Analysis of Kinetics Using a Hybrid Maximum-Entropy/Nonlinear-Least-Squares Method: Application to Protein Folding

Peter J. Steinbach,* Roxana Ionescu,† and C. Robert Matthews‡

*Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, †Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802, and ‡Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

ABSTRACT A hybrid analysis that combines the maximum entropy method (MEM) with nonlinear least squares (NLS) fitting has been developed to interpret a general time-dependent signal. Data that include processes of opposite sign and a slow baseline drift can be inverted to obtain both a continuous distribution of lifetimes and a sum of discrete exponentials. Fits by discrete exponentials are performed with initial parameters determined from the distribution of lifetimes obtained with the MEM. The regularization of the parameter space achieved by the MEM stabilizes the introduction of each successive exponential in the NLS fits. This hybrid approach is particularly useful when fitting by a large number of exponentials. Revision of the MEM "prior" based on features in the data can improve the lifetime distribution obtained. Standard errors in the mean are estimated automatically for raw data. The results presented for simulated data and for fluorescence measurements of protein folding illustrate the utility and accuracy of the hybrid algorithm. Analysis of the folding of dihydrofolate reductase reveals six kinetic processes, one more than previously reported.

INTRODUCTION

The analysis of kinetic signals in terms of exponential processes is pervasive, with numerous applications found in physics, chemistry, biophysics, and medicine (Istratov and Vyvenko, 1999). In biological systems, chemical changes and atomic motions occur on timescales ranging from femtoseconds to seconds. The motions characteristic of a protein are the link between the protein's structure and its biological function. Consequently, the kinetics of many functionally important motions such as local protein conformational changes, ligand rebinding, and protein folding have been monitored extensively by time-resolved spectroscopic methods. The goal of such studies is a detailed model of the kinetics that accounts for all transitions among intermediate states and any other processes that contribute to the measured signal. Before such a model can be formulated, an accurate kinetic description must be obtained from the data. The number of kinetic processes reflected in the data and the amplitude and characteristic rate of each process must be determined, and it is this problem we address here. Extracting the number of processes necessary to describe the data can be difficult and requires that the noise in the data be well characterized. Moreover, the rates recovered from the measurement depend on the number of kinetic processes used.

Kinetics are commonly interpreted as a sum of discrete exponential processes. Nonlinear least squares (NLS) or maximum likelihood (ML) algorithms are typically used to obtain the amplitude and lifetime for each process. The parameter values sought are those that maximize the probability of measuring the data set (to within the noise) given the parameterized model. This optimization is most rigorously performed for kinetics with normally distributed errors using NLS fitting and for kinetics with Poisson noise using alternative ML approaches (Baljzer and Prendergast, 1992). Conventional NLS and ML algorithms require a priori specification of the number of exponentials to be included in the model. In addition, these procedures become increasingly hindered by the "multiple-minimum problem" as the number of kinetic processes increases. Because exponentials are not orthogonal along the real axis, a change in one parameter can be compensated for by changes in the others (Istratov and Vyvenko, 1999). When searching for optimal parameter values on a multidimensional probability surface possessing a large number of minima, the values obtained will depend on the initial values assigned.

Alternatively, quasi-continuous lifetime distributions can be obtained from the kinetics using a regularization method (Press et al., 1992). Two well-known approaches are the Tikhonov–Phillips regularization (Phillips, 1962), implemented in the program CONTIN (Provencher, 1982a,b), and the maximum entropy method (MEM) (Livesey and Brochon, 1987; Lavalette et al., 1991; Steinbach et al., 1991, 1992; Lambright et al., 1993; Johnson et al., 1996; Hogiu et al., 1997; Plaza del Pino et al., 1997; Kumar et al., 2001). The MEM has also been used to simultaneously fit multiple curves measured at different temperatures to obtain a two-dimensional distribution of activation enthalpy and entropy (Steinbach, 1996). Regularization algorithms offer several advantages over discrete approaches such as NLS or ML. The number of kinetic components need not be specified. By optimizing the regularizing functional for a given quality of fit, structure is introduced into the lifetime dis-

Submitted September 7, 2001, and accepted for publication January 4, 2002.

Address reprint requests to Peter J. Steinbach, 12 South Dr. MSC 5626, Bldg. 12A, Rm. 2051, Bethesda, MD, 20892-5626. Tel.: 301-496-1100; Fax: 301-402-2867; E-mail: steinbac@helix.nih.gov.

0006-3495/02/04/2244/12 $2.00

tribution conservatively. To this end, CONTIN minimizes the sum of the squares of the second differences (discretized form of the second derivative) of the distribution while MEM algorithms maximize the entropy of the distribution. For a given quality of fit, both the CONTIN and MEM solutions are unique; there is no multiple-minimum problem. However, the use of a regularizing functional biases the distribution and consequently fails to completely separate the desired signal from the noise, yielding nonrandom residuals. In addition, the MEM requires that a prior distribution be specified; this prior is defaulted to in the limit of low signal-to-noise and, ideally, incorporates existing knowledge into the distribution to be obtained.

Here, we present a hybrid MEM/NLS analysis applicable to a general time-dependent signal. The new software package, MemExp, accommodates kinetics possessing positive and negative slope and a slowly varying baseline. A continuous description that evolves according to the MEM is used to guide a series of discrete NLS fits during which one exponential is added at a time. Thus, the hybrid algorithm provides an automated and objective approach to the multiple-minimum problem that plagues conventional parametric fitting (NLS and ML) when many kinetic processes are present. Consequently, MemExp is particularly useful when a large number of discrete processes is an appropriate kinetic description. By interpreting kinetics simultaneously from the complementary perspectives of discrete exponentials and continuously distributed lifetimes, a comprehensive assessment of the data is performed conveniently. Also, the use of several different priors is supported to improve the fidelity of the distributions recovered. One continuous and one discrete representation are recommended, and results are plotted to facilitate evaluation of the results.

In the next section, the MemExp algorithm is described. The software is then tested using simulated data in the distributed, distributed plus discrete, and discrete kinetic regimes. Finally, the folding kinetics of dihydrofolate reductase (DHFR) in the presence of methotrexate (MTX) is analyzed using MemExp. Six kinetic processes are recovered from the fluorescence data, one more than have been previously identified.

MemExp: a MEM/NLS algorithm

Kinetics measured at times t_i can be fit by

$$\mathcal{F}_i = D_0 \int_{-\infty}^{\infty} \mathrm{d}\log\tau\,(g(\log\tau) - h(\log\tau))e^{-t_i/\tau} + \sum_{k=0}^{3}(b_k - c_k)(t_i/t_{max})^k, \quad (1)$$

where $g(\log\tau)$ and $h(\log\tau)$ are the lifetime distributions that correspond to decaying and rising kinetics, respectively, and a polynomial accounts for the baseline. The constant t_{max} is approximately the longest time measured and prevents baseline coefficients from differing greatly in magnitude. Formally, Laplace transforms involve an integral over the rate coefficient $k = 1/\tau$, but when analyzing data that span several decades in time, it is more convenient to express the underlying distribution as a function of $\log\tau$ (Livesey and Brochon, 1987). Upon discretizing Eq. 1 for a computer, the g and h distributions become vectors. All **g**, **h**, **b**, and **c** parameters may be restricted to positive values without loss of generality. The normalization constant D_0 can be estimated from the data, provided that the temporal window spanned by the measurement is sufficient to include all kinetic processes. In the presence of experimental uncertainties, the unconstrained numerical inversion of Eq. 1 is known to be an ill-posed problem; infinitely many solutions exist that fit the data to within the noise (Provencher, 1992). Alternatively, a fit can be done by n_e exponentials:

$$\mathcal{F}_i = D_0 \sum_{j=1}^{n_e} A_j e^{-t_i/\tau_j} + \sum_{k=0}^{3} B_k (t_i/t_{max})^k, \quad (2)$$

where A_j and τ_j are the amplitude and lifetime of the jth exponential, respectively, and B_k is the kth coefficient in the polynomial approximation of the baseline.

The MEM part of MemExp is an extension of previous work (Steinbach, 1996) and is based on the algorithm of Cornwell and Evans (1985). It fits Eq. 1 to kinetics by an iterative, second-order optimization of the entropy S (Skilling, 1989),

$$S(\mathbf{f}, \mathbf{F}) = \sum_{j=1}^{M} [f_j - F_j - f_j \ln(f_j/F_j)], \quad (3)$$

subject to constrained values of χ^2,

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathcal{F}_i - D_i}{\sigma_i}\right)^2, \quad (4)$$

and, optionally, the normalization I,

$$I = \Delta\left(\sum_{j=1}^{M_1} g_j - \sum_{j=1}^{M_2} h_j\right). \quad (5)$$

The constrained optimization is achieved by maximizing the function Q,

$$Q \equiv S - \lambda\chi^2 - \alpha I, \quad (6)$$

where λ and α are Lagrange multipliers. The natural logarithm in Eq. 3 requires that all elements of the desired image **f** be positive. Here, **f** is the concatenation of up to four vectors: **g**, **h**, **b**, and **c** (Eq. 1). M_1 and M_2 pixels are used to discretize the g and h distributions, respectively, and Δ is the spacing in $\log\tau$. Implicit in the use of χ^2 is the assumption that the standard errors, σ_i, in the measured data, D_i, can be assumed Gaussian. Note also that, if **f** is normalized ($\sum_{j=1}^{M} f_j$ = constant) and if all F_j are set equal to a constant,

then maximizing S (Eq. 3) boils down to maximizing a more familiar expression for the entropy

$$S = -\sum_{j=1}^{M} f_j \ln(f_j).$$

Unconstrained maximization of S ($\lambda = \alpha = 0$) with respect to f_j yields $f_j = F_j$. That is, **F** is the prior image that is defaulted to in the absence of good data. The ability to manipulate **F** lends flexibility to MEM inversions not available to other regularization methods. In contrast, this added flexibility requires that **F** be chosen wisely. A uniform (flat) **F** was typically used in early applications of the MEM. However, in some instances, artifactual ripples are present in distributions obtained with a flat **F**. For example, Provencher (1992) demonstrated the problem encountered by regularization methods when analyzing kinetics made up of an exponential process proximal in time to a broad process. Whether the analysis was performed with CONTIN or with the MEM using a uniform **F**, ripples were superimposed on the broad feature in the lifetime distribution. Such artifacts arise because regularizers are designed primarily to suppress abrupt changes in the lifetime distribution, favoring broad distributions with minimal higher-order derivatives instead. When analyzing kinetics with both discrete and distributed phases, the use of a regularizing functional that depends equivalently on all regions of the lifetime distribution (as do CONTIN and the MEM with a flat **F**) will tend to oversmooth the sharp features (e.g., exponentials) and undersmooth the slowly varying features. In many cases, unwarranted ripples can be smoothed out of the MEM image by deriving a data-dependent **F** from a preliminary **f**. A simple approach is to convolute **f** with a Gaussian and to restart the MEM calculation with this blurred image as the prior (Gull, 1989). In the context of recovering lifetime distributions, this uniform blurring of **f** to create a revised **F** incorporates broad kinetic features into the definition of S.

Given that traditional regularization algorithms can falter when faced with a mixture of sharp and broad features, lifetime distributions recovered by the MEM might be improved considerably if features in **f** were incorporated into **F** in a manner dependent on how large and well resolved they are. MemExp allows the larger, well-resolved peaks to be included in **F** unchanged while the smaller, less-resolved features are blurred before being added to **F**. Moreover, the well-resolved peaks can be included in the prior after being sharpened; a Gaussian of the same area but smaller full width at half maximum (FWHM) can be used to represent them in **F**. As shown below, this differential blurring can lead to significantly smoother distributions when sharp and broad kinetic features are close in time.

To avoid excessive curvature being attributed to the baseline in the MEM fits, the values of F_j associated with the linear, quadratic, and cubic coefficients are very small (e.g., $b_1 = b_2 = b_3 = c_1 = c_2 = c_3 = 10^{-20}$) compared to the F_j values used for the constant coefficients (e.g., $b_0 = c_0 = 10^{-3}$). If a data-dependent **F** is obtained by blurring **f**, only the **g** and **h** elements of **f** are blurred; the baseline elements of **F** are simply assigned the **f** values.

The MEM calculation is iterative, starting with an initial image having maximum entropy: **f** = **F** (initially, a constant); $\lambda \approx \alpha \approx 0$. Newton–Raphson optimizations of Q at fixed values of λ and α are followed by automatic adjustments of the Lagrange multipliers. The multipliers are changed gradually enough to ensure that the gradient of Q remains sufficiently small (Cornwell and Evans, 1985). Thus, **f** evolves from a flat distribution (with very large χ^2) into a series of maximum-entropy distributions that become increasingly structured as χ^2 and I approach the desired values. Whenever the prior **F** is derived from **f**, **f** is set equal to the new prior and the Lagrange multipliers are reset to near zero. The optimization of Q and multiplier updating are iterated until χ^2 stops changing appreciably.

When χ^2 reaches a specified value (e.g., 1.2), the current **f** distribution is written to a file, and MemExp performs the first NLS fit (Eq. 2) with one exponential included for each MEM peak having an appreciable area and a mean within a specified log τ range. Because distributions produced by the MEM are quite smooth and noiseless, simple identification of minima and maxima in **f** is sufficient to detect peaks in the lifetime distribution. Let i_- and i_+ denote the location of the minima immediately preceding and following the ith local maximum in **f**, respectively. Then the area and mean of this MEM peak are obtained by numerical integration, as in

$$\text{Area}_i = \Delta \sum_{j=i_-}^{i_+} g_j, \tag{7}$$

$$\text{Mean}_i = \frac{\Delta}{\text{Area}_i} \sum_{j=i_-}^{i_+} g_j \log \tau. \tag{8}$$

For the outermost MEM peaks (fastest and slowest processes), the range of integration extends to the distribution edge. Of course, the portions of **f** corresponding to **g** and **h** are searched for peaks independently. (Also, the first and last terms in Eqs. 7 and 8 are multiplied by ½ as in the trapezoidal rule.) The initial amplitude and log lifetime of the ith exponential in the NLS fit are then taken as Area_i and Mean_i, respectively. During NLS optimization of the discrete parameters, the lifetimes are free to stray outside the specified log τ range. This range is used to exclude exponentials from the NLS fit that correspond to undesirable ripples in the MEM distribution at either end of the log τ axis.

The analysis of local maxima performed by MemExp also serves to identify relatively sharp, well-resolved peaks. In addition to the area and mean, two ratios are calculated for each local maximum in the lifetime distribution. The value of **f** at the local maximum is divided by each of the

values of **f** at the adjacent minima. For a small ripple superimposed on a broad peak, at least one of these ratios is approximately unity. For a well-resolved peak, both ratios are large, perhaps exceeding 100. For a narrow peak partially overlapping a broad peak, one ratio will be rather large (e.g., 100) and the other will be smaller (e.g., 2), depending on the extent of overlap. If differential blurring of **f** into **F** is to be performed and if the *i*th MEM peak is found to be sufficiently large and sufficiently well resolved, then it is subtracted from **f** in the range from i_- to i_+. After all such peaks are subtracted, the remainder of **f** is blurred uniformly. Then each subtracted peak is accounted for, either by adding it to **F** unchanged or by adding a Gaussian with the same area and a reduced FWHM. For peaks that partially overlap a neighboring feature, the peak subtracted from **f** before blurring is obtained by essentially "folding" the peak over from the side opposite the overlapping feature so that a symmetric peak is introduced into **F**.

As the MEM convergence continues, this analysis (and storage) of the **f** distribution is repeated periodically to determine the current number of MEM peaks n, the peak means and areas, and the maximum-to-minimum ratios. If n has not changed since the last time the image was analyzed, the MEM calculation is resumed. If n has increased by one, the areas and means of the current MEM peaks are characterized (Eqs. 7 and 8) and a new NLS fit is performed with one more exponential than was used in the previous NLS fit. If n has increased by more than one, the $n + 1$ MEM peaks of largest area are introduced as exponentials in an NLS fit. Then, the next largest MEM peak is accounted for in an additional NLS fit, and so on. Thus, the MEM is used to introduce one exponential at a time into a series of NLS fits; as features are resolved in the continuous kinetic description, exponentials are added to the discrete description. The MEM algorithm may converge before the corresponding discrete fit describes the data well. Consequently, upon convergence of the MEM calculation, the widest MEM peak is identified and an NLS fit is done in which this peak is represented as two discrete exponentials, each initially accounting for half its area and displaced from its mean on either side.

Another important detail of the hybrid algorithm is the preservation of the sign of each kinetic process during the NLS fitting. The sign s_j is $+1$ for peaks in the g distribution and -1 for peaks in h. The amplitude of the jth exponential is expressed as $A_j = s_j a_j^2$, and the a_j parameters, not the A_j, are varied. Also, a second value of χ^2 is reported for the NLS fits to facilitate their comparison to each other in which N in the denominator of Eq. 4 is reduced by N_p, the number of NLS parameters.

Beginning near the end of the calculation (e.g., $\chi^2 \approx 1.2$), the MEM and NLS calculations are characterized numerically and summarized graphically to facilitate subsequent evaluation by the program and by the user. The data, fit, baseline, residuals of the fit R, autocorrelation function of the residuals C, and lifetime distribution are plotted. The autocorrelation of the residuals is calculated as

$$C_j = \frac{\langle (R_i - \langle R \rangle)(R_{i+j} - \langle R \rangle) \rangle}{\langle R^2 \rangle - \langle R \rangle^2}, \tag{9}$$

where angle brackets denote averaging over the index i and

$$R_i = \frac{\mathcal{F}_i - D_i}{\sigma_i}. \tag{10}$$

The correlation length of the residuals, τ_c, is approximated by summing $|C_j|$ over the first three-quarters of the autocorrelation function. Because τ_c is a measure of systematic deviations of the fit from the data, this quantity is helpful in choosing the iteration that is to represent the desired solution.

Upon completion of the MEM convergence, MemExp recommends a distributed and a discrete fit as the optimal kinetic descriptions. Selecting $\chi^2 = 1$ is not necessarily optimal; this choice can result in overly conservative underfitting (Gull, 1989; Titterington, 1985). In addition, a selection based solely on the value of χ^2 can be unduly sensitive to the error estimates, σ_i, which are often not known exactly. Instead, MemExp recommends distributed and discrete solutions based on changes in χ^2 and τ_c. Here, the discrete fit recommended by MemExp corresponds to the greatest number of exponentials n_e for which χ^2 has decreased by more than 1% from the fit to $n_e - 1$ exponentials. The recommended distributed fit is the last for which both χ^2 and τ_c have decreased by more than 1% from the preceding distribution analysis. The choice of the continuous description is somewhat arbitrary, because the changes in χ^2 and τ_c depend on the frequency at which images are analyzed. Note that the recommended continuous description is not, in general, the one from which the recommended discrete description was initiated. These selection criteria were adopted here because they worked very well for all simulated data analyzed (see below). For real data, where baseline drifts or other systematic errors might render Eqs. 1 and 2 more approximate, the recommended solutions may tend to overfit the data. We therefore chose the smallest number of exponentials recommended by MemExp for any of the many fluorescence measurements analyzed with a given baseline description.

One more feature of MemExp warrants description. A multi-step analysis of raw data has been automated in which one or more preliminary MEM-only inversions is performed to estimate the standard errors of the data. Initially, all standard errors are set equal to a constant and a MEM-only inversion is performed. The last MEM fit for which both χ^2 and τ_c have decreased is identified. This fit's deviations from the data are analyzed in a narrow range (e.g., $m = 2$)

centered about each data point to estimate the standard errors,

$$\sigma_i = \sqrt{\frac{1}{2m}\sum_{j=i-m}^{i+m}(D_j - \mathcal{F}_j)^2}. \quad (11)$$

These estimates are then smoothed and input to another inversion. This use of a MEM calculation to estimate the noise in the data can be repeated, but the error estimates generally converge quickly, and one preliminary MEM inversion often suffices. The selection criteria described above in terms of changes in χ^2 and τ_c were found to perform well with errors estimated in this iterative way and with the true errors used to generate simulated data.

Linear algebra calculations done during the MEM optimization are performed using routines in the public-domain LINPACK library. NLS fitting is done using public-domain MINPACK routines. The MemExp software, documentation, and sample input files are available online. MemExp runs under several operating systems.

Simulated data and measured fluorescence

MemExp was tested using simulated data in three different kinetic regimes: purely distributed, distributed plus discrete, and purely discrete. First, test data were simulated with realistic noise from the continuous distribution corresponding to a stretched exponential, $\exp(-(t/\tau)^\beta)$, with $\beta = 0.5$ (Götze, 1991). Second, following Provencher (1992), data were simulated as the sum of an exponential process at $\log \tau = -6$ and a broadly distributed process centered at $\log \tau = -3$. These data were all given the same standard error, $\sigma = 0.001$. Third, two groups of ten kinetic traces were simulated from a model involving five exponentials (one negative and four positive amplitudes) and a constant offset. This discrete model is similar to the measured protein-folding kinetics. The first group was simulated with errors determined by MemExp for an actual measurement. The second was simulated with these errors increased tenfold.

The protocol for the expression and purification of wild type DHFR from *Escherichia coli* was described elsewhere (Jennings et al., 1993). The buffer contained 10 mM phosphate, 0.2 mM K_2EDTA and 1 mM β-mercaptoethanol at pH 7.8. Protein dissolved in 5 M urea/buffer was rapidly mixed with MTX/buffer with a stopped-flow instrument (model SX.18 MV, Applied Photophysics Ltd., Leatherhead, U.K.). The temperature of the mixing cell was controlled by a water bath. The final concentrations of the reactants in the mixing cell were: DHFR, 3 μM; MTX, 15 μM; urea 0.83 M. Similar results were obtained when the ratio [MTX]/[DHFR] was 3 or 10, for final [DHFR] of 10 and 5 μM, respectively. The excitation wavelength was 280 nm and the emission was recorded using a cut-off filter at 320 nm. MTX quenches the protein fluorescence upon binding, thus increasing the amplitudes of the fluorescence

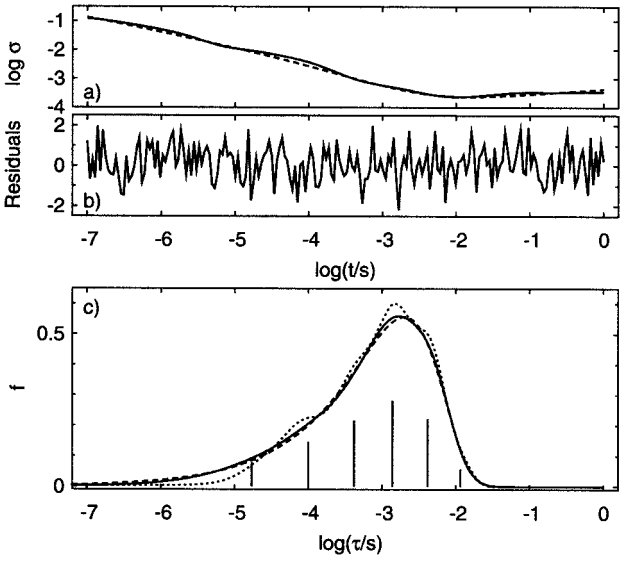

FIGURE 1 Analysis of stretched exponential kinetics. (*a*) Standard errors used to generate data (*dashed*) and estimated by MemExp (*solid*). (*b*) Residuals of fit shown as solid curve in (*c*). (*c*) MemExp recommendations. Distributions using $F_j = 0.001$ (*dotted*) and data-dependent **F** (*solid*). The six-exponential fit (*vertical lines*) reflects the true distribution (*dashed*).

changes in the folding process. The data were acquired approximately equally spaced in log time, ranging from 1 ms to 1000 s. Measurements were performed at four temperatures (12, 15, 18, and 21°C), and MemExp was applied to five (unaveraged) data sets at each temperature to permit an assessment of the kinetic processes recovered numerically. Control experiments were performed by mixing unfolded protein in 5 M urea with 5 M urea in the absence and presence of MTX, and by mixing MTX with buffer or urea.

Data-dependent **F** distributions were obtained by blurring **f** at $\chi^2 \leq 1.1$ with a Gaussian having a FWHM of 0.5 (simulated five-exponential and DHFR data) or 2.0 (simulated stretched-exponential data). For the data simulated in the distributed plus discrete regime, **F** was obtained by representing the local maximum in **f** with area greater than 0.05 and maximum-to-minimum ratios in excess of 30 (only one found) as a Gaussian with FWHM of ⅓ the original MEM peak. The remainder of **f** was convoluted with a Gaussian with FWHM = 1. No error estimates were input to MemExp to obtain the results reported here; errors were estimated using one to three preliminary MEM inversions.

RESULTS

The MemExp analysis of the simulated stretched-exponential kinetics is shown in Fig. 1. The errors estimated agree well with the true errors (Fig. 1 *a*) except for small undulations. Each of the three fits in Fig. 1 *c* fit the data well, with χ^2 between 0.73 and 0.86. Although both distributions obtained resemble the true distribution (Fig. 1 *c*), the one

obtained with a constant **F** has small shoulders near log $\tau = -4$ and -2.6. If **f** is blurred uniformly to obtain a data-derived **F**, these shoulders are eliminated, and a distribution much closer to the true one is recovered. The amplitudes of the six discrete exponentials recovered vary smoothly, also reflecting the shape of the MEM distributions. Thus, the user would be likely to conclude correctly that this data set is best described by distributed lifetimes. Clearly, the number of exponentials sufficient to fit stretched-exponential kinetics increases with increasing signal-to-noise and with decreasing β.

Figure 2 also illustrates the kind of improvement that can be obtained with a judicious choice of the MEM prior **F**. These test data were simulated from the same lifetime distribution used by Provencher (1992). Figure 2 *a* shows some of the graphical summary generated by MemExp, with the MEM calculation proceeding from top to bottom. Each time the MEM image is analyzed and stored, a row of plots is appended to this summary. In this analysis, the prior was updated when χ^2 reached 1.1 (*top*) by differentially blurring **f**. The one local maximum found to be well resolved was incorporated into **F** as a Gaussian having a FWHM equal to ⅓ that of the MEM peak. The remainder of **f** was blurred by a Gaussian with FWHM = 1 before incorporation into **F**. Upon restarting the calculation with the new prior, χ^2 increases temporarily (*middle*). The solution obtained with the revised prior **F** (*bottom*) is free of the ripples present in the distribution obtained with a flat prior (*top*) and produces a better fit (reduced values of χ^2 and τ_c). Note in the top left panel that the residuals obtained using a flat prior are especially large in the temporal region between the sharp and broad features in the kinetics. This local ringing in the residuals reflects the inability of a regularizer to optimally recover a delta function and the broad peak adjacent to it. Similarly unsatisfactory results were obtained when **F** was derived by blurring **f** uniformly (data not shown). If instead, a sharp feature is incorporated directly into **F** (*middle*), the subsequent maximization of S is focused on smoothing the remainder of the lifetime distribution (*bottom*). As a result, the artifacts introduced when using either CONTIN or the MEM with a flat **F** (Provencher, 1992) are successfully suppressed. Good results were also obtained after repositioning the exponential at log $\tau = -5$, one decade closer to the broad process (data not shown). Although the benefits of differentially blurring **f** into **F** decreased upon moving the exponential to log $\tau = -4$ and to log $\tau = -3$ (data not shown), the artifacts were also diminished because the data could then be better fit by a more homogeneously varying lifetime distribution of the kind favored by regularization methods.

The continuous and discrete descriptions recommended by MemExp for the two groups of data simulated with five exponentials are shown in Figs. 3 and 4. For each of the forty runs (ten data sets at two noise levels using two **F**), MemExp correctly recommended the fit by five exponentials. All eighty recommended fits (continuous and discrete) produced values of χ^2 between 0.75 and 0.92. A constant baseline was used to synthesize the data and to fit them, but very similar results were obtained when fitting with all four baseline parameters varied (data not shown). Thus, use of a cubic baseline did not impair the recovery of the kinetic processes.

The discrete descriptions of the data recommended by MemExp agree well with the true description, even at noise levels ten times greater than those achieved experimentally.

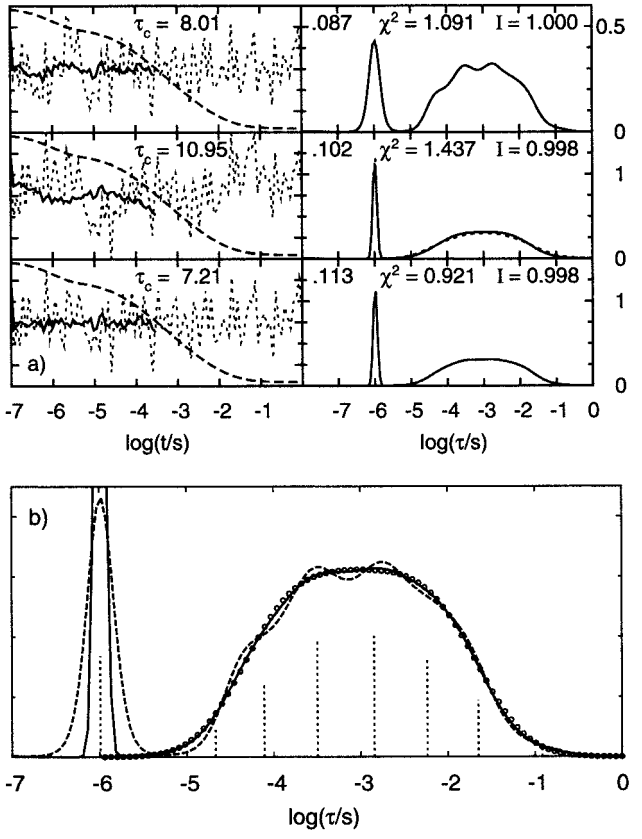

FIGURE 2 Analysis of kinetics simulated as the sum of one exponential and one distributed process. (*a*) Sample MemExp output. The calculation proceeds from top to bottom. The data and fit (*dashed*, indistinguishable on this scale) are plotted in the left column, along with the residuals of the fit (*dotted*) and autocorrelation of the residuals (*solid*). The correlation length τ_c is given. The lifetime distribution (*solid*) is plotted in the right column, along with the prior **F** upon its derivation from **f** (*dotted*, *middle panel*). These plots are labeled by the extension to the image file name (equal to the number of Q evaluations) and the values of χ^2 and image normalization I. Upon reaching a specified value of χ^2 (here, 1.1), **F** is derived from **f** using peak-dependent blurring. The one well-resolved local maximum in **f** is incorporated into **F** as a Gaussian with FWHM = ⅓ of the MEM peak's FWHM, and the rest of **f** is blurred by a Gaussian with FWHM = 1 before being added to **F**. (*b*) Lifetime distributions recovered using $F_j = 0.001$ (*dashed*, from *top panel* of *a*), and **F** obtained from **f** after sharpening the well-resolved peak (*solid*, from *bottom panel* of *a*). The true distribution is the sum of one exponential (amplitude = ⅙, log lifetime = −6) and one broad peak (*circles*). The recommended eight-exponential fit is shown as dotted vertical lines.

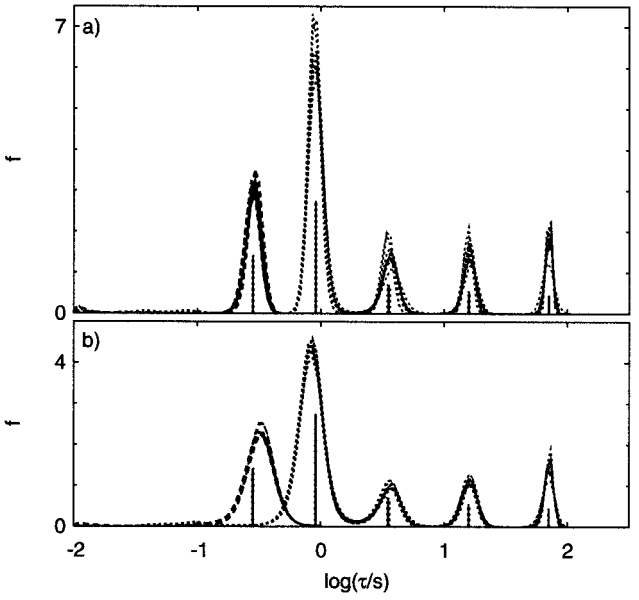

FIGURE 3 Lifetimes recovered for kinetics simulated with realistic noise from five exponentials (*solid vertical*). The fastest exponential has a negative amplitude. Ten distributions are plotted, with rising (*dashed*) and decaying processes (*dotted*). The ten discrete descriptions (*dotted vertical*) are nearly identical to the true description. (*a*) $F_j = 0.001$. (*b*) Data-dependent **F**.

The advantages of MemExp over the MEM alone are evident at the high noise level (Fig. 4), where the continuous descriptions disagree with regard to the number of kinetic processes, but the discrete do not. The MEM was able to fit several of the data sets without resolving both the second ($\log \tau = 0.55$) and third ($\log \tau = 1.2$) processes in the true **g** distribution. However, when NLS fits by discrete exponentials were performed with initial parameters determined from the MEM inversion, five exponentials were always needed. The ability to resolve adjacent peaks in a continuous description clearly increases with the signal-to-noise ratio, as is illustrated by this analysis at two noise levels from both the continuous and discrete perspectives. Also, the variance of exponential lifetimes recovered at high noise (Fig. 4 *c*) underscores the benefits of analyzing multiple raw data files.

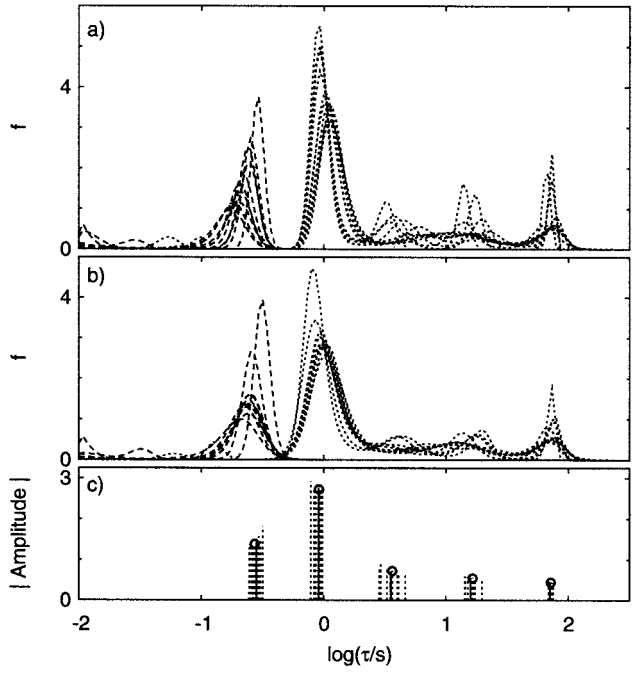

FIGURE 4 Lifetimes recovered for discrete kinetics simulated with 10× more noise than in Fig. 3. (*a*, *b*) Continuous descriptions plotted as in Fig. 3. (*c*) Discrete descriptions (*dashed*), their average (*circles*), and the true description (*solid*).

The effects of using a data-derived **F** are also illustrated in these simulations. After **f** has been blurred uniformly to create **F**, the MEM inversion is biased toward one or more relatively short, broad peaks. When the kinetics is in fact broadly distributed, use of this data-dependent **F** can improve the results by smoothing out unnecessary undulations in the recovered lifetime distribution (Fig. 1 *c*). When the slope of the kinetics changes sign, this bias may cause adjacent peaks of opposite sign to overlap (Fig. 3 *b*). A slight shift of the first two peaks toward each other is also seen in this case. Clearly, only $\mathbf{g} - \mathbf{h}$ can be recovered from the data, not **g** and **h** independently. Here, the MEM fit is overparameterized, more so than for monotonic kinetics, and a range of peak positions and widths can fit the data in regions where the slope changes sign. Thus, the small but significant differences in MEM peak positions obtained in Fig. 3, *a* and *b*, for the first two processes indicate the small uncertainty in the lifetimes recovered for the two adjacent processes of opposite sign. Although use of the data-dependent **F** resulted in a biased shift of the first two MEM peaks, the corresponding five-exponential fits obtained agree extremely well with the true kinetic model. Analysis from both the distributed and discrete perspectives clearly demonstrates that the data allow excellent fits by slightly different peak positions and widths, particularly in regions where the slope of the data changes sign.

Representative fluorescence measurements of DHFR folding at 12 and 21°C are shown in Fig. 5. The data rise at short times and decay at long times, exhibiting too many features to be described by a stretched exponential. Note the unsatisfactory fit in Fig. 5. Also shown is the fluorescence measured for a control experiment in which (unfolded) DHFR in 5 M urea was mixed with MTX in 5 M urea, yielding final concentrations of 5.5 μM DHFR and 50 μM MTX. These control data indicate that the multiple processes observed in the folding experiments are indeed due mainly to folding events. The control also illustrates the need to consider a slowly varying baseline.

A MemExp analysis of fluorescence data highlights the advantages of interpreting kinetics from both the distributed and discrete perspectives. The four fits shown in Fig. 6

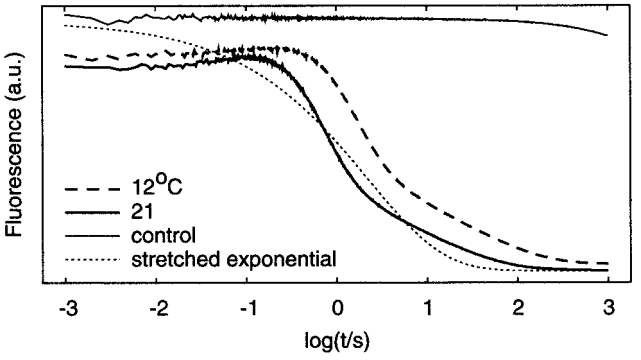

FIGURE 5 DHFR folding in the presence of MTX. Fluorescence measured during folding at 12 (*dashed*) and 21°C (*thick solid*) are plotted. Also shown are a control measurement (*thin solid*) during which DHFR remains unfolded and an unsatisfactory fit by a stretched exponential (*dotted*).

produce χ^2 values between 0.84 and 0.92. Small features exist in the recommended continuous distributions of lifetimes that do not correspond to any exponentials in the recommended discrete fits. Because the NLS fits evolve in a quantized fashion (with the addition of one exponential at a time), the choice of optimal discrete fit is generally clearer than the choice of optimal distributed fit. For the kinetics analyzed in Fig. 6, the recommended continuous fits overinterpret the data somewhat. Thus, the discrete fits performed by MemExp help to distinguish true kinetic processes from artifacts associated with over-fitting. These extraneous MEM features tend to be most prominent at the ends of the log τ axis, but they can also appear at intermediate values of log τ (Fig. 6 *b*). Note also that the slowest

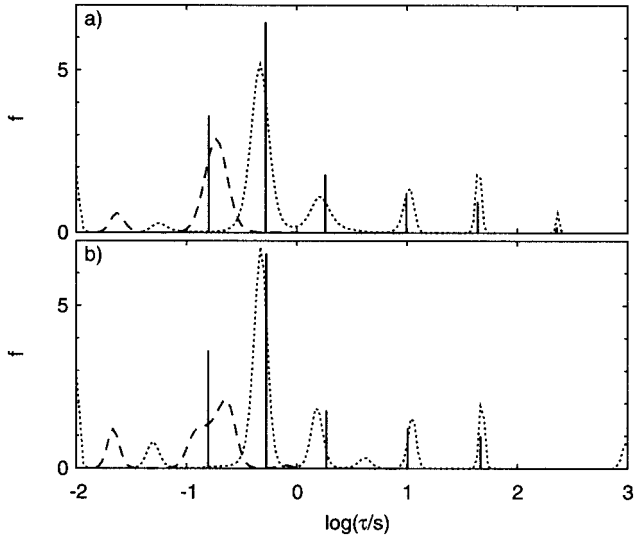

FIGURE 6 Lifetimes of DHFR folding in the presence of MTX at 21°C. The continuous and discrete (*vertical lines*, ×7.0) representations recommended with a constant (*a*) and a cubic (*b*) baseline. Distributions governing rising (*dashed*) and decaying kinetics (*dotted*) are shown. The fastest exponential has a negative amplitude.

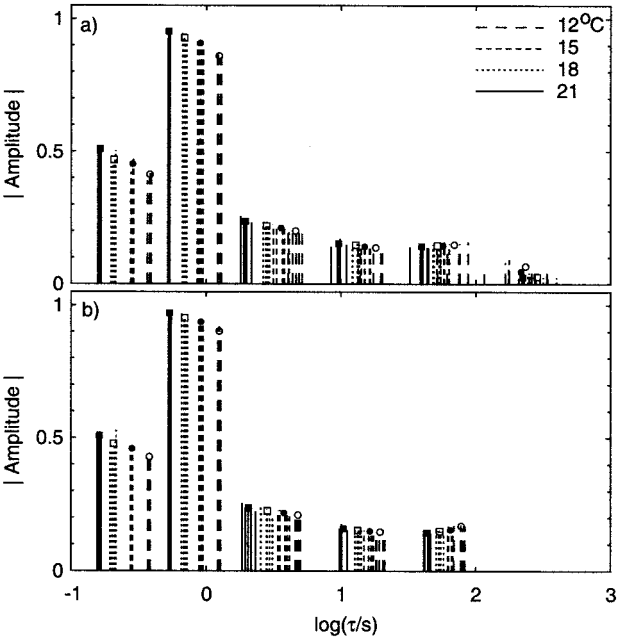

FIGURE 7 Discrete lifetimes of DHFR folding in the presence of MTX for five measurements at each of four temperatures. Symbols mark the mean log lifetime and mean amplitude. The fastest exponential has a negative amplitude. (*a*) Six exponentials recovered with a constant baseline. (*b*) Five exponentials recovered with a cubic baseline.

exponential recovered with a constant baseline is not present in the fits with a cubic baseline.

The MemExp analysis of five fluorescence measurements at each of four temperatures is shown in Fig. 7. The forty χ^2 values obtained by these discrete fits range from 0.84 to 0.95. The discrete fits using a constant baseline (Fig. 7 *a*) call for six exponentials, one rising and five decaying processes. The rate recovered for the slowest process ($\tau \approx$ 200 s) does not increase monotonically with temperature as do the other rates. When a cubic baseline was used, this slow process was interpreted as baseline drift (Fig. 7 *b*), but the rates of the remaining five exponentials changed very little from those obtained with a constant baseline.

Under similar solvent conditions, the folding of DHFR at 15°C has been reported to involve five exponentials, denoted as τ_1 to τ_5, with τ_1 being the slowest phase (Jennings et al., 1993; Touchette et al., 1986). Instrumental limitations in the earlier work required that the fastest phases be measured by stopped-flow fluorescence and the slowest phases (τ_1 and τ_2) by manual mixing absorption. The two fastest and the slowest lifetimes recovered at 15°C with MemExp using a constant baseline agree very well with the processes denoted earlier as τ_5, τ_4, and τ_1, respectively. However, the data at intermediate times require three additional processes in the current analysis, whereas two processes have sufficed in the past. The observation of an additional phase in the MemExp analysis suggests not only an improved numerical approach but an enhanced signal-to-noise ratio produced

with a more stable fluorescence instrument. It is possible that the slowest decay is not well characterized because the maximum measurement time permitted by the fluorescence instrument is 1000 s.

Between 12 and 21°C, the rates of the five fastest processes increase with increasing temperature, and fits by the Arrhenius law yield activation enthalpies ranging from 11 to 16 kcal/mol. Because our protein-folding measurements cannot be performed over a large range in temperature, this estimate of barrier heights is admittedly crude and serves only as a semi-quantitative consistency check. Still, these estimated activation enthalpies are reasonable and are close to those reported in the literature for other proteins (Bilsel and Matthews, 2000).

DISCUSSION

The MEM is said to be "maximally noncommittal" with regard to unavailable information. This property has been illustrated by the so-called kangaroo problem: given that one-third of all kangaroos are left handed and one-third of all kangaroos are blue eyed, what fraction of kangaroos is both blue eyed and left handed? Any answer between zero and one third is consistent with the limited information given. Let x denote the fraction of kangaroos that are both blue eyed and left handed. Then the fraction that are blue eyed and right handed is $⅓ - x$, the fraction that are not blue eyed but are left handed is $⅓ - x$, and the fraction that are neither blue eyed nor left handed is $⅓ + x$. Taking F_j to be constant, $S = -\sum_{j=1}^{4} f_j \ln(f_j) = -x \ln(x) - 2(⅓ - x)\ln(⅓ - x) - (⅓ + x)\ln(⅓ + x)$. Maximizing S with respect to x yields $x = ⅑$. Thus, no correlation is imposed between the eye color and handedness of kangaroos; the MEM with a uniform prior is maximally noncommittal when given incomplete information.

This conservative, noncommittal nature of the MEM has its drawbacks. At a given value of χ^2, the image **f** is biased to remain close to the prior **F**. Consequently, the residuals of a MEM fit are not random, and the image may be excessively flattened. This bias can be especially problematic when a flat prior is used and the data call for both sharp and slowly varying features (Fig. 2 *a*, *top*). The ultimate goal of complete separation of signal from noise is not reached. Concomitantly, MEM images may include undesirable ripples (Figs. 2 and 4).

Because any regularization method biases the distributed solution sought, its ability to recover a lifetime distribution free of artifacts will depend on whether the distribution underlying the given kinetics trace is consistent with the bias imposed. CONTIN enforces smoothness by minimizing the sum of the squared second differences of **f**. The MEM biases the image to remain close to the prior **F** by maximizing the entropy (Eq. 3). In this light, the need to specify **F** in Eq. 3 is actually an advantage of the MEM relative to other regularization methods; the regularizer and its bias can be adjusted by changing **F**. MemExp facilitates the convenient use of several different priors. A flat prior is maximally noncommittal but does not enforce smoothness; each pixel in **f** is biased toward the same value, regardless of the values of **f** at neighboring pixels. A prior obtained by uniformly blurring **f** enforces smoothness but may still under-fit sharp features. A prior obtained by differentially blurring **f** as implemented in MemExp can improve the distributed fit to sharp features but introduces some subjectivity into the analysis. Because MemExp allows convenient manipulation of the prior, it is relatively easy to assess the sensitivity of **f** to changes in **F**.

The Pixon method of image reconstruction (Puetter and Yahil, 1998) overcomes many of the regularization problems encountered in conventional image processing. Because the point-spread function in typical deconvolution applications is relatively narrow, a given pixel in the image affects only a small region of the data. The Pixon method exploits this property by using a local goodness of fit to recover a multiresolution image of considerably reduced complexity (fewer parameters) relative to a MEM image. Residuals obtained by the Pixon method lack the spatial correlation present in MEM residuals, and the Pixon method has been applied successfully to many imaging problems. However, the analog to the point-spread function in deconvolution is an exponential decay in Laplace inversion. Consequently, a pixel in a distribution of lifetimes influences the fit to many data points, not just those in a narrow temporal range. This nonlocal character of the Laplace transform and the nonorthogonality of exponentials make it more difficult to use a local goodness of fit to establish differences in resolution in a lifetime distribution.

Nonetheless, MemExp is similar in spirit to the Pixon method. Both methods are extensions to the MEM that describe the data using fewer parameters than are used in a MEM reconstruction. When applied to the deconvolution of blurred images, the Pixon method identifies regions in the data with low signal-to-noise ratio and reduces the number of pixels used in these regions. Similarly, MemExp identifies regions of the lifetime distribution with enhanced intensity, and switches from a continuous distribution of peaks of finite width to one of infinitely narrow peaks and far fewer parameters.

For a given value of χ^2 and a given default image **F**, the MEM image is unique for linear image-data transformations. This uniqueness is especially advantageous as the number of kinetic processes becomes large. In this regime, the dependence of the final parameter values on the initial values complicates multiexponential fitting using NLS or ML (Baljzer and Prendergast, 1992) alone. By initializing the NLS fit parameters based on the largest MEM features within a specified lifetime range, the important choice of initial parameter values is done objectively and reasonably using MemExp.

By using two traditional fitting techniques in concert, the MemExp algorithm is an attempt to get the best of both and the worst of neither. Maximizing the entropy solves the multiple-minimum problem but at the expense of biasing the solution to the point that small, unwarranted features may appear (Figs. 2 and 4) and larger peaks may be shifted slightly (Fig. 3 *b*). When fitting with nonlinear least squares, χ^2 is reduced without bias, but the answer may depend on the initial values chosen for the parameters. In recovering a discrete representation of the kinetics, MemExp maximizes the entropy only to get in the neighborhood of the answer, and the final answer is obtained by an unbiased minimization of χ^2 alone. The attractive properties of the MEM are exploited, but the associated side effects are avoided. The simple, robust description of the data in terms of multiple exponentials is optimized with confidence that the multiple-minimum problem has been addressed effectively.

The analysis of simulated and experimental data validates the protocol implemented in MemExp. Several strengths of this MEM/NLS algorithm have been demonstrated. First, a rather complicated analysis is automated in MemExp to a considerable extent, and the important choice of prior **F** can be made conveniently from among several alternatives. It is generally accepted that any existing knowledge of the measurement should be incorporated into **F** whenever possible. Here, we have seen that knowledge of the analysis, i.e., the tendency of regularization methods to oversmooth sharp features proximal to slowly varying ones can also be included in the prior to good effect. Because the numerical inversion of Eq. 1 is ill-posed in the presence of noise, it is prudent to analyze a data set more than once, considering a uniform prior and priors derived from **f** with and without differential blurring. If features in the lifetime distribution disappear upon deriving the prior from **f** (Figs. 1 and 2), their existence is not established by the measurement. Given the subjectivity inherent in differential blurring, it seems best to choose the distribution obtained in this way only when it possesses fewer local maxima than those recovered using other priors (Fig. 2). The detailed graphical summary of each calculation provided by MemExp facilitates the interpretation of the kinetics.

Second, the errors estimated by MemExp (Fig. 1 *a*) are verified by the observation that the χ^2 values of the fits chosen from the final MEM/NLS inversions are typically in the range of 0.7 to 1.0 and that the corresponding residuals (Fig. 1 *b*) are reasonably uniform in magnitude throughout the entire temporal range of the data.

Third, MemExp uses an evolving lifetime distribution obtained by the MEM to guide a series of discrete fits to the data. A distributed and a discrete interpretation are recommended automatically. MemExp correctly identified the five-exponential model for each of the twenty simulated data sets, including the ten simulated with noise ten times greater than that of a fluorescence experiment.

Fourth, MemExp facilitates the verification of features in the MEM distribution. The reduced number of parameters in the discrete fit leads to a more robust optimization. When the number of exponentials used in the NLS fit is less than the number of peaks in the MEM description at comparable values of χ^2 (Fig. 6), the MemExp user knows which of the MEM features are necessary to fit the data and which are undesirable artifacts.

Fifth, analysis from both the continuous and discrete perspectives helps assess the resolvability of kinetic processes given the noise in the data (Alcala et al., 1987). Because the MEM is biased in favor of a minimally structured distribution, it may underinterpret data for which the processes are truly discrete exponentials (Fig. 4). By contrast, the NLS fits are biased to introduce structure only via infinitely narrow peaks. When these two approaches produce comparable structure at comparable values of χ^2 (Fig. 3), the signal-to-noise is sufficient to unambiguously establish the number of kinetic processes. Parameter uncertainties can be estimated from the MEM peak widths and from any differences in peak positions. When the number of peaks in the MEM description is less than the number of exponentials used in the NLS fit at comparable values of χ^2 (Fig. 4), the data are consistent with two qualitatively different interpretations, and additional or better data are needed to definitively determine the number of processes.

Whether a distributed or a discrete interpretation might be favored a priori depends on the experiment being analyzed. Distributed kinetics is seen in biomolecular systems when the reaction rate depends on a "conformational coordinate" and relaxation along this coordinate is slow compared to the reaction under study. For example, freezing a carboxymyoglobin (MbCO) solution in a glycerol/water solvent produces a heterogeneous system in which individual proteins are trapped in different conformations possessing different lifetimes of the CO-dissociated state (Austin et al., 1975). Following flash photolysis of the CO ligand at temperatures below ~180 K, protein conformations do not interconvert on the time scale of CO rebinding. Consequently, the low-temperature kinetics is naturally described by distributed lifetimes. In this low-temperature regime, the MEM can be used to simultaneously fit rebinding kinetics measured at multiple temperatures by a two-dimensional distribution of activation enthalpy and entropy (Steinbach, 1996). Distributed rebinding of ligands to myoglobin is generally seen at high solvent viscosity, whether or not this high viscosity was induced by cooling (Austin et al., 1975; Hagen et al., 1995). In fact, nonexponential kinetics is seen for MbCO even in aqueous solution at room temperature, indicating that relevant protein relaxations occur during rebinding (Tian et al., 1992).

A revised version of the "new view" of protein folding stresses the multidimensional nature of the free energy landscape in which the thermodynamic and kinetic properties of the protein can be determined by the projection of the

potential energy surface onto one or two important coordinates (Brooks et al., 2001). For funnel-like energy landscapes, simple theoretical models predict single-exponential relaxation kinetics (Bicout and Szabo, 2000). More complex responses are expected for rougher surfaces. Experimentally, the folding of a few proteins has been described in terms of stretched exponentials (Sabelko et al., 1999; Morozova-Roche et al., 1999; Leeson et al., 2000). However, the overwhelming majority of protein-folding kinetics has been analyzed in terms of exponential decay(s) (Bilsel and Matthews, 2000), even at temperatures as low as −15°C (Gillespie and Plaxco, 2000). The complex, multiexponential character of DHFR refolding was initially reported by Touchette et al. (1986), and different models have been proposed in terms of sequential or parallel pathways of folding (Touchette et al., 1986; Frieden, 1990; Jennings et al., 1993). Multiexponential refolding comprising four exponentials has also been reported for staphylococcal nuclease (Walkenhorst et al., 1997) and the α-subunit of tryptophan synthase (Bilsel et al., 1999). Recently, five exponentials were reported for the refolding of β-lactoglobulin (Kuwata et al., 2001). The folding of DHFR in the presence of MTX is clearly not described by a stretched exponential (Fig. 5). Rather, the distributions obtained by the MEM for DHFR folding exhibit considerable structure (Fig. 6), indicating that the description of DHFR folding requires a discrete set of folding intermediates. The characterization of the energy surface underlying the implied intermediates and transition states requires additional kinetic and thermodynamic data. The MTX binding site is at the interface of the two domains in the structure of DHFR, and the observed kinetics may represent domain/domain repacking events, as suggested by Wallace and Matthews (2002).

CONCLUSIONS

The interpretation of time-dependent signals in terms of continuous and discrete lifetime distributions permits a comprehensive analysis of measured kinetics. The MemExp software reported here is particularly useful when analyzing kinetics that involve a large number of discrete exponentials. The use of a MEM prior derived from the data can significantly improve the distributed solution obtained for this ill-posed problem. Raw data can be analyzed using an automated multistep protocol, in which standard errors are estimated accurately. The hybrid algorithm and analysis protocol performed well when tested with simulated data corresponding to distributed, distributed-plus-discrete, and discrete kinetics.

When applied to fluorescence measurements of the folding of DHFR in the presence of MTX, MemExp resolved six processes, one more than previously reported. The measurements presented here were fit well by one exponential rise and five exponential decays. A more precise description of the slowest component calls for data at longer times. This improved quantitative characterization of the data is an essential step toward understanding the folding process in detail. Future modeling of the processes that contribute to the measured signal, including folding intermediates and the transitions among them, will need to account for the six processes identified here.

Finally, the analysis developed here could be readily generalized in two ways. Any linear image-data transformation could be interpreted similarly in terms of continuous and discrete descriptions. Also, a maximum likelihood algorithm could be added as an alternative to NLS, improving the analysis of data with noise that is not well approximated as normally distributed.

We thank Per L. Hansen, Sinisa Pajevic, Charles D. Schwieters, and Osman Bilsel for stimulating discussions and the National Science Foundation (grant MCB-0081076) for support.

REFERENCES

Alcala, J. R., E. Gratton, and F. G. Prendergast. 1987. Resolvability of fluorescence lifetime distributions using phase fluorometry. *Biophys. J.* 51:587–596.

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.

Baljzer, Ž., and F. G. Prendergast. 1992. Maximum likelihood analysis of fluorescence data. *Methods Enzymol.* 210:200–237.

Bicout, D. J., and A. Szabo. 2000. Entropic barriers, transition states, funnels, and exponential folding kinetics: a simple model. *Protein Sci.* 9:452–465.

Bilsel, O., J. A. Zitzewitz, K. E. Bowers, and C. R. Matthews. 1999. Folding mechanism of the α-subunit of tryptophan synthase, an α/β barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry.* 38:1018–1029.

Bilsel, O., and C. R. Matthews. 2000. Barriers in protein folding reactions. *Adv. Protein Chem.* 53:153–207.

Brooks, C. L., III, J. N. Onuchic, and D. J. Wales. Taking a walk on a landscape. 2001. *Science.* 293:612–613.

Cornwell, T. J., and K. F. Evans. 1985. A simple maximum entropy deconvolution algorithm. *Astron. Astrophys.* 143:77–83.

Frieden, C. 1990. Refolding of *Escherichia coli* dihydrofolate reductase: sequential formation of substrate binding sites. *Proc. Natl. Acad. Sci. U.S.A.* 87:4413–4416.

Gillespie, B., and K. W. Plaxco. 2000. Nonglassy kinetics in the folding of a simple single-domain protein. *Proc. Natl. Acad. Sci. U.S.A.* 97: 12014–12019.

Götze, W. 1991. Aspects of structural glass transitions. *In* Liquids, Freezing and Glass Transition, Vol. 1. J. P. Hansen, D. Levesque, and J. Zinn-Justin, editors. North-Holland, Amsterdam. 287–503.

Gull, S. F. 1989. Developments in maximum entropy data analysis. *In* Maximum Entropy and Bayesian Methods. J. Skilling, editor. Kluwer Academic, Norwell, MA. 53–71.

Hagen, S. J., J. Hofrichter, and W. A. Eaton. 1995. Protein reaction kinetics in a room-temperature glass. *Science.* 269:959–962.

Hogiu, S., M. Enescu, and M.-L. Pascu. 1997. Dynamic and thermodynamic effects of glycerol on bovine serum albumin in aqueous solution: a tryptophan phosphorescence study. *J. Photochem. Photobiol. B.* 40: 55–60.

Istratov, A. A., and O. F. Vyvenko. 1999. Exponential analysis in physical phenomena. *Rev. Sci. Instrum.* 70:1233–1257.

Jennings, P. A., B. E. Finn, B. E. Jones, and C. R. Matthews. 1993. A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. *Biochemistry.* 32:3783–3789.

Johnson, J. B., D. C. Lamb, H. Frauenfelder, J. D. Müller, B. McMahon, G. U. Nienhaus, and R. D. Young. 1996. Ligand binding to heme proteins. VI. Interconversion of taxonomic substates in carbonmonoxy-myoglobin. *Biophys. J.* 71:1563–1573.

Kumar, A. T. N., L. Zhu, J. F. Christian, A. A. Demidov, and P. M. Champion. 2001. On the rate distribution analysis of kinetic data using the maximum entropy method: applications to myoglobin relaxation on the nanosecond and femtosecond timescales. *J. Phys. Chem. B.* 2001: 7847–7856.

Kuwata, K., R. Shastry, H. Cheng, M. Hoshino, C. A. Batt, Y. Goto, and H. Roder. 2001. Structural and kinetic characterization of early folding events in β-lactoglobulin. *Nature Struct. Biol.* 8:151–155.

Lambright, D. G., S. Balasubramanian, and S. G. Boxer. 1993. Dynamics of protein relaxation in site-specific mutants of human myoglobin. *Biochemistry.* 32:10116–10124.

Lavalette, D., C. Tetreau, J.-C. Brochon, and A. Livesey. 1991. Conformational fluctuations and protein reactivity: determination of the rate constant specturm and consequences in elementary biochemical processes. *Eur. J. Biochem.* 1956:591–598.

Leeson, D. A., F. Gai, H. M. Rodriguez, L. M. Gregoret, and R. B. Dyer. 2000. Protein folding and unfolding on a complex energy landscape. *Proc. Natl. Acad. Sci. U.S.A.* 97:2527–2532.

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.

Morozova-Roche, L. A., J. A. Jones, W. Noppe, and C. M. Dobson. 1999. Independent nucleation and heterogeneous assembly of structure during folding of equine lysozyme. *J. Mol. Biol.* 289:1055–1073.

Phillips, D. L. 1962. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.* 9:84–97.

Plaza del Pino, I. M., A. Parody-Morreale, and J. M. Sanchez-Ruiz. 1997. Maximum entropy analysis of kinetic processes involving chemical and folding–unfolding changes in proteins. *Anal. Biochem.* 244:239–255.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in Fortran. 2nd ed. Cambridge University Press, Cambridge, U.K.

Provencher, S. W. 1982a. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comp. Phys. Comm.* 27:213–227.

Provencher, S. W. 1982b. CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comp. Phys. Comm.* 27:229–242.

Provencher, S. W. 1992. Low-bias macroscopic analysis of polydispersity. *In* Laser Light Scattering in Biochemistry. S. E. Harding, D. B. Sattelle, and V. A. Bloomfield, editors. The Royal Society of Chemistry, Cambridge, U.K. 92–111.

Puetter, R. C., and A. Yahil. 1998. The pixon method of image reconstruction. *In* Astronomical Data Analysis Software and Systems VIII. D. M. Mehringer, R. L. Plante, and D. A. Roberts, editors. ASP Conference Series Vol. 172. ASP, San Francisco. 307–316.

Sabelko, J., J. Ervin, and M. Gruebele. 1999. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 96:6031–6036.

Skilling, J. 1989. Classic maximum entropy. *In* Maximum Entropy and Bayesian Methods. J. Skilling, editor. Kluwer Academic, Norwell, MA. 45–52.

Steinbach, P. J., A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B. R. Cowen, D. Ehrenstein, H. Frauenfelder, J. B. Johnson, D. C. Lamb, S. Luck, J. R. Mourant, G. U. Nienhaus, P. Ormos, R. Philipp, A. Xie, and R. D. Young. 1991. Ligand binding to heme proteins: connection between dynamics and function. *Biochemistry.* 30:3988–4001.

Steinbach, P. J., K. Chu, H. Frauenfelder, J. B. Johnson, D. C. Lamb, G. U. Nienhaus, T. B. Sauke, and R. D. Young. 1992. Determination of rate distributions from kinetic experiments. *Biophys. J.* 61:235–245.

Steinbach, P. J. 1996. Two-dimensional distributions of activation enthalpy and entropy from kinetics by the maximum entropy method. *Biophys. J.* 70:1521–1528.

Tian, W. D., J. T. Sage, V. Šrajer, and P. M. Champion. 1992. Relaxation dynamics of myoglobin in solution. *Phys. Rev. Lett.* 68:408–411.

Titterington, D. M. 1985. General structure of regularization procedures in image reconstruction. *Astron. Astrophys.* 144:381–387.

Touchette, N. A., K. M. Perry, and C. R. Matthews. 1986. Folding of dihydrofolate reductase from *Escherichia coli*. *Biochemistry.* 25: 5445–5452.

Walkenhorst, W. F., S. M. Green, and H. Roder. 1997. Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. *Biochemistry.* 36:5795–5805.

Wallace, L. A., and C. R. Matthews. 2002. Highly divergent dihydrofolate reductases conserve complex folding mechanisms. *J. Mol. Biol.* 315: 193–211.