Critical Evaluation of Kinetic Models for Bacteriorhodopsin Photocycles

Richard I. Shrager

Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892

Richard W. Hendler*

Laboratory of Cell Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

Received: October 24, 2002; In Final Form: December 10, 2002

How can the field of bacteriorhodopsin (BR) photocycle kinetics come up with so many conflicting models? A brief description of the common features of various methods is given, along with the hazards of their use. We describe how our methods are chosen to avoid those hazards. Finally, a set of requirements is suggested for a good model of the BR photocycle.

1. Plan of the Work

In 1982, we introduced a least-squares method based on singular value decomposition (SVD) for the deconvolution of overlapping spectra from 3D data.[1] Later, we made some refinements,[2] and we recently published a tutorial.[3] This approach, and related linear algebra techniques, have become widely used in a variety of applications, including kinetics. The use of such techniques is not straighforward and is subject to restrictions at every turn. The present paper is devoted to an evaluation of linear algebra methods, most involving SVD at some stage, that have been used to model the kinetics of the bacteriorhodopsin (BR) photocycle. There are two central questions: Why do similar techniques applied to similar data yield such a variety of models? What are the requirements for a proper solution?

Published studies of BR kinetics share a common formalism that will be explained here in brief. A detailed discussion of this formalism and its problems is given in the Supporting Informationof ref 4. In section 2, we define the relevant matrices and outline their relations. In section 3, we discuss the various approaches that have been used and give our assessment of their difficulties. In section 4, we explain our own model,[4] which is unidirectional with two or more parallel cycles, and the restricted data that it currently describes. In section 5, we list some published results of the methods under discussion. In section 6, we list other experimental evidence in support of the existence of parallel cycles. Finally, in section 7, we discuss the critical features that any model must have in order to be a serious candidate for BR kinetics.

2. Matrices in Question

2.1. Data. In all of the work we will discuss, the same general kind of data is used. A set of wavelengths w_1, w_2, . . . and a set of measurement times t_1, t_2, . . . are chosen. The experimental data consist of several spectra, which may be any combination of optical, resonance Raman, or infrared, measured at times t after the initiation of the photocycle by a laser pulse. These data are stored in the columns of a matrix A. Each column of A is a mixture of spectra of the various states (species) of BR. The amplitude of each species' spectrum at each time is proportional to the concentration of that species. Each element of A is the measured spectral value (e.g. $A_{3,5}$ is the value at wavelength w_3 and time t_5).

There are two versions of A: the data, represented by the symbol A, and any estimate of A, represented by hat notation $\hat{A}$. All other matrices described below will be estimates, so hat notation for those matrices will be omitted for simplicity.

2.2. Model. The goal is to produce a model of the BR photocycle. These models are represented by a system of linear ordinary differential equations (ODEs) of the first degree, which may be expressed in matrix terms as

$$dY(t)/dt = JY(t),\ Y(0) = y^{[0]} \tag{1}$$

The matrix Y contains the time courses of each reacting species in its rows. There is one row of Y for each species in the reaction, and one column of Y for each time t_i. The particular order of species is determined by the model (e.g., each model cited in Table 1 should have an ODE for each species listed there). The only possible exception is the ground-state BR, which may be represented by an ODE or calculated from the conservation of total BR.

The matrix Jacobian J contains all of the rate constants of the system so that J and $y^{[0]}$ completely determine the system. The off-diagonal elements $J_{i,j}$, $i \neq j$, are the rate constants for species j converting to species i. The on-diagonal elements $J_{i,i}$ are the negative sums of all of the rate constants leading from species i, including intrasystem conversions and sink rates if any.

The column vector of initial conditions $y^{[0]}$ consists of mostly zeros, with a few presumed initial states for which $y_i^{[0]} > 0$. To have parallel cycles, there must be some nonzero $y^{[0]}_i$ values whose subsequent species are completely separate from those in other cycles. These initial states should be the excited states of BR just after the laser flash, and all subsequent species in the cycle should be included in the model, although limitations of time resolution often require the omission of earlier species.

2.3. Time-Course Decomposition. If we have estimated the proper model, J and $y^{[0]}$, then the resulting time-course estimates

* Corresponding author. E-mail: rwh@helix.nih.gov.

10.1021/jp0273070 CCC: $25.00
Published on Web 01/22/2003

TABLE 1: Partial List of Models from Recent Papers in Order of Publication[a]

9	at pH between 5 and 8.6: $L \rightleftharpoons M \rightleftharpoons N \rightarrow O \rightarrow BR$
11	$L \rightleftharpoons M_1 \rightarrow M_2 \rightleftharpoons N \rightarrow O \rightarrow BR$ and $L \rightarrow N$ see also Table 1 in ref 11
12	$K \rightleftharpoons L \rightleftharpoons X \rightleftharpoons M \rightleftharpoons N \rightleftharpoons O \rightarrow BR$
13	at low pH: $L \rightleftharpoons M_1 \rightleftharpoons M_2 \rightleftharpoons N \rightleftharpoons O \rightarrow BR$ at high pH: $L \rightleftharpoons M_1 \rightleftharpoons M_1' \rightleftharpoons M_2' \rightleftharpoons N' \rightleftharpoons O' \rightarrow BR$
14	$K \rightleftharpoons L \rightleftharpoons M \rightleftharpoons N \rightleftharpoons O \rightarrow BR$
15	$K \rightarrow L \rightarrow M_1 \rightarrow M_1' \rightarrow BR_1' \rightarrow BR$ and $M_1 \rightleftharpoons M_2 \rightarrow M_2' \rightarrow BR_2' \rightarrow BR$ and $BR_2' \rightleftharpoons N$
6	rapid cycle: $K^r \rightleftharpoons L^r \rightleftharpoons M_1^r \rightleftharpoons M_2^r \rightarrow BR$ slow cycle: $K^s \rightleftharpoons L^s \rightleftharpoons M_1^s \rightleftharpoons M_2^s \rightarrow BR$
4	fast cycle: $L_f \rightarrow M_f \rightarrow N \rightarrow O \rightarrow BR$ slow cycle: $L_s \rightarrow M_s \rightarrow BR$
16	$K \rightleftharpoons L \rightleftharpoons M_1 \rightleftharpoons M_2 \rightleftharpoons N \rightleftharpoons O \rightarrow BR$

[a] Reference numbers are in the left column.

Y can be generated directly using the procedure given in the Supporting Information of reference 4. Then, a least-squares estimate of the data matrix A can be computed from the equation

$$\hat{A} = DY \tag{2}$$

where $D_{i,k}$ is the value of the kth species spectrum at the ith wavelength. The spectra D are deduced from equation 2 by

$$D = AY^{+} \tag{3}$$

where Y^{+} is the Moore−Penrose pseudo-inverse of Y.

In some methods, J is not the starting point from which the other matrices are deduced. Instead, one might attempt to estimate D directly from A by a process currently called self-modeling (SM)[5,6] and work one's way back to J. In that case, one might employ equation 2 in the reverse direction:

$$Y = D^{+}A \tag{4}$$

where D^{+} is the psuedo-inverse of D. SM, as currently used, has inherent inaccuracies and is limited in the number of spectra that it can resolve. Because of these limitations, SM has been applied only to mutant data with fewer intermediates. To date, there are no published results in which SM has been applied to the entire wild-type BR system.

2.4. Exponential Decomposition. Systems described by equation 1 can be modeled (with few exceptions) by sums of exponentials:

$$\hat{A} = CX \tag{5}$$

where each ith row of X contains a single decaying exponential function $\exp(-k_i t_j)$, and the corresponding column of C contains the least-squares coefficients of that exponential so that CX is a least-squares approximation of A.

Although C and X are not sufficient to figure the system out completely, these matrices do contain some useful information. The fitted constants k_i are the macroscopic rate constants (i.e., the observable rates of decay, which would be actual rate constants if the system were irreversible, and the corresponding columns of C would be the "spectra" associated with that decay, as explained below). But reversibility or no, if the exponential approximation in equation 5 is accepted as a working hypothesis, then any J we select must have eigenvalues that agree with the fitted rate constants k_i.

Equation 5 serves as the starting point of some methods,[4,7,8] ours among them. First, a series of exponential fits is used to determine the minimal representation of A by exponentials. Note that this exponential phase assumes only that the kinetic steps are first order. It does not presume any particular model beyond that. Then, with the required exponentials in hand, one selects from only those (usually numerous) models that produce the required exponentials.

The columns of the matrix factor C in equation 5 would be actual species spectra if we were studying a mixture of species that were decaying independently. The relations $Y = X$ and $D = C$ would both hold in such a case. If the model were sequential (one branch) and unidirectional with widely spaced rate constants in decreasing order, say $k_{i+1}/k_i < 0.2$, then the columns of C would be approximate transition difference spectra from one species to the next. No one in the field believes that either of these models describes the BR photocycle. Still, the "spectra" in the columns of C may be sufficient to give hints about the mechanism, but any such hints would have to be verified by more quantitative methods.

2.5. Interrelations. Our matrices so far: data A, Jacobian J, initial conditions $y^{[0]}$, spectra D, time courses Y, exponentials X, and coefficients C, are intricately related so that by knowing as few as two of them the others can be fitted or simulated provided certain conditions, namely, accuracy and proper conditioning, are met. It is the violations of those conditions that we feel are responsible for the current confusion in the field.[4] Some methods involve additional matrices (e.g., the SVD of A and the eigenvectors of J) that are mainly of procedural interest and will not be discussed in this paper. However, we will recount some of the methods in the next section, along with our reasons for suspecting that they do not always work well in the BR system.

3. Methods in Question

In the methods and studies we will discuss, we cannot know what if anything went wrong without access to the data and specifics of the calculations, which are rarely present in the publications. Yet we know that of all the models of wild-type BR that are applicable to standard conditions of temperature and pH at most one such model can be right. So something is going wrong in most cases. What follows in this section is our assessment of the possible problems associated with the various methods.

The methods that have been applied to BR kinetics can be divided into three classes: class 1, class 2, and hybrid, as described in Hendler and Shrager.[4] The use of matrix methods in BR kinetics was significantly advanced when the work of Ames and Mathies[9] was analyzed by Nagle,[10,11] who also specified some requirements for fitting models. Since then, several investigators have used SVD and related methods,[6,12−16] assured by Nagle's work that the task, if not the BR kinetic mechanism, was at last understood.

3.1. Methods of Class 1. The methods of class 1 start with an assumed model or set of models (J and $y^{[0]}$), adjusting the parameters within those models to fit the data in the least-squares sense. Nagle[10] showed that the methods of class 1 require data at three or more temperatures to resolve all of the rate constants in J, assuming that the spectra are constant with temperature. Nagle's proposed treatment is remarkable in that his set of models includes all models that can be expressed with a given size of J. That is, aside from the number of intermediates, nothing need be assumed about J, although any constraints on the rate constants (e.g., that some of them must be zero) would certainly help. The number of intermediates is usually estimated in advance (e.g., by the successive fitting of exponentials[7] or by SVD[1]). This generality is clearly an appealing feature. The

downfall of such methods, we feel, is the requirement that relative spectral shape be constant with temperature. Spectral change with temperature is commonly observed, as seen , for example, in the work of Demchenko.[17] Although that work is mainly in the ultraviolet range, effects are also seen in the visible (Figure 4.12 in ref 17). This concern is particularly important in the case of BR. The retinal chromophore is very sensitive to its environment. The wavelength of maximum absorbance of protonated retinal Schiff base analogues is at 446 nm,[18] but the peak absorbances of the intermediates cover a range of 412 to 640 nm. These changes reflect the influence of the environment on the absorption characteristics of retinal. It is reasonable to suspect that temperature can alter some aspects of the retinal environment. Such changes can be minimized by restricting temperature to a very narrow range, but this leads to experiments that resemble experiments at one temperature, effectively reintroducing the problems described by Nagle.[10] However, if temperature affects the species' spectra, a wide range of temperature will exaggerate those spectral changes. There may be no practical middle ground. Despite these considerations, it is not certain that thermal variation is a fatal flaw. Nagle et al.[19] have presented evidence that the spectra of at least some intermediates are stable within the temperature range used. All told, however, the derivation of a generally acceptable model by method 1 remains elusive.

3.2. Methods of Class 2. The methods of class 2 start from the knowledge of the spectra (columns of *D*) or the time courses (rows of *Y*), from which *J* and $y^{[0]}$ can sometimes be computed. Nagle[10] calls these direct methods, and he points out that they require data at only one temperature, thus avoiding potential complications due to temperature variation. Again, aside from the number of intermediates, nothing is assumed about *J*. As in method 1, some method is used to estimate the number of species in advance.

Several things interfere with class 2 methods. A knowledge of *D* or *Y* is not available directly from data in the visible range, although there is the possibility (as yet unfulfilled) that vibrational spectroscopy could provide adequate data of this type.[10] Often, *D* is inferred from the data *A* by an approximate process called self-modeling (SM). The seminal paper on SM by Lawton and Sylvestre[5] describes the resolution of two spectra. As the figure on page 627 of that reference indicates, the result is not a unique solution. Rather, one is faced with a range of solutions within which the actual solution must lie. Every subsequent use of SM[6,19,20] has admitted some form of doubtful procedure producing a range of possible spectra. If absolute spectra are sought, then the regions of acceptability can be narrowed considerably by constraining the absorbances to positive values, but this does not necessarily produce proper spectra. (See more about constraints below.) Biases from the SM-deduced spectra propagate into the other matrices, making the conclusions suspect. Furthermore, if the resulting *D* matrix is ill-conditioned (see below), then errors in *D* and *A* will tend to be greatly magnified in *Y*.

To make matters worse, even an exact knowledge of *D* does not guarantee that class 2 methods starting from *D* will work. One difficulty is that the number of spectra in *D* may be less than the number of species. It is commonly believed that there are at least two forms of the M intermediate and that the spectra are very similar if not identical. For example, most of the models in Table 1 have at least two forms of M. If two or more similar spectra are combined as one in *D*, then eq 4 will produce a composite time course for all species having that spectrum. The composite time course is inherently impossible to interpret on the basis of spectral information alone because there are an infinite number of ways to distribute that single composite row of *Y* into the two or more time courses of the actual species. Some guiding assumptions about the model are required. Such assumptions are the essence of the hybrid methods discussed below.

Another difficulty of class 2 methods is ill conditioning, which is the sensitivity of the results to small changes in input. In class 2 methods, *D* may be ill-conditioned so that eq 4 produces inaccurate time courses. Such ill conditioning of *D* comes about when some columns of *D* are nearly linearly dependent (e.g., when two or more spectra resemble each other). For example, if the spectra of the varous M intermediates are regarded as very similar rather than identical so that each M spectrum is represented separately in *D*, then the resulting *Y* might well be uselessly inaccurate. We are thus presented with a procedural dilemma. If we merge similar spectra into a single column of *D*, then we must somehow unmerge the corresponding row of *Y*. However, if we keep the similar spectra in separate columns of *D*, then *D* is ill-conditioned. Even without such diabolical near-duplicates, the visible spectra of BR and N are believed to have their respective maxima between 560 and 570 nm, rather close compared to their widths, so the condition of *D* in the BR system is not a trivial consideration.

3.3. Hybrid Methods. Hybrid methods use information from both the model and the species spectra. The motivation behind hybrid methods is to do away with the three-temperature requirement for pure class 1 methods while also avoiding the uniqueness and conditioning problems of class 2 methods. Most of the methods in use for BR kinetics have some hybrid aspect to them simply because the pure methods are impractical. It is not clear that one can have the best of both worlds, but experience with our own hybrid method (described below) has been encouraging. It is clear, however, that using biased quantities (e.g., the results of SM) in a hybrid method, as in any method, will still bias the results. All hybrid methods are not created equal.

3.4. Initial Conditions. In class 1 and hybrid methods, $y^{[0]}$ must be chosen at some point, perhaps repeatedly. When the apparatus is too slow to observe some early intermediate, say everything before L, it may be tempting to use L as the nonzero initial condition. This is not the proper way to handle rapid, unobservable intermediates. Rather, one should postulate some initial state for each subcycle and let the equations determine the concentrations at observation times *t*.

3.5. Constraints. One tool of potential value in estimation methods is the use of constraints. There are natural constrants in many of the quantities under consideration. In particular, the following quantities are expected to be nonnegative:

(1) *k*, the macroscopic constants in the expressions $\exp(-k_i t_j)$;

(2) $J_{i,j}$, the off-diagonal elements of *J*, which are the rate constants of the system;

(3) $y^{[0]}$, the initial concentrations;

(4) *Y*, the time-dependent concentrations; and

(5) *D*, the absorbances, if the method is designed to find absolute spectra. As with any tool in modeling, there are hazards in the use of constraints. For example, in a search for best-fit parameters, a method may temporarily try negative parameters but finally wind up with positive results. However, if constraints are imposed, then the search may be blocked from proceding to the best solution and may remain stuck at a local minimum created by the constraints. This possibility can be lessened by fitting several times with a variety of first estimates. It can be

eliminated completely by trying every constrained fit without constraints as well.

Sometimes it is best to tolerate some negative values at the solution, even when you know that they should not be there. For example, suppose we use eq 3 or some other equation that propagates noise into *D*, which is set up to produce absolute spectra. If we insist that the model produce no negative *D* whatever, in effect we force some regions of *D* to have all positive errors, where without the constraints they would have both positive and negative errors. This is a classic recipe for biased results. Constraints are useful in some cases and deceptive in others.

4. Some Reported Results

Table 1 contains a partial list of models that have been proposed by various groups using methods of the types described above. The most disturbing aspect of this list is that there is no trend, no discernible convergence or refinement. Starting in the 1990s, many papers have been heavily influenced by the work of Nagle,[10] whose work on the properties and limitations of class 1 and 2 methods prevented the field from wasting much time on impossible modeling procedures. However, his subsequent modeling efforts[11] showed that pure methods of class 2, using the best available data at that time, were not going to produce clear-cut answers. In his Table 1, several models are tested at five pH values about 1 unit apart. The best model at each pH was different from the best models at all other pH's. By "different", we are not talking about merely the values of some rate constants here but rather the very existence of pathways.

Lozier et al.[12] used the class 1 approach with data at four temperatures. As a result (we believe), their deduced spectra show cross contamination (i.e., double peaks where there should be only one). In their conclusion, they state that "limitations imposed by the optical spectra make it unlikely that further details of the cycle can be resolved simply by fitting more accurate or more complete visible absorption data." Our own feeling is that their approach used too much data, as was explained above and will be illustrated below. The appeal of model-independent deduction drew investigators into SM, essentially in an effort to use class 2 or hybrid methods. Indications are that SM, as it is now conceived, has an inherent lack of accuracy that cannot be overcome by constraints or by any foreseeable technique. In addition, current SM methods are severely limited in the number of spectra they can hope to resolve. The most optimistic users of SM claim to resolve four spectra, but most investigators are convinced that the BR system has five or more spectra. Thus, SM does not hold out much hope of resolving a full system.

In an effort to get something useful from SM, Zimanyi et al.[6] used a mutant system that did not generate the N or O species. Despite the obvious differences between their mutant system and the wild type, their proposed mechanism bears some similarity to ours, but in the absence of the N and O species, those spectra cannot be resolved. Also, it is not certain that the resolved mutant spectra are identical to those in the wild type. Zimanyi et al.[6] conclude, based on "Work in Progress on the SVD-SM Analysis of Data for the Wild-Type BR" (unpublished at this writing), that their parallel model "is not a feature of the wild-type photocycle." Since we have not seen that evidence, we reserve judgment.

Hessling et al.[14,15] used Fourier transform infrared (FTIR) spectra to augment the information in the visible spectra. In 1993,[14] they proposed a reversible sequential model for wild-type BR with one form each of K, L, M, N, O, and BR. But when they reexamined the system[11] (see Table 1), they proposed a branched model with four forms of M, no O, and three forms of BR. Using FTIR makes a lot of sense because FTIR could confirm or expand the findings in the visible range. However, there could be difficulties. We have also looked at FTIR data, and we have found that events that are prominent in FTIR can be weak or absent in the visible region. This presents a challenge in the interpretation of the union of such data. For example, the most prominent feature in Figures 3 and 4 of ref 15 is a trough at 1527 cm^{-1}. We find that this trough remains kinetically active long after major activity in the visible has ceased.

Aside from the problem of FTIR interpretation, Hessling's methods are SM-based. Recall that SM is limited to four spectra. Hessling gets around this limitation by including data at several pH's. He uses information from one pH to reduce the dimensionality of the problem at the next pH, which could work, given certain assumptions about constancy from one pH to the next. Then again, these assumptions could be as faulty as the assumptions of constancy with temperature in class 1 methods.

The work of van Stokkum and Lozier[16] returns to a class 1 method and deduces a specific model (see Table 1) with specific rate constants. We checked their model against our own data at room temperature and neutral pH using eq 3 to generate the species spectra. We found several problems. Most glaring was the interchange of the N and O spectra. We also found that the N spectrum contained M as well as O and that the M_1 and M_2 spectra both contained a substantial BR peak in addition to the expected M peak.

In summary, there have been several ingenious methods designed to deduce a unique mechanism for the BR photocycle. Each has its flaws. SM does not produce unique answers and is dependent on the rank of the problem (i.e., the number of linearly independent spectra) being four or less. A natural way around this is to exploit some kind of invariance over several pH values, as Hessling et al.[11] suggest. This is also the way to get a unique solution from the methods of class 1 using several temperatures, as Nagle[10] suggests. The inclusion of FTIR data is yet another extension designed to supplement visible spectra. Each extension, whether by pH, temperature, or wavelength, is a promising innovation, yet perilous. If the added data introduce added phenomena not in accord with invariance assumptions, then the result will likely be biased. We feel that this has happened in every attempt to extend the data so far.

5. Our Model

In reaction to the difficulties described above, we have scaled back the complexity of both the model and the data. With respect to the data, we restricted our attention to room temperature and neutral pH to avoid any variation in spectra or rates that pH and temperature variation might induce. With respect to the model, we decided to concentrate on forward rates exclusively. We recognize that the true model may have some reverse rates. If those reverse rates are small enough, then the true model will agree with ours quite closely. Even if the reverse rates are significant, our model could still resemble the true model sufficiently to guide further modeling efforts. However, a reversible model, being more complex, will have to explain significantly more of BR kinetics than our current model does.

Forward-rate models have a pleasant property: they can all be reduced to a union of forward sequential models (single path, no reverse rates) as we explained in detail.[4] The first phase of our procedure is to establish an exponential fit of *A*. Since there are only a finite and manageable number of forward sequential

models that conform to the fitted exponentials, they can all be scanned for signs of validity. A valid sequence should use some of the fitted exponentials and no others and should produce reasonable difference spectra. The final model (i.e., the union of some of the reasonable sequences) should use all of the fitted exponentials and should produce reasonable absolute spectra. Through this process, we found only one model that met all of the criteria (see the entry for reference 4 in Table 1). Although our method requires some spectral information to avoid the three-temperature requirement derived by Nagle,[10] we do not want to specify questionable spectra. Here again, our approach is to simplify. We use the least specific kind of visible spectral information, namely, that each species' absorbance is low far from its peak absorbance.

Our result, listed in Table 1, is two parallel subcycles. The quality of the fit to A is exactly the same as that of any other model (forward or reversible) using the same macroscopic rate constants. The deduced absolute visible spectrum for each species has its peak at the expected wavelength, and the species appear in the expected order. Using the same methods and an exhaustive search procedure, we could produce many forward rate models that fit A equally well, but only one such model produced the right spectra.

6. Evidence for Parallel Cycles

An independent criterion for a correct model is provided by the fact that the intensity of the actinic flash influences the ratio of the fast form of M (M_f) to the slow form of M (M_s). See, for example, Hendler et al.[20] This eliminates most of the models that have been proposed, namely, any model with all forms of M in a single cycle (i.e., with a single nonzero initial condition in $y^{[0]}$). The reason for the elimination is that all such models are influenced by actinic light intensity in only one way, namely, by changing the fraction of BR cycling. This changes the scale of the deduced concentrations but not their ratios at any given time. That is, in a one-cycle model, all time courses in the sequence will be scaled by the same factor when the actinic light intensity changes. It might seem that one-cycle models can be salvaged by providing a mechanism whereby actinic light intensity can influence the rate constants, but such a mechanism would have to be quite novel. The individual time courses for M_f and M_S do not change with actinic light, except for the scale. That is, only the M_f/M_s ratio changes. It is hard to see how a one-cycle model could accommodate this behavior. We are aware of only one publication that proposes a model of this type.[21] We are not convinced that the conclusion reached in that paper has been adequately demonstrated for a number of reasons: (1) Time-course data were measured at only two wavelengths. (2) Time profiles for the intermediates were obtained from these two wavelengths using the assumed kinetic model:

$$\rightarrow M_1 \rightarrow M_2 \rightleftharpoons N^{(-1)} \rightleftharpoons N^{(0)} \rightarrow BR$$

This model is no longer proposed for BR kinetics. (3) The only quantitative data in Table 1 shows a minimal effect of actinic light in which the mole fraction of the faster form of M decreased from 0.87 to 0.70 in going from low to high intensity. (4) The time constant for the slow form of M did change from 70 to 55 ms. (5) No fitting statistics are shown. The only models left standing from Table 1 are (6) for a mutant system without N or O and (4) for the wild type.

Although only two forms of M are commonly seen in the photocycle of wild-type BR, there are two other kinetic forms of M that are present under other circumstances. A species with $\tau \approx 10$ ms becomes prominent in mutant BR.[22] The perturbation of purple membranes with dilute Triton X-100 causes the appearance of a fourth species with $\tau \approx 70$ ms.[23] All slower species can be converted to the fastest species M_f by increasing the hydrophobicity of the membrane with decane. Conversely, M_f can be converted to slower forms by mutating any or all of the four aspartates on the cytoplasmic side, which form a tight ring around the shared inner trimer space.[22] In intact Halobacteria cells under conditions of increased electrochemical potential for protons, the 10-ms and 70-ms species are seen as well as additional M forms.[23,24]

A major difference between M_f and the other three M species (with τ's of 5, 10, and 70 ms) is that the decay path of M_f includes the O intermediate whereas the other three M species decay directly to the ground state.[21,25,26] The relative amounts of these four M species are regulated under a variety of changes of external conditions. It has long been known that for low actinic light M_f is the predominant species and that an increase in light intensity favors the participation of M_s.[21,27] Increases in pH dramatically increase the participation of M_f while decreasing that of M_s.[22] In intact Halobacteria cells, an increase of the electrochemical potential for protons decreases the participation of M_f whereas the addition of uncoupler makes M_f the predominant species.[24] In addition to the above observations relating to the intermediate M, evidence for at least two forms of L exists.[4,28] No single continuous photocycle has been proposed that could account for all of the above observations.

7. Critical Features for Any Model to Be Seriously Considered

As stated in section 1, the fundamental matrix equation that relates a set of acquired spectral/kinetic data to a particular model is $A = DY$. One perversity of such an equation is that it has no knowledge of the processes being modeled and can always provide a least-squares estimate for one of the matrices given the other two. If, say, A is known (observed) and Y is deduced, albeit incorrectly, then $D = AY^{+}$ will be a least-squares estimate of the spectra predicated with the wrong model. Similarly, if D is deduced by a biased form of SM, $Y = D^{+}A$ will be a biased estimate of the time courses. No warnings will be issued in the process. It is up to the investigator to know when a given solution is acceptable. Fortunately, there is a lot of information to go on.

(1) Within the time resolution of the measurements, D should contain absolute spectra for the basic species BR, K, L, M, N, and O, with maxima close to their published values and decreasing absorptions as the wavelength varies away from the peak values.

(2) The time courses in Y should be consistent with the macroconstants that have been observed in many laboratories throughout the world.

(3) The photocycle(s) specified in J should link the states in the accepted sequence: BR, K, L, M, N, O, with possible branches, shunts, and parallelisms.

(4) The photocycle(s) should contain at least two forms of M and should accommodate the ability of the M species to change proportion as the conditions described above change.

Only the parallel unidirectional model[4] has been shown to meet all of these criteria.

A thermodynamic argument has been raised against a model with essentially irreversible steps. If one interprets irreversibility to mean that the reverse rates are exactly zero, then the ΔG for such changes would be infinite. However, if the forward/reverse

rate constant ratio is, say, 1000/1, then ΔG is 4140 calories. The amount of energy in one photochemical equivalent (i.e., 1 einstein) at 532 nm is 53 741 calories. At a quantum efficiency of 0.6, 32244 calories are available. This is sufficient for more than seven steps with forward/reverse ratios of 1000/1. The free energy necessary to maintain an electrochemical gradient of 300 mV is 3463 calories, and a 200-mV gradient can be maintained with 2309 calories. There is no obvious reason that energy needs to be conserved by having some or many highly reversible low-energy steps.

References and Notes

(1) Shrager, R. I.; Hendler, R. W. *Anal. Chem.* **1982**, *54*, 1147−1152.
(2) Shrager, R. I. *Chemom. Intell. Lab. Syst.* **1986**, *1*, 59−70.
(3) Hendler, R. W.; Shrager, R, I. *J. Biochem. Biophys. Methods* **1994**, *28* 1−33.
(4) Hendler, R, W.; Shrager, R. I. *J. Phys. Chem. B* **2001**, *105*, 3319−3328. See Supporting Information at http://pubs.acs.org.
(5) Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 617−633.
(6) Zimanyi, L.; Kulcsar, A.; Lanyi, J. K.; Sears, D. F., Jr.; Saltiel, J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4414−4419.
(7) Xie, A. H.; Nagle, J. F.; Lozier, R. H. *Biophys. J.* **1987**, *51*, 627−635.
(8) Szundi, I.; Lewis, J. W.; Kliger, D. S. *Biophys. J.* **1997**, *73*, 688−702.
(9) Ames, J. B.; Mathies, R. A. *Biochemistry* **1990**, *29*, 7181−7190.
(10) Nagle, J. F. *Biophys. J.* **1991**, *59*, 476−487.
(11) Nagle, J. F. *Photochem. Photobiol.* **1991**, *54*, 879−903.
(12) Lozier, R. H.; Xie, A.; Hofrichter, J.; Clore, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3610−3614.
(13) Zimanyi, L.; Varo, G.; Chang, M.; Needleman, R.; Lanyi, J. K. *Biochemistry* **1992**, *31*, 8535−8543.
(14) Hessling, B.; Souvignier, G.; Gerwert, K. *Biophys. J.* **1993**, *65*, 1929−1941.
(15) Hessling, B.; Herbst, J.; Rammelsberg, R.; Gerwert, K. *Biophys. J.* **1997**, *73*, 2071−2080.
(16) van Stokkum, I. H. M.; Lozier, R. H. *J. Phys. Chem. B* **2002**, *106*, 3477−3485.
(17) Demchenko, A. P. *Ultraviolet Spectroscopy of Proteins*; Springer-Verlag: Berlin, 1981; Chapter 4.
(18) Hamm, P.; Zurek, M.; Roschinger, T.; Patzelt, H.; Oesterhelt, D.; Zinth, W. *Chem. Phys. Lett.* **1996**, *263*, 613−621.
(19) Nagle, J. F.; Zimanyi, L.; Lanyi, J. K. *Biophys. J.* **1995**, *68*, 1490−1499.
(20) Zimanyi, L.; Lanyi, J. K. *Biophys. J.* **1993**, *64*, 240−250.
(21) Hendler, R. W.; Dancshazy, Zs.; Bose, S.; Shrager, R. I.; Tokaji, Zs. *Biochemistry* **1994**, *33*, 4604−4610.
(22) Hendler, R. W.; Bose, S. 2002, submitted to *Eur. J. Biochem.*
(23) Joshi, M, K.; Bose, S.; Hendler, R. W. *Biochemistry* **1999**, *38*, 8786−8793.
(24) Hendler, R. W.; Drachev, L. A.; Bose, S.; Joshi, M. K. *Eur. J. Biochem.* **2000**, *267*, 5879−5890.
(25) Eisfeld, J. H.; Pusch, C.; Diller, R.; Lohrmann, R.; Stockburger, M. *Biochemistry* **1993**, *32*, 7196−7215.
(26) Luchian, T.; Tokaji, Zs.; Dancshazy, Zs. *FEBS Lett.* **1996**, *386*, 55−59.
(27) Shrager, R. I.; Hendler, R. W.; Bose, S. *Eur. J. Biochem.* **1995**, *229*, 589−595.
(28) Rodig, C.; Chizov, I.; Weidich, O.; Siebert, F. *Biophys. J.* **1999**, *76*, 2687−2701.