precise computation of the rate distribution function $f(\lambda)$ is not a straightforward task. Whereas it is easy to calculate $N(t)$ from a given $f(\lambda)$ with Eq. 2 by performing essentially a Laplace transformation, the inverse operation is ill conditioned (10). Consequently, the inverse transform of an experimental data set, which is inevitably incomplete and noisy, leads to ambiguity in $f(\lambda)$. A multitude of distribution functions will agree equally well with the experimental data. In some of the solutions, $f(\lambda)$ will assume negative values, which is physically unreasonable. Others will exhibit wild oscillations that are not justified by the experimental data. To select a single distribution $f(\lambda)$, we must impose additional constraints on $f(\lambda)$. Various techniques have been developed to obtain reasonable distributions $f(\lambda)$ (11). In this paper, we shall discuss the numerical techniques that we use to analyze flash photolysis data on heme proteins. A sophisticated inversion method employs the maximum entropy method (MEM), a data analysis technique that has found wide application in various fields, including radio astronomy (12), neutron scattering (13), fluorescence (14), and ligand rebinding (15, 16). The basic idea of the MEM is simple (17, 18): a rate distribution $f(\lambda)$ is represented by a discrete set of data $f(\lambda_j)$. The amount of uncertainty involved in the specification of $f(\lambda_j)$ is measured by the Shannon-Jaynes entropy S. By maximizing S under the constraint that the misfit statistic $\chi^2 = 1$, the MEM does not introduce spurious correlations into $f(\lambda)$ and yields the rate distribution function having the smoothest features compatible with both the experimental data and the noise. We will demonstrate the capabilities of the method by applying it to kinetic data taken with a large dynamic range [four decades in $N(t)$] over a wide range of temperatures. Subtle undulations in the rebinding curves are resolved into individual peaks describing different rebinding processes.

The inversion of the data obtained in the time domain into a distribution of rate coefficients is the first step in the kinetic analysis; yet it alone does not lead to a full understanding of the physical mechanisms that give rise to the distribution. Although the mathematical structure of Eq. 2 may suggest parallel processes observed for an inhomogeneous ensemble of molecules with different rebinding rates, the distribution $f(\lambda)$ may likewise arise from sequential processes. The dependence of $f(\lambda)$ on temperature and other external parameters gives additional information that may finally lead to a physical model describing the kinetic process.

In the present paper, we shall discuss several numerical methods used to approximate $f(\lambda)$ from $N(t)$. Particular emphasis is given to the maximum entropy method. Examples will demonstrate how these techniques are applied to the analysis of flash photolysis experiments on proteins.

NUMERICAL INVERSION TECHNIQUES

Derivative approximation

For broad rate distributions that extend smoothly over several decades, there is a simple way of getting a rough idea of $f(\lambda)$ by using the approximation (19)

$$f(\lambda) \approx -N(t)\,\frac{d\log N(t)}{d\log t}. \tag{6}$$

Eq. 6 is justified with Eq. 2,

$$-N(t)\frac{d\log N(t)}{d\log t} = -\frac{dN(t)}{\ln 10\, d\log t} = \int d\log\lambda\, f(\lambda)R(\lambda, T) \quad (7)$$

with

$$R(\lambda, t) = \lambda t e^{-\lambda t}. \quad (8)$$

If $R(\lambda, t)$ were a δ-function on a logarithmic scale, Eq. 6 would be exact. However, $R(\lambda, t)$ is an asymmetric, bell-shaped curve with a full width at half maximum of $\lambda t = 2.4$ and a maximum at $\lambda t = 1$. Therefore, the convolution of the distribution function $f(\lambda)$ with $R(\lambda, t)$ blurs sharp features. Furthermore, $f(\lambda)$ is obtained locally form the slope of $N(t)$. Numerical differentiation of noisy data results in a large scatter of $f(\lambda)$.

Fit with model functions

In this method, the actual rate distribution function $f(\lambda)$ is approximated by simple analytic functions. These functions may be based on theoretical considerations, as in the case of the Gaussian on a logarithmic λ-scale (20) or the gamma distribution (21). They can, however, also be viewed as mere parameterizations of $f(\lambda)$ with a minimal set of parameters. Multimodal distributions are represented by a sum of discrete distributions. Nonlinear least-squares fitting varies the parameters of the model function $f(\lambda)$ so that its transform with Eq. 2 gives the best fit to the experimental $N(t)$ curve. Here the derivative approximation is useful to estimate the number of discrete distributions and the initial parameters of the fit. In our work, we frequently parameterize $f(\lambda)$ with a sum of Gaussians on a logarithmic scale:

$$f(\lambda) = \sum_i a_i \exp\left[-b_i(\log\lambda - \log\lambda_i^c)^2\right]. \quad (9)$$

The areas, widths, and center positions of the Gaussians are determined by a_i, b_i, and λ_i^c, respectively.

Maximum entropy method

The MEM is quite general and can be used whenever the data one measures represent some transform of the function of interest. From the many solutions that fit the inevitably noisy and incomplete data equally well, the MEM selects a unique distribution free of spurious correlations.

From the kinetics experiment we obtain a set of data points $\mathcal{T}^{\exp}(t_i)$. We assume that the $\mathcal{T}^{\exp}(t_i)$ have Gaussian errors with standard deviations $\sigma(t_i)$. We divide the log λ axis into a large number of bins with equal spacing on a logarithmic λ-scale. The continuous $f(\lambda)$ distribution is then discretized in J values $f(\lambda_j)$. In our application, where the data are linear in $f(\lambda)$, the best fit to I data points $\mathcal{T}^{\exp}(t_i)$ is obtained by maximizing the Bayesian posterior probability distribution that is proportional to (13, 17)

$$\mathcal{P} = \exp(\alpha S - I\chi^2/2), \quad (10)$$

where α is a dimensional coefficient that is not known initially. The χ^2-statistic measures how well a trial set $\mathcal{T}^{\text{fit}}(t_i)$ fits the experimental data set $\mathcal{T}^{\exp}(t_i)$,

$$\chi^2 = \frac{1}{I}\sum_{i=1}^{I}\left[\frac{\mathcal{T}^{\text{fit}}(t_i) - \mathcal{T}^{\exp}(t_i)}{\sigma(t_i)}\right]^2. \quad (11)$$

The Shannon-Jaynes entropy S (22) is defined by

$$S = -\sum_{j=1}^{J} f(\lambda_j)\left[\ln\frac{f(\lambda_j)}{F(\lambda_j)} - 1\right]. \quad (12)$$

The function $F(\lambda_j)$ incorporates any prior knowledge we have about the form of $f(\lambda_j)$ into the fit. In the absence of any data, maximizing the entropy yields $f(\lambda_j) = F(\lambda_j)$, as can be seen by taking the derivative $\partial S/\partial f(\lambda_j)$, setting it equal to zero and solving for $f(\lambda_j)$. In the absence of any prior knowledge we choose a constant F distribution so as not to introduce any structure into f not warranted by the data. The distribution $f(\lambda_j)$ is quite insensitive to the constant value chosen for F because of the high quality of the experimental data. For all the fits shown in this paper $F(\lambda_j)$ was set equal to 10^{-2} or 10^{-3}.

Finding the maximum in $\mathcal{P}$ is equivalent to finding the maximum of $\ln\mathcal{P} = \alpha S - I\chi^2/2$. Different algorithms have been applied to this task (for a comparison see reference 18). Following the work of Gull and Daniell (12), we use a simple algorithm that maximizes the function

$$Q = S - L_0\chi^2. \quad (13)$$

L_0 is a Lagrange multiplier that is chosen such that $\chi^2 \cong 1$. The fit to the data at each time, t_i, is calculated in transmittance space[2]

$$\mathcal{T}^{\text{fit}}(t_i) = 10^{-\Delta A_{\max}\sum_{j=1}^{J}\Delta\log\lambda_j f(\lambda_j)\exp(-\lambda_j t_i)}. \quad (14)$$

where $\Delta A_{\max}$ is the maximum change in the absorbance of the sample that occurs immediately after photolysis, and $\Delta\log\lambda_j$ is the spacing between neighboring $f(\lambda_j)$ points. Setting $\partial Q/\partial f(\lambda_j)$ to zero we obtain

$$f(\lambda_j) = F(\lambda_j)\exp\left(L\sum_{i=1}^{I}\mathcal{T}^{\text{fit}}(t_i)\frac{\mathcal{T}^{\text{fit}}(t_i) - \mathcal{T}^{\exp}(t_i)}{\sigma^2(t_i)}e^{-\lambda_j t_i}\right). \quad (15)$$

A number of coefficients are lumped together into L such that $L = 2\ln 10\, L_0\Delta A_{\max}\Delta\log\lambda_j/I$. The adopted sign convention requires that L be positive for absorbing signals ($\Delta A_{\max} > 0$) and negative for bleaching signals ($\Delta A_{\max} < 0$). Eq. 15 represents J coupled nonlinear equations. For all the data in this paper J equalled 100. The task of solving 100 coupled nonlinear equations is done by iteration. Because of the nonlinearity of the equations, successive iterations are combined, where typically $C = \leq 0.1\%$ of the current iteration was added to the previous iteration of $f(\lambda_j)$. L and the weight C are adjusted so that χ^2 approaches a value acceptably close to 1 while a convergence test approaches 0:

$$\text{TEST} = \frac{1}{2}\left|\frac{\nabla S}{|\nabla S|} - \frac{\nabla\chi^2}{|\nabla\chi^2|}\right|^2. \quad (16)$$

The gradients in Eq. 16 are evaluated with respect to the $f(\lambda_j)$ in the $J = 100$-dimensional space. For a given value of χ^2, *TEST* is zero at the unique point in $f(\lambda_j)$-space where the gradients S and χ^2 are parallel; there is no component of ∇S along the surface of constant χ^2. The situation for the case of only two $f(\lambda_j)$ is illustrated in Fig. 2 where

---

[2] The transmittance is related to the kinetics signal $N(t)$ by: $\mathcal{T} = 10^{-\Delta A_{\max}N(t)}$ (see Applications). Because the noise is assumed Gaussian in transmittance, it is necessary to fit in $\mathcal{T}$-space rather than in $N(t)$-space.

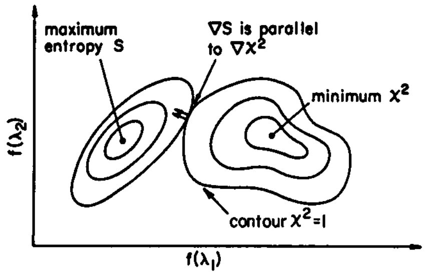

FIGURE 2 A contour plot of entropy, S, and the χ^2-statistic for a two-dimensional $f(\lambda)$. The maximum entropy solution corresponds to the point where the gradients of S and χ^2 are parallel, and $\chi^2 \approx 1$.

contour plots of S and χ^2 are shown. The maximum entropy solution corresponds to the point on the contour $\chi^2 = 1$ where S is maximal. Because the statistical uncertainty, $\sigma(t_i)$, is not known precisely, the final χ^2 is generally chosen somewhat larger than unity in order to be conservative. Decreasing χ^2 further tends to sharpen peaks with little change of peak positions and areas. Experience also indicates that maps producing a *TEST* greater than 0.001 may suffer from unacceptable artifacts, i.e., structure not justified by the signal-to-noise ratio of the data. Lavalette et al. (16) recently suggested the addition of a heavily weighted data point at early times to enforce the initial condition $N(0) = 1$. We have chosen not to do so because the normalization factor, ΔA_{max}, is not known precisely. Our $f(\lambda)$ distributions, therefore, do not normalize to unity, and we do not interpret the feature in $f(\lambda)$ that is related to the shortest times too seriously.

Our MEM algorithm has been tested using synthetic data with signal-to-noise characteristics similar to that found in our flash photolysis data. First a $f(\lambda_j)$ distribution, shown as points in Fig. 3 *c*, was generated and used to calculate transmittance data by Eq. 14. Gaussian noise was added to the data. The maximum entropy algorithm produced the distribution shown as the solid line in Fig. 3 *c*. The fit is shown as a solid line in Fig. 3 *a*. The MEM does an excellent job of recovering $f(\lambda_j)$, particularly peak positions, given the limited, noisy data. The residuals are shown in Fig. 3 *b*. (They are defined as $[\mathcal{T}^{exp}(t_i) - \mathcal{T}^{fit}(t_i)]/\sigma(t_i)$ and should possess a standard normal distribution. Thus 68% of the data points should fall with ± 1, and 95% should fall within ± 2.) As expected, where the signal-to-noise ratio is poor (at large values of λ), the $f(\lambda_j)$ distribution is less accurately reproduced.

A second test shows the influence of noise on the sharpness of the features in the rate distribution recovered by the MEM. The most extreme case for $f(\lambda)$ is a δ-function on a logarithmic scale, corresponding to an exponential $N(t)$. We simulated the transmittance signal starting from 0.1 at $t = 0$ and approaching 1 at long times. Different amounts of Gaussian noise (independent of time) were added. Fig. 4 shows that with large amounts of noise the obtained rate distributions are rather broad. As the signal-to-noise ratio increases, the rate distribution approaches the δ-function.

These two tests clearly show that the MEM does not introduce correlations that are not inherent in the data. Broad rate distributions are obtained from limited, noisy data (see Fig. 4) because the MEM is maximally noncommittal with regard to unavailable information. In other words, a distribution is "presumed smooth until proven spikey."

Coupled rate processes

The previous subsections addressed numerical inversion techniques for single rate processes. Often, the kinetic schemes are more complicated, involving coupled sequential and parallel processes. As an example we consider the following kinetic scheme:

$$\begin{array}{c} B_0 \rightarrow A_0 \\ \upharpoonleft\downharpoonright \\ B_1 \rightarrow A_1 \end{array} \tag{17}$$

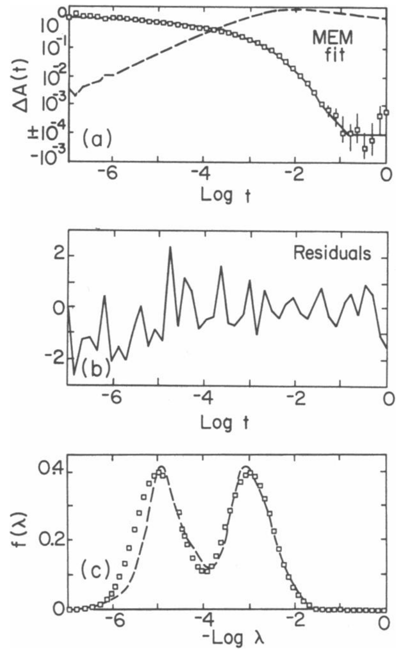

FIGURE 3 Test of the maximum entropy method. (*a*) Squares, synthetic data with realistic signal-to-noise ratio (S/N) generated from the rate distribution in *c*; solid line, fit with the maximum entropy method; dashed line, $10^{-3} \times$ (S/N). (*b*) Residuals of the fit (in transmittance). (*c*) Squares, rate distribution used to generate the data in *a*; dashed line, rate distribution recovered by the MEM from the data in *a*.

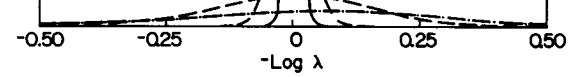

FIGURE 4 Test of the maximum entropy method. Synthetic data simulating an exponential rate process were generated by calculating a transmittance signal rising from $\mathcal{T} = 0.1$ at $t = 0$ to 1 at long times. Gaussian noise, independent of time, was added with standard deviations of 10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4}. The corresponding rate distribution functions $f(\lambda)$ are shown. As the signal-to-noise ratio increases, $f(\lambda)$ approaches a δ-function.

The processes $B \rightarrow A$ are coupled owing to an additional exchange between A_0 and A_1. This scheme can be modelled with the following two equations:

$$N_0(t) = \int d \log \lambda [f_0(\lambda) - f_e(\lambda)] \exp(-\lambda t), \tag{18}$$

$$N_1(t) = \int d \log \lambda [f_1(\lambda) + \alpha f_e(\lambda)] \exp(-\lambda t). \tag{19}$$

The response functions $N_0(t)$ and $N_1(t)$ decay from 1 to 0 during the time course of the process and represent the populations that are to move into states A_0 and A_1. The rate distributions for the processes $B \rightarrow A$ are denoted by $f_0(\lambda)$ and $f_1(\lambda)$; and $f_e(\lambda)$ represents the exchange process. The signs in front of $f_e(\lambda)$ are chosen such that $f_e(\lambda)$ is positive when a net flow occurs from A_0 to A_1. The coefficient α gives the ratio of the two species A_0, A_1 in equilibrium. An example for such an exchange process will be discussed later.

To perform the inversion of Eqs. 18 and 19 with the maximum entropy method, we cannot apply the entropy definition of Eq. 12 directly. We generalize Eq. 12 by defining

$$S = - \sum_{n=0,1} \sum_j f_n(\lambda_j) \left[\log \left(\frac{f_n(\lambda_j)}{F_n(\lambda_j)} \right) - 1 \right] - \sum_j f_e(\lambda_j) \left[\log \left(\frac{f_e(\lambda_j)}{F_e(\lambda_j)} \right) - 1 \right]. \tag{20}$$

This approach yields accurate results for $(f_0 - f_e)$ and $(f_1 + \alpha f_e)$, but the individual distributions f_0, f_1, and f_e cannot be extracted unambiguously for the following reasons: (*a*) The entropy is maximized by assigning the smallest possible exchange $A_0 \rightarrow A_1$; thus the exchange peak determined by the MEM may be smaller than the true peak. (*b*) The entropy is maximized by making the region where $f_0(\lambda)$ and $f_e(\lambda)$ are both nonzero very small. The true distribution may have some overlap of $f_e(\lambda)$ and $f_0(\lambda)$ giving rise to peaks in $f_e(\lambda)$ that may be shifted in position relative to the true position.

As an alternative to the MEM approach we can also use model functions for f_0, f_1, and f_e as described previously. For example, f_0 and f_1 can be represented by a small number of Gaussians on a logarithmic scale and the exchange distribution by an additional Gaussian:

$$N_0(t) = \int d \log \lambda \cdot \left[\sum_{j=1}^{J} g(\lambda; a_{0j}, b_{0j}, \lambda_{0j}^c) - g(\lambda; a_e, b_e, \lambda_e^c) \right] \exp(-\lambda t), \tag{21}$$

$$N_1(t) = \int d \log \lambda \cdot \left[\sum_{j=1}^{J} g(\lambda; a_{1j}, b_{1j}, \lambda_{1j}^c) + \alpha g(\lambda; a_e, b_e, \lambda_e^c) \right] \exp(-\lambda t). \tag{22}$$

The parameters of the Gaussians are varied to find the best fit to the data for $N_0(t)$ and $N_1(t)$ using a Levenberg-Marquardt nonlinear least-squares algorithm.

APPLICATIONS

Experimental approach

The kinetics of the binding reaction of carbon monoxide to sperm whale myoglobin after flash photolysis,

$$\mathrm{Mb} + \mathrm{CO} \rightleftharpoons \mathrm{MbCO} \tag{23}$$

has been investigated in great detail. The covalent bond between the heme iron and the CO ligand is broken with unit quantum efficiency on absorption of a visible photon (23). The subsequent ligand rebinding is accompanied by changes in the absorption spectrum of the sample; therefore it can be followed in time by monitoring the intensity of light, $\mathscr{I}$, passing through the sample before and after the photolyzing laser pulse arrives at time $t = 0$. The transmittance $\mathscr{T}(t_i, T)$ at a temperature T is recorded for a set of time t_i:

$$\mathscr{T}(t_i, T) \equiv \frac{\mathscr{I}(t_i, T)}{\mathscr{I}(0^-, T)} = 10^{-\Delta A(t_i, T)} \tag{24}$$

where $\Delta A(t_i, T) = A(t_i, T) - A(0^-, T)$ is the change in the absorbance of the sample. The fraction of proteins yet to rebind a ligand is given by $N(t_i, T) = \Delta A(t_i, T) / \Delta A_{\mathrm{max}}(T)$, where $\Delta A_{\mathrm{max}}(T)$ is the initial absorbance change upon complete photolysis. We use two different flash photolysis systems to measure these absorbance changes; one monitors the absorbance of the heme group in the visible, and the other monitors the absorbance in the infrared $\sim 1{,}950\ \mathrm{cm}^{-1}$ ($\approx 5\ \mu\mathrm{m}$) due to the stretch vibration of the CO molecule.

In the visible, binding of Mb and CO in a 75% glycerol/water solvent at pH 6.8 was measured in the Soret band at 440 nm. The protein concentration was $\sim 10\ \mu\mathrm{M}$ and the CO pressure either 0.05 or 1 bar. The MbCO samples were dissociated with a 30-ns, 530-nm pulse generated by a frequency-doubled, Q-switched, Nd-doped phosphate glass laser. Two dichroic mirrors preferentially guided the 530-nm light to the sample while failing to reflect the 1,060-nm light. The energy of the green pulse reaching the sample was ~ 250 mJ. A 24-V tungsten lamp (Oriel Corp., Stratford, CT) provided the monitoring beam. Before reaching the sample, the monitor light passed through a monochromator set at 440 nm. A 440-nm interference filter and two other blue glass filters between the sample and the photomultiplier tube (PMT) reduced the amount of 530-nm light that scattered into the PMT, degrading the signal at short times. The signal from the PMT (model R928; Hamamatsu Corp., Middlesex, NJ) was amplified and recorded by both a 10-ns digitizer (model TR8818, LeCroy, Chestnut Ridge, NY) and a home-built 1-μs digitizer (24) with logarithmic time base en route to a DEC μPDP 11/73.

Flash photolysis experiments with monitoring in the infrared were performed on MbCO samples in a 75% glycerol/water solution at pH 5.7. The protein concentration was ~ 15 mM. Photodissociation was achieved using a 300-ns, 530-nm pulse (broad band) from a dye laser (model DL2100C; Phase-R, New Durham, NH). The source for the monitoring beam was a lead salt laser diode (Laser Photonics, Analytics Div., Bedford MA) that was tunable in the range between 1,900 and 2,000 cm^{-1}. The laser diode and the sample were kept at cryogenic temperatures with a closed-cycle helium refrigerator (CTI-Cryogenics Div., Helix Technology Corp., Waltham, MA). The IR beam was monitored with a photovoltaic InSb detector (Infrared Associates, Cranbury, NJ). The output signal was amplified and digitized by our home-built logarithmic digitizer.

Because the inversion of the data using the maximum entropy method depends on the estimation of experimental uncertainties, we outline this procedure. The logarithmic digitizer uses two analog-to-digital converters (12 bit, 2 μs) that take turns measuring the detector signal, thus acquiring a voltage sample every microsecond. A number of voltage samples, always a power of two, is averaged together in real time by the digitizer. P such averages are performed during a period in which the time after the start pulse doubles. Then the number of voltage samples to be averaged is doubled, and the next P averages are calculated. Each of these cycles is called a doubling period. The extent to which the data are averaged is changed by varying the parameter P, called points per doubling. Typically, we use $P = 12$. This method of data collection produces approximately equally spaced points on a logarithmic time axis.

Normally, several files containing P points per doubling are averaged down to a single file containing M points per doubling with $1 \leq M \leq P$. M is chosen so as to maximize signal averaging while maintaining sufficient time resolution. Typical values of M range from 1 to 4. Since the number of samples averaged together by the logarithmic digitizer changes only at the doubling period boundaries, each doubling period has its own intrinsic signal-to-noise ratio and is treated independently. A line is fitted to all points in a doubling period, giving a calculated transmittance $\mathscr{T}^{\text{lin}}(t_i)$. The scatter about the line is used to estimate the standard deviation in the data. The uncertainty is estimated as $\sigma(t_i) \approx \sqrt{M}\sqrt{\Sigma_{q=i}^{Q}[\mathscr{T}^{\text{exp}}(t_i) - \mathscr{T}^{\text{lin}}(t_i)]^2/[P(P-2)]}$. Here, Q is the total number of points from all data files that are averaged together to give M final data points in a particular doubling period.

Model for ligand binding

The simplest kinetic scheme that describes the basic features of the binding reaction of ligands to heme proteins involves a single adiabatic energy surface for the protein–ligand system that has three sequential wells along the reaction coordinate: the bound state A, the pocket state B, and the solvent state S. Neglecting distributed barriers and thermal dissociation of the ligand from the heme ($k_{AB} \approx 0$), we obtain the system depicted in Fig. 5 (25).

Initially, the ligand is bound to the heme iron. A short laser pulse breaks the Fe-CO bond and the ligand moves into the heme pocket. There the ligand has two choices: either it rebinds internally (process I) or it escapes into the solvent. Subsequently, a ligand from the solvent will enter the protein molecule and bind to the iron. We call this process S.

Protein molecules are complex systems that can assume a large numbers of conformational substates (CS) (2, 26–28). CS vary slightly in their three-dimensional structures and generally function with different rates. Consequently, the rate coefficients for the processes sketched in Fig. 5 are not single valued but are characterized by distributions. Experiments suggest that the CS are arranged in a hierarchy of tiers (28–31). Sperm whale MbCO possesses three substates in the highest tier (CS0). These substates are called A substates and characterized by different infrared absorption bands (~ 1,950 cm^{-1}) because of the stretch mode of the bound CO ligand. Both x-ray crystallography (32) and linear dichroism (33) have resolved structural inhomogeneity in the orientation of the bound CO molecule.

Monitoring in the infrared allows us to measure the binding kinetics of the three CS0 separately. They perform the same function but with different rates: A_0 rebinds CO faster than A_1, which rebinds faster than A_3 (34). The geminate rebinding to each of the A substates at low temperatures is nonexponential and is characterized by a distributed enthalpy barrier. This inhomogeneity is explained by conformational substates of tier 1, CS1. Enthalpy distributions and preexponentials of the three A substates have been measured in the infrared using two different techniques: Ansari et al. (34) monitored the transmittance at a fixed temperature and a single infrared wavelength with a microsecond flash-

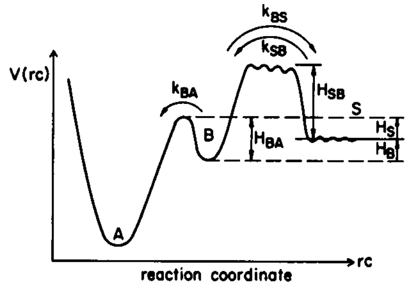

FIGURE 5 The 3-well model. The protein–ligand system may occupy one of three states along the reaction coordinate, *rc*. The ligand may be bound at the heme iron (*A*), be dissociated in the pocket (*B*), or be in the solvent (*S*).

photolysis system. Berendzen and Braunstein (35) used temperature-derivative spectroscopy to monitor rebinding across the entire A substate spectrum while ramping the temperature.

Low-temperature rebinding monitored in the Soret band

At temperatures below 200 K, the ligand cannot leave the protein molecule and rebinds via the internal process I. The rebinding rate is governed by an enthalpic barrier of height H_{BA}. The reaction occurs either by quantum-mechanical tunneling (36, 37) (dominant below ~50 K in MbCO) or by thermal activation. In the latter case, the rate coefficient $\lambda = k_{BA}(H_{BA}, T)$ is given by the Arrhenius relation (1):

$$k_{BA}(H_{BA}, T) = A_{BA}(T/T_0)e^{-H_{BA}/RT}, \tag{25}$$

where $A_{BA}(T/T_0)$ is the frequency factor and T_0 is set to 100 K. For a unique value of the preexponential and the enthalpic barrier we expect exponential rebinding according to

$$N(t) = e^{-k_{BA}(H_{BA},T)t}. \tag{26}$$

In contrast, process I is clearly nonexponential. We explain this observation by postulating that the protein molecules exist in many conformational substates with different rebinding barriers H_{BA}. Below 160 K, each molecule is frozen in a particular substate with a certain barrier height H_{BA} on the time scale of the experiment. Instead of describing the kinetics with rate distribution functions $f(\lambda)$ as in Eq. 2, we can model the reaction with a single, temperature-independent distribution $g(H_{BA})$ of enthalpic barriers,

$$N(t, T) = \int_0^\infty dH_{BA}\, g(H_{BA})e^{-k_{BA}(H_{BA},T)t}. \tag{27}$$

The temperature dependence of $N(t, T)$ enters solely through the Arrhenius relation for the rate coefficient. For the activation enthalpy distribution, $g(H_{BA})$, model functions are usually employed (20, 21). The enthalpy distribution and the Arrhenius prefactor A_{BA} can be determined by fitting a single $g(H_{BA})$ to a series of rebinding curves taken at various temperatures.

The maximum entropy method offers a different strategy of analyzing the barrier distribution. The rate distribution functions $f(\lambda) \equiv f(\lambda, T)$ can be determined individually at each temperature and then converted into enthalpy distributions using the relations

$$H_{BA} = RT \ln 10 \left(\log \frac{A_{BA}T}{T_0} - \log \lambda\right), \tag{28}$$

$$g(H_{BA}, T) = f(\lambda, T)\frac{d \log \lambda}{dH_{BA}} = \frac{f(\lambda, T)}{RT \ln 10}. \tag{29}$$

We have added the variable T in $g(H_{BA}, T)$ to indicate that it is determined separately at each temperature.

Fig. 6 shows the $g(H_{BA}, T)$ distributions obtained by the MEM for the temperatures 60, 70, 80, and 90 K. They are in good agreement with each other and support the temperature independence of $g(H_{BA})$. We also plot the gamma distribution that was obtained by fitting all kinetic traces between 60 and 160 K. It gives a good overall representation of the $g(H_{BA})$ distribution, especially the high-enthalpy tail, but details differ. The MEM distribution is much more symmetric than the gamma distribution. The major part can be described by a Gaussian with a peak enthalpy of 10.5 kJ/mol and a width of 6 kJ/mol.

The systematic deviation from a Gaussian at high enthalpies is explained by the A_3 substate, which has a peak enthalpy of ~18 kJ/mol (34, 35). However, the kinetics monitored in the visible are in fact the sum of the contributions of the three A substates having different preexponentials A_{BA} and enthalpy distributions $g(H_{BA})$ (25). For the calculation of the $g(H_{BA}, T)$ distributions, we used an overall $A_{BA} = 10^{8.8}$ s^{-1} that was determined from fitting a gamma distribution to all the kinetic traces between 60 and 160 K. The simplified treatment of the Soret kinetics using Eqs. 25 and 27 introduces errors. It nevertheless is a good approximation because A_0 and A_1 have similar preexponentials, $10^{8.7}$ and $10^{8.9}$ s^{-1}, respectively, and because they account for ~90% of the population at pH 6.8. Using $A_{BA} = 10^{8.8}$

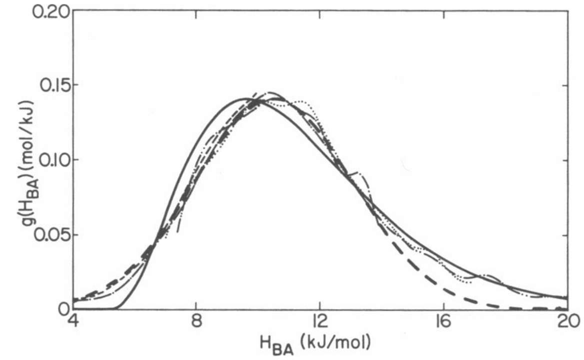

FIGURE 6 Enthalpy barrier distributions $g(H_{BA}, T)$ calculated from the rebinding kinetics of MbCO at 60 K (– – –), 70 K (— · —), 80 K (······), and 90 K (— ·· —), using the MEM. The solid line represents the $g(H_{BA})$ that was obtained from a least-squares fit of a gamma distribution to rebinding data betwen 60 and 160 K (25). The heavy dashed line is a Gaussian fit to the $g(H_{BA}, T)$ distributions obtained with the MEM.

s^{-1}, instead of the actual value $10^{10.4}$ s^{-1} for A_3 (25), results in an apparent peak enthalpy of 15–16 kJ/mol.

Rebinding above 200 K monitored in the Soret band

Above 200 K, the ligands can surmount the barrier between B and S and escape from the protein molecules. The fraction of proteins that rebind subsequently via process S increases with temperature. Fig. 7 *a* shows a rebinding curve of MbCO monitored in the Soret band at 250 K. Process I is nonexponential and extends from the earliest times to $\sim 10^{-3}$ s. Process S is close to exponential and responsible for the rapid drop-off around 10^{-2} s. In Fig. 7 *b–d*, we show the rate distribution functions $f(\lambda)$ that were obtained with the numerical inversion techniques presented in previous subsections.

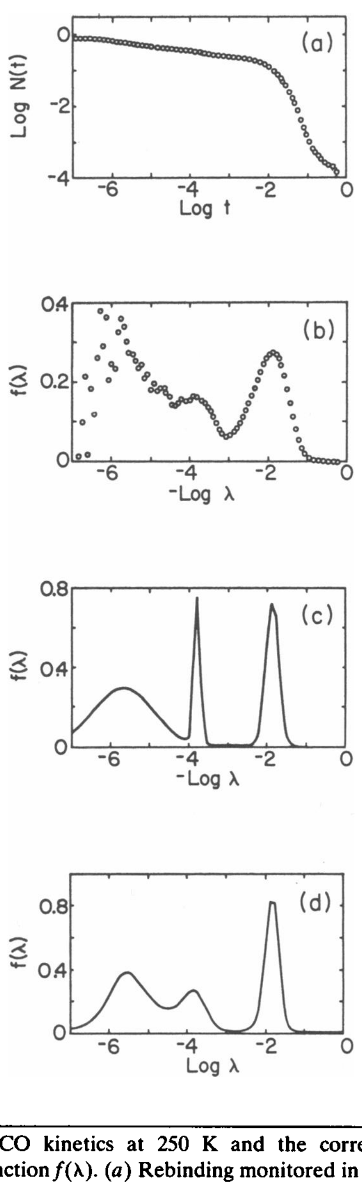

FIGURE 7 MbCO kinetics at 250 K and the corresponding rate distribution function $f(\lambda)$. (*a*) Rebinding monitored in the Soret band at 440 nm. The rate distribution function $f(\lambda)$ was computed using (*b*) the derivative method, (*c*) a fit by Gaussians on a logarithmic scale, and (*d*) the maximum entropy method.

Fig. 7 *b* shows $f(\lambda)$ obtained with the derivative approximation (Eq. 6). It has large scatter at early times where the statistical error of the data is large. Although the method involves an intrinsic blurring of sharp features, three maxima in the rate distribution function are easily distinguished. Using these data as start parameters, a fit with three log-Gaussians was performed (Eq. 9). The resulting $f(\lambda)$ is shown in Fig. 7 *c* and has, as expected, narrower peaks than the $f(\lambda)$ obtained with the derivative approximation. The $f(\lambda)$ calculated with the maximum entropy method (Fig. 7 *d*) also shows three maxima. The middle peak is markedly broader than for the Gaussian fit. While the Gaussian fit chooses the three Gaussians most compatible with the experimental data, the MEM imposes no number or shapes of peaks, but rather yields the smoothest, broadest features compatible with the data.

Fig. 8 shows rate distributions $f(\lambda)$ obtained with the MEM from flash photolysis experiments at 240, 250, and

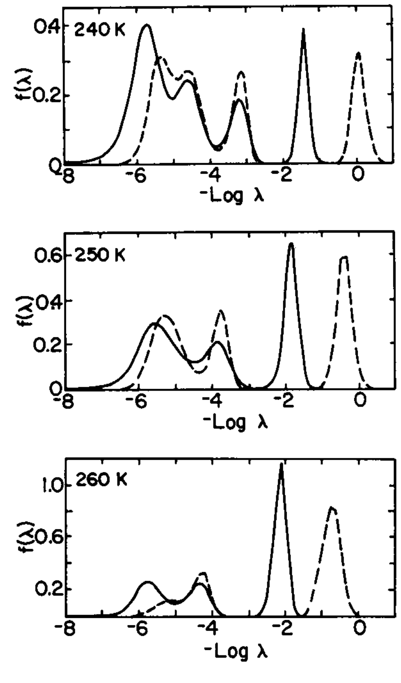

FIGURE 8 High-temperature features resolved by the MEM. Rate distributions for rebinding measured in the Soret band at 240, 250, and 260 K. Solid lines, MbCO sample with CO pressure of 1 bar; dashed lines, CO pressure 0.05 bar. The fastest kinetic process is not recovered well because of the limited time range and uncertainties in the normalization factor ΔA_{max}.

260 K. Two MbCO samples with different CO concentrations were used. It is obvious that only the position of the rightmost peak depends on the CO concentration: for samples with higher CO concentration this peak shifts to the left. Therefore, it is attributed to rebinding from the solvent, process S. At first glance, the time dependence of process I at 250 K resembles a straight line in the log $N(t)$ versus log t plot, Fig. 7 *a*, which is described by a power law. The corresponding rate distribution is unimodal (see Eqs. 3 and 4). In contrast, the rate distribution function clearly reveals two peaks for process I at 250 K. These are reflected in the time domain as subtle undulations on the $N(t)$ curve. These undulations have already been noticed by Austin et al. in 1975 (2), who attributed them to multiple wells along the reaction coordinate. Application of the MEM to ligand rebinding data at 5-K intervals reveals three peaks between 230 and 245 K in addition to the peak for process S. At $T \geq 250$ K, the two peaks with the larger rate coefficients merge into a single peak.

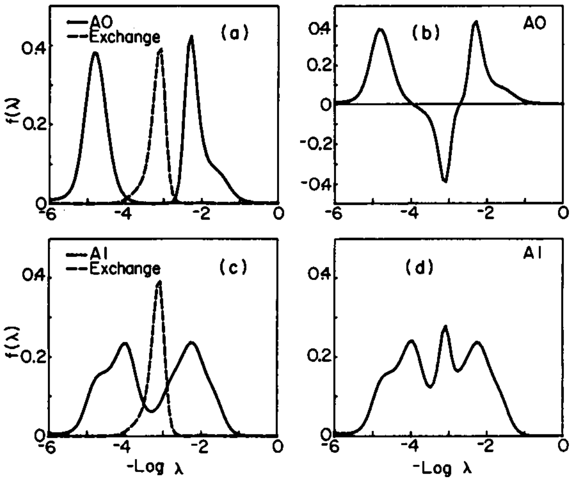

FIGURE 10 Maximum entropy fits at 250 K for the coupled rebinding process of A_0 and A_1 showing $f_0(\lambda)$, $f_1(\lambda)$, and the exchange distribution, $f_c(\lambda)$. (*a*) $f_0(\lambda)$ and $f_c(\lambda)$; (*b*) $f_0(\lambda) - f_c(\lambda)$; (*c*) $f_1(\lambda)$ and $f_c(\lambda)$; and (*d*) $f_1(\lambda) + \alpha f_c(\lambda)$.

Flash photolysis with monitoring in the CO stretch bands

In Fig. 9 we show the rebinding signal of A_0 and A_1 at 250 K. (In this sample, which has a pH of 5.7, only a minor fraction of molecules are in the A_3 substate. For simplicity, we shall neglect A_3 in the following discussion.) While $N_1(t)$ decays monotonically, $N_0(t)$ decays up to 10^{-4} s after the photolysis flash, then it increases again before decaying finally. This novel behavior can be explained by invoking an exchange between A_0 and A_1, as previously discussed. At 250 K, the A substates are in dynamic equilibrium. After photodissociation, ligand rebinding to A_0 is faster than to A_1. Therefore, A_0 will become overpopulated with respect to equilibrium. This nonequilibrium situation prevails for times shorter than the inverse rates for thermal fluctuations between A_0 and A_1. At longer times, these transitions establish equilibrium between A_0 and A_1 and lead to identical kinetics for A_0 and A_1.

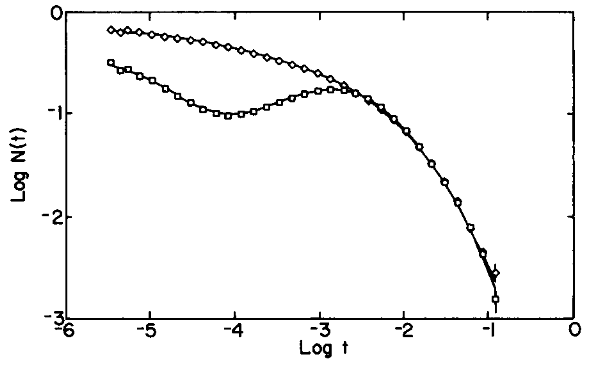

FIGURE 9 Rebinding at 250 K of MbCO monitored in the A_0 (1,966 cm^{-1}) and A_1 (1,947 cm^{-1}) CO stretch bands. The solid lines give the fits obtained with the MEM.

The inversion of the kinetic traces in Fig. 10 was performed using the maximum entropy algorithm for coupled rate processes already presented. Three peaks stand out for both A_0 and A_1. The dashed peaks in Fig. 10, *a* and *c*, correspond to the exchange process. The peaks to the left with rate coefficients around 10^5 s^{-1} are the internal processes, while the peak for the solvent process is located around 10^2 s^{-1}. Essentially identical results were obtained when fitting the data with parameterized distributions according to Eqs. 21 and 22.

CONCLUSIONS

Kinetic processes in proteins are often not adequately described by a small number of exponential components. When analyzing the data, it is advantageous to transform the experimental data numerically from the time domain into rate space. Different techniques are available to perform this task. The inversion technique employing the maximum entropy method is particularly useful, because the resulting rate distribution function $f(\lambda)$ has the minimal features compatible with both the data and the statistical errors. No additional correlations are introduced. The physical interpretation of $f(\lambda)$, therefore, needs to consider only features that are warranted by the experimental data. A note of caution must be added to these remarks. No technique can derive unambiguous conclusions from noisy data. As the tests discussed previously show, data of high quality are required to distinguish a distribution from a "single exponential." A multifaceted approach is consequently recommended for analyzing complex reaction phenomena as in biomolecules. Features that emerge essentially unchanged when employing different techniques, such as fits with parameterized distributions, the maximum entropy method, and global (e.g., multiple temperature) models, can then be used to elucidate the physics behind the observed kinetic processes.

We thank Alastair K. Livesey for his helpful discussions of the maximum entropy method, and Joel Berendzen, David Ehrenstein, Imre Kovacs, Judith R. Mourant, Pal Ormos, Robert Philipp, and Aihua Xie for their advice and criticism. This work was supported in part by the National Science Foundation (Grant DMB87-16476), the National Institutes of Health (Grants GM-18051 and GM-32455), and the Office of Naval Research (Grant N00014-89-J-1300). G. U. Nienhaus thanks the Alexander von Humboldt Foundation for a Feodor Lynen Fellowship.

Received for publication 29 July 1991 and in final form 9 September 1991.

REFERENCES

1. Dlott, D. D., H. Frauenfelder, P. Langer, H. Roder, and E. E. Dilorio. 1985. Nanosecond flash photolysis study of carbon monoxide binding to the β chain of hemoglobin Zürich [β63(E7)His→Arg] *Proc. Natl. Acad. Sci. USA.* 80:6239–6243.
2. Austin, R. H., K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. 1975. Dynamics of ligand binding to myoglobin. *Biochemistry.* 14:5355–5373.
3. Kohlrausch, R. 1854. Theorie des elektrischen Rückstandes der Leidner Flasche. *Poggendorff's Annalen der Physik und Chemie.* 91:56–82, 179–214.
4. Kohlrausch, F. 1863. Ueber die elastische Nachwirkung bei der Torsion. *Poggendorff's Annalen der Physik und Chemie.* 119:337–368.
5. Williams, W., and D. C. Watts. 1970. Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function. *Transactions of the Faraday Society.* 66:80–85.
6. Iben, I. E. T., D. Braunstein, W. Doster, et al. 1989. Glassy behavior of a protein. *Phys. Rev. Lett.* 62:1916–1919.
7. Scher, H., M. F. Shlesinger, and J. T. Bendler. 1991. Time-scale invariance in transport and relaxation. *Physics Today.* 44(January):26–34.
8. Lindsey, C. P., and G. D. Patterson. 1980. Detailed comparison of the Williams-Watts and Cole-Davidson functions. *J. Chem. Phys.* 73:3348–3357.
9. Bendler, J. T. 1984. Lévy (stable) probability densities and mechanical relaxation in polymers. *Journal of Statistical Physics.* 36:625–637.
10. McWirther, J. G., and E. R. Pike. 1978. On the numerical inversion of the Laplace transform and Fredholm integral equations of the first kind. *J. Phys. A Math. Gen.* 11:1297–1745.
11. Provencher, S. W. 1979. Inverse problems in polymer characterization: Direct analysis of polydispersity with photon correlation spectroscopy. *Makromol. Chem.* 180:201–209.
12. Gull, S. F., and G. J. Daniell. 1978. Image reconstruction from incomplete and noisy data. *Nature (Lond.).* 272:686–690.
13. Sivia, D. S. 1990. Bayesian inductive inference maximum entropy and neutron scattering. *Los Alamos Science.* Summer 1990:180–206.
14. Livesey, A. K., and J. C. Brochon. 1987. Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. *Biophys. J.* 52:693–706.
15. Steinbach, P. J., H. Frauenfelder, J. B. Johnson, and T. B. Sauke. 1989. CO rebinding to sperm whale myoglobin: mapping the A states into the Soret. *Biophys. J.* 55:565*a*. (Abstr.)
16. Lavalette, D., C. Tetreau, J.-C. Brochon, and A. Livesey. 1991. Conformational fluctuations and protein reactivity: determination of the rate constant spectrum and consequences in elementary biochemical processes. *Eur. J. Biochem.* 1956:591–598.
17. Livesey, A. K., and J. Skilling. 1985. Maximum entropy theory. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* A41:113–122.
18. Skilling, J., and R. K. Bryan. 1984. Maximum entropy image reconstruction: general algorithm. *Monthly Notices of the Royal Astronomical Society.* 211:111–124.
19. Primak, W. 1955. Kinetics of processes distributed in activation energy. *Phys. Rev.* 100:1677–1689.
20. Agmon, N., and J. J. Hopfield. 1983. CO binding to heme proteins: A model for barrier height distributions and slow conformational changes. *J. Chem. Phys.* 79:2042–2053.
21. Young, R. D., and S. F. Bowne. 1984. Conformational substates and barrier height distributions in ligand binding to heme proteins. *J. Chem. Phys.* 81:3730–3737.
22. Jaynes, E. T. 1983. Papers on Probability, Statistics and Statistical Physics. R. D. Rosenkranz, editor. D. Reidel, Dordrecht, The Netherlands.
23. Bücher, T., and J. Kaspers. 1947. Photochemische Spaltung des Kohlenoxydmyoglobins durch ultraviolette Strahlung (Wirksamkeit der durch die Proteinkomponente des Pigments absorbierten Quanten). *Biochim. Biophys. Acta* 1:21–34.
24. Berendzen, J., H. Frauenfelder, T. Sauke, and R. Scholl. 1989. A logarithmic-timebase transient recorder. *Bull. Am. Phys. Soc.* 34:880*a* (Abstr.).
25. Steinbach, P. J., A. Ansari, J. Berendzen, et al. 1991. Ligand binding to heme proteins: the connection between dynamics and function. *Biochemistry* 30:3988–4001.
26. Elber, R., and M. Karplus. 1987. Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. *Science (Wash. DC).* 235:318–321.
27. Frauenfelder, H., F. Parak, and R. D. Young. 1988. Conformational substates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* 17:451–479.
28. Ansari, A., J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young. 1985. Protein states and proteinquakes. *Proc. Natl. Acad. Sci. USA.* 82:5000–5004.
29. Noguti, T., and N. Gō. 1989. Proteins: structural basis of hierarchical multiple substates of a protein. I. Introduction. *Proteins.* 5:97–103.

30. Hong, M. K., D. Braunstein, B. R. Cowen, et al. 1990. Conformational substates and motions in myoglobin: external influences on structure and dynamics. *Biophys. J.* 58:429–436.

31. Frauenfelder, H., P. J. Steinbach, and R. D. Young. 1989. Conformational relaxation in proteins. *Chem. Scr.* 29A:145–150.

32. Kuriyan, J., S. Wilz, M. Karplus, and G. A. Petsko. 1986. X-ray structure and refinement of carbon-monoxy (FeII)-myoglobin at 1.5 Å resolution. *J. Mol. Biol.* 192:133–154.

33. Ormos, P., D. Braunstein, H. Frauenfelder, M. K. Hong, S.-L. Lin, T. S. Sauke, and R. D. Young. 1988. Orientation of carbon monoxide and structure-function relationship in carbonmonoxymyoglobin. *Proc. Natl. Acad. Sci. USA.* 85:8492–8496.

34. Ansari, A., J. Berendzen, D. Braunstein, et al. 1987. Rebinding and relaxation in the myoglobin pocket. *Biophys. Chem.* 26:337–355.

35. Berendzen, J., and D. Braunstein. 1990. Temperature-derivative spectroscopy: a tool for protein dynamics. *Proc. Natl. Acad. Sci. USA.* 87:1–5.

36. Alberding, N., R. H. Austin, K. W. Beeson, S. S. Chan, L. Eisenstein, H. Frauenfelder, and T. M. Nordlund. 1976. Tunneling in ligand binding to heme proteins. *Science (Wash. DC).* 192:1002–1004.

37. Alben, J. O., D. Beece, S. F. Bowne, et al. 1980. Isotope effect in molecular tunneling. *Phys. Rev. Lett.* 44:1157–1160.