ANNEX

Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 4408–4413, April 1999
Biophysics

Singular value decomposition with self-modeling applied to determine bacteriorhodopsin intermediate spectra: Analysis of simulated data

LÁSZLÓ ZIMÁNYI*†, ÁGNES KULCSÁR*, JANOS K. LANYI‡, DONALD F. SEARS, JR.§, AND JACK SALTIEL§¶

*Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, Szeged, H-6701, Hungary; ‡Department of Physiology and Biophysics, University of California, Irvine, CA 92697; and §Department of Chemistry, Florida State University, Tallahassee, FL 32306-4390

Communicated by Michael Kasha, Florida State University, Tallahassee, FL, February 16, 1999 (received for review October 27, 1998)

ABSTRACT **An *a priori* model-independent method for the determination of accurate spectra of photocycle intermediates is developed. The method, singular value decomposition with self-modeling (SVD-SM), is tested on simulated difference spectra designed to mimic the photocycle of the Asp-96 → Asn mutant of bacteriorhodopsin. Stoichiometric constraints, valid until the onset of the recovery of bleached bacteriorhodopsin at the end of the photocycle, guide the self-modeling procedure. The difference spectra of the intermediates are determined in eigenvector space by confining the search for their coordinates to a stoichiometric plane. In the absence of random noise, SVD-SM recovers the intermediate spectra and their time evolution nearly exactly. The recovery of input spectra and kinetics is excellent although somewhat less exact when realistic random noise is included in the input spectra. The difference between recovered and input kinetics is now visually discernible, but the same reaction scheme with nearly identical rate constants to those assumed in the simulation fits the output kinetics well. SVD-SM relegates the selection of a photocycle model to the late stage of the analysis. It thus avoids derivation of erroneous model-specific spectra that result from global model-fitting approaches that assume a model at the outset.**

A general problem in spectroscopy is the dissection of spectra of mixtures of unknown composition into the spectra of the pure constituents, thereby determining the relative amount of the components. In a typical experiment, many spectra are measured, and the variation of an experimental parameter provides a systematic change in the contribution of the pure components to each mixture spectrum. The spectra are arranged in a matrix so that the experimental parameter varies along one of the dimensions. Various algebraic procedures can be used to determine the number of pure components (equal to the effective rank) of the data matrix and to reduce the random noise content at the same time. Principal component analysis (PCA) yields orthonormal spectral eigenvectors, and the corresponding combination coefficients are determined as dot products between the eigenvectors and the mixture spectra (1). Singular value decomposition (SVD) derives the same orthonormal eigenvectors as well as another orthonormal vector set, which, when multiplied by the singular values, provides the same combination coefficients as does PCA (2).

The condition that fluorescence or absorption spectra have no negative intensities permits their normalization before the analysis, ensuring that the derived spectra of the pure components also are normalized (3). The combination coefficients of the normalized spectra of a rank-two matrix are points along a normalization line, as one coefficient is plotted versus the other. The combination coefficients of the pure spectra are sought on the same line beyond the points corresponding to measured spectra during self-modeling (SM) (3, 4). When three pure forms are present, points defined by the combination coefficients of mixture spectra fall within a triangle on the normalization plane in three-dimensional space. The sides represent two component mixtures, and the vertices represent the pure components, as in a phase diagram (5–11). Once the SM procedure locates the spectra of the pure components, reverse normalization provides their actual amplitude.

We describe an application of SVD-SM (analogous to PCA-SM) to the determination of the spectra of the intermediates in the bacteriorhodopsin photocycle. On light excitation, bacteriorhodopsin (BR), the light-driven proton pump in the cell membrane of *Halobacterium salinarium*, exhibits a series of spectrally distinct metastable intermediates labeled as J, K, L, M, N, and O before returning to the initial state (BR) (for reviews, see refs. 12, 13). Although transitions between the intermediates appear to be first-order reactions, back reactions and/or parallel pathways result in mixtures rather than pure intermediates at all times during the photocycle. Moreover, the intermediate spectra strongly overlap in the visible spectral range. Hence, decomposing the measured difference spectra into difference (and absolute) spectra of the intermediates, and their time-dependent concentrations (kinetics), is a mathematically underdetermined problem (14, 15). Global model fits are in principle capable of determining the spectra and the kinetics simultaneously, but in practice they are hampered by spurious local minima of the optimization routines in the case of noisy data (15). In addition, errors in the models lead to model-specific erroneous spectra optimally adjusted to fit the assumed kinetics sequence. The goal in this and the following paper (16) is to determine accurate model-independent spectra of the intermediates at the outset, thereby defining their time evolution. Relegation of the selection of the best model for the kinetics to the final step should enhance the understanding of the proton transfer mechanism.

RESULTS

Generation of Simulated Data. SVD-SM was tested on simulated data resembling the simpler case of the photocycle of the Asp-96 → Asn (D96N) mutant BR with four spectrally distinct intermediates, K, L, M_1, and M_2, in the submicrosecond to 100-millisecond range. The data were generated by modification of the procedure described in ref. 17. A measured visible absorption spectrum of light-adapted BR was shifted on

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: BR, bacteriorhodopsin; PCA, principal component analysis; PCR, photocycling ratio; SM, self-modeling; SP, stoichiometric plane; SVD, singular value decomposition.

†Parts of this work were done during L.Z.'s visit to Florida State University.

¶To whom reprint requests should be addressed. e-mail: jsaltiel@chemmail.chem.fsu.edu.

the wavenumber scale to provide spectra at appropriate energies for the K, L, and M_1 intermediates. The amplitudes of the new spectra were changed, and their half widths were modified by convolution with Gaussians to yield spectra consistent with previous information. M_1 was shifted by 6 nm to the blue, and its amplitude was decreased by 5% to generate the M_2 spectrum. The two Ms were originally introduced to fit a kinetic model for the wild-type protein (18) and to account for the slight spectral shift of the maximum of measured difference spectra on D96N BR during the rise of the amount of M (19). Time-dependent concentrations of the intermediates were simulated by integration of the rate equations corresponding to the photocycle scheme $K \rightleftharpoons L_1 \rightleftharpoons L_2 \rightleftharpoons M_1 \rightarrow M_2 \rightarrow BR$ plus $M_1 \rightarrow BR$. The two L forms were assumed to be spectrally indistinguishable. This scheme is adapted from ref. 20, with the additional branch from M_1 to BR introduced here to account for the new finding of biphasic BR recovery kinetics in the D96N mutant (16). Table 1 shows the rate constants used in the simulation.

Products of the difference spectra between the pure intermediates and BR and their kinetics, sampled at logarithmically quasiequidistant time points, provided mixture-difference spectra. These were attenuated by a photocycling ratio (PCR) of 0.15, corresponding to excitation of 15% of the sample to create the noise-free data matrix with the difference spectra arranged as column vectors. The individual mixture spectra were further multiplied with factors deviating from 1 by normally distributed random numbers, with mean 0 and variance 0.01, to account for variations of the laser intensity, as in the case of measured data. Finally, normal distribution random noise was added to the data points to yield a simulated noisy data matrix. This spectral noise has increasing amplitudes toward the blue and red ends of the spectrum as well as decreasing amplitudes in five steps with increasing time. The former models lower light intensity at both ends of the experimental spectra, and the latter models the increasingly longer gate pulses (and, therefore, longer light integration times) with increasing delay times of the optical multichannel analyzer instrument (21). Fig. 1 shows the input absorption spectra, the input intermediate kinetics, and the noisy data matrix.

Analysis of Simulated Data: SVD. SVD of the noise-free data matrix recovers four significant components, with the rest containing fluctuations reflecting rounding errors. SVD of the noisy data matrix provides eigenvectors with nonmonotonously varying autocorrelations beyond the first three. This is the result of the assumed nonuniform noise along both the spectral and time dimensions. The rotation algorithm of Henry and Hofrichter (2) was used to reorder eigenvectors 4–8. A new fourth vector pair was obtained that carries significant signal as revealed by its high autocorrelation. Subsequent eigenvectors contain only random noise. The data matrix was reconstructed with reduced noise by using these first four eigenvectors. SVD treatment of this matrix provided new,

Table 1. Input and recovered rate constants

Reaction	Input k, s^{-1}	Output k, s^{-1}
$K \rightarrow L_1$	5.00×10^5	5.13×10^5
$K \leftarrow L_1$	1.00×10^5	1.00×10^5
$L_1 \rightarrow L_2$	5.00×10^3	5.13×10^3
$L_1 \leftarrow L_2$	1.00×10^3	1.00×10^3
$L_2 \rightarrow M_1$	1.00×10^7	$1.23 \times 10^{6*}$
$L_2 \leftarrow M_1$	1.00×10^7	$1.17 \times 10^{6*}$
$M_1 \rightarrow M_2$	1.00×10^3	1.00×10^3
$M_2 \rightarrow BR$	2.00×10^0	2.09×10^0
$M_1 \rightarrow BR$	2.00×10^2	1.48×10^2

Output rate constants were obtained by fitting the SVD-SM-derived output kinetics for the noisy simulated spectra to the input reaction scheme. Rate constants denoted with an asterisk are minimum values.

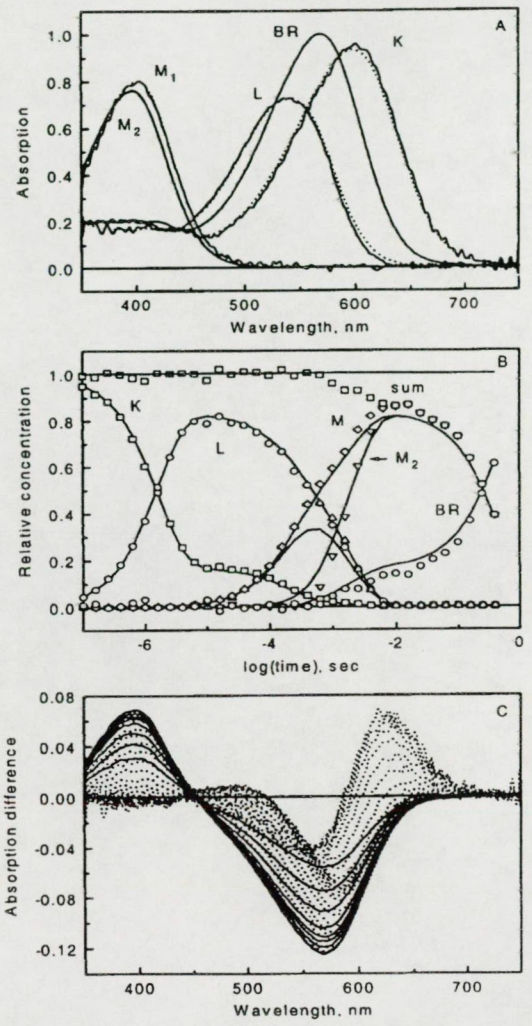


FIG. 1. Input absorption spectra (*A*, dotted lines) and time evolution (*B*, solid lines) of the photocycle intermediates used in the simulation. Combination of the input spectra and kinetics, with noise added, yielded the mixture difference spectra (*C*), with solid lines representing the final decay of the signal. Output intermediate spectra (*A*, solid lines) and kinetics (*B*, symbols) were obtained from the analysis of the data in *C*. In *B*, "sum" means the total intermediate concentration, and "M" means $M_1 + M_2$.

orthonormal eigenvectors (the orthonormality was lost during the rotation procedure):

$$D = U \cdot S \cdot V^T, \qquad [1]$$

where D ($n \times m$) is the reconstructed data matrix whose elements, $D_{ij} = D(\lambda_i, t_j)$, are the absorption difference values at wavelength λ_i and time t_j after the start of the photocycle. Matrices U ($n \times 4$) and V ($m \times 4$) consist of the orthonormal spectral eigenvectors and the orthonormal kinetics vectors, respectively, and the S (4×4) diagonal matrix contains the significant singular values. The product

$$A^T = S \cdot V^T \qquad [2]$$

defines the A ($m \times 4$) matrix, which is equivalent to the combination coefficient matrix in PCA-SM and whose elements were designated previously as α_j, β_j, γ_j, and δ_j (6).

The Stoichiometric Plane. The elements of the data matrix are products of the difference spectra, $\Delta\varepsilon_k$, of the pure intermediates and their time-dependent concentrations, c_k:

$$D(\lambda_i,t_j) = \sum_{k=1}^{r} \Delta\varepsilon_k(\lambda_i) \times c_k(t_j), \text{ or, in matrix form, } D = \Delta\varepsilon \cdot c^T, \quad [3]$$

where r is the number of intermediates, generally greater than or equal to the rank of matrix D. Both the difference spectra and the concentrations on the right-hand side of Eq. 3 are unknown. There exists an unknown transformation T, however, that converts the SVD basis sets to the respective real spectra and kinetics:

$$D = (U \cdot T) \cdot (T^{-1} \cdot A^T) \quad [4]$$

so that

$$\Delta\varepsilon = U \cdot T \text{ and } c = A \cdot R, \quad [5]$$

where R is the transpose of the inverse of matrix T. Up to a certain time in the photocycle, no recovery of the BR initial state takes place, and the sum of the intermediate concentrations is constant, or is unity with proper normalization:

$$\sum_{k=1}^{r} c_k(t_j) = 1, \quad j = 1, \ldots, l \leq m. \quad [6]$$

Eqs. 5 and 6 together yield for the combination coefficients:

$$\sum_{k=1}^{r} R_k A_{j,k} = 1, \quad j = 1, \ldots, l \leq m, \quad [7]$$

where $R_k = \Sigma_{i=1}^{r} R_{k,i}$ are time-independent constants for $k = 1, \ldots, r$.

The set of Eq. 7 is analogous to the equation of a plane in three-dimensional space. We therefore designate the r-1 D surface of points that obey Eq. 7 the stoichiometric plane (SP). Before the onset of the recovery of the initial state in the photocycle, each point in this space that corresponds to a mixture difference spectrum must fall on the SP. Accordingly, the combination coefficients belonging to the unknown difference spectra of the pure intermediates also must fall on this plane. The stoichiometric criterion for the pure intermediate spectra introduced by Nagle *et al.* (22) is a consequence of Eq. 7.

Identification of the SP is based on the combination coefficients contained in matrix A. Because its columns contain increasing amounts of noise, a new transformation is helpful in the accurate determination of the SP, which, by properly mixing the columns of A (and U), provides new combination coefficient vectors with more evenly distributed noise while the new spectral basis vectors obtained from U are still orthonormal:

$$U \cdot A^T = (U \cdot P) \cdot (P^T \cdot A^T). \quad [8]$$

For a three- and four-component system, respectively, the corresponding P matrices are as follows:

$$P = 1/\sqrt{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix} \quad P = 1/2\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \quad [9]$$

After this transformation, the first 4, 5, . . . , m equations in the Eq. 7 are solved consecutively for R_k in the least squares sense, and, in each case, the standard deviations of the corresponding 4, 5, . . . , m points from the derived SP are calculated. The parameter l, i.e., the number of spectral points before any recovery of the initial state, is identified as the one before the deviations start to systematically increase. In other words, l corresponds to the number of spectra that gives the minimum standard deviation of the spectral points from the least squares plane defined by these points. Although this approach reveals the early BR recovery for noise-free data, it yields the incorrect result of no BR recovery until the 27th data point, when the entire spectral region of the noisy input data is analyzed. This result is attributable to the small difference between the M spectra. Therefore, the SP was searched for on truncated data matrices in the >540-nm range, where the two M − BR difference spectra are identical and the matrices behave as robust, three-component systems.

Table 2 shows the standard deviation from unity of the left-hand side of Eq. 7, as fitted to the first 11, 12, . . . , 34 spectral points of the noisy data. An early increase of this deviation is followed by a plateau before the final increase. For the noise-free and the noisy truncated data, the number of data points before the early recovery was estimated as 18 and 20, and the R_k parameters in the equation of the SP (Eq. 7) for l = 18 and l = 20, respectively, were determined. The standard deviation before the early recovery of BR for the noisy data is consistent with the standard deviation from unity introduced in the simulation to model the laser intensity fluctuations.

Estimation of the Photocycling Ratio. The PCR (also considered unknown, as in the case of real experiments) was obtained in a general way without using the pure M_2 − BR difference spectra that are expected in the millisecond time domain in the D96N mutant. In the first method, the PCR was varied until the best SP was found in the least squares sense that fits the first 20 (18 for noise-free data) truncated difference spectra augmented with the PCR-scaled negative of the BR absorption spectrum, in the spectral range 540–750 nm. The latter is equivalent to the pure M − BR difference spectra in this spectral interval and should fall on the SP of the first 20 mixture spectra when scaled by the proper PCR. The second method calculates the dot product between the truncated (−BR) spectrum and the first three spectral eigenvectors (columns of U) from the SVD output of the first 20 truncated difference spectra. The resulting combination coefficients are substituted into Eq. 7 to yield the reciprocal of the PCR. The averages of these values, 14.99% for the noise-free data and 14.43% for the noisy data, were accepted as the true PCR.

The pure M_2 absorption spectrum was obtained from the average of the late SVD reconstructed difference spectra of the full data matrix by using the criterion that adding a properly scaled BR absorption spectrum to the difference spectra contributed by M_2 alone should give uniform baseline for wavelengths >540 nm. Normalization by the scaling factor provides the M_2 absorption spectrum with the proper amplitude.

SM. SM, as tailored here for the BR problem, is the technique of searching for the pure intermediate spectra on

Table 2. The stoichiometric plane and BR recovery

No. of spectra	Standard deviation	No. of spectra	Standard deviation
11	0.011186	23	0.018126
12	0.014940	24	0.019683
13	0.014354	25	0.021780
14	0.013981	26	0.023265
15	0.013530	27	**0.023681**
16	0.013372	28	0.026399
17	0.013183	29	0.030809
18	0.013412	30	0.036071
19	0.013085	31	0.044445
20	**0.012974**	32	0.061808
21	0.014920	33	0.088526
22	0.017599	34	0.124710

Standard deviation of the stoichiometric plane fits to the combination coefficients of the first 11, 12, . . ., 34 noisy difference spectra. Bold numbers represent the last point considered to be on the plane and the point where the main phase of BR recovery starts.

the SP by using geometric criteria as well as criteria regarding the relative displacement on the wavelength axis of the pure intermediate spectra. First, a truncated matrix consisting of difference spectra 1–20, augmented with the negative of the BR spectrum times the PCR, was created for the 540- to 750-nm spectral region. The last column of this matrix represents the pure M_1 − BR difference spectrum, and the remaining 20 represent varying mixtures of the K − BR, L − BR, and M_1 − BR spectra. SVD on this matrix for both the noise-free and the noisy data returned a rank of 3, as expected. The parameters for the SP follow from Eq. 7.

Fig. 2*A* shows the plot of the second combination coefficient versus the first one (the third coefficient is automatically determined by the equation of the SP). Early points show a progression from K toward L, mostly; then, a turn represents the onset of the accumulation of M. The LM side of the triangle corresponding to pure L + M_1 mixtures was located first. The line connecting the first and the last (pure M_1) points was divided to yield nine equidistant starting points on the SP, and a direction was determined that roughly corresponds to the tangent of the trajectory of the early spectral points (a good approximation of the KL line). This direction represents decreasing amounts of K and increasing amounts of L anywhere within the KLM triangle. Moving in this direction along the parallel lines in Fig. 2*A* should lead to points on the LM line. The search for these points was based on the expectation that there is a spectral region in which both M and L absorptions (and any combination thereof) are negligible and only K (and, potentially, BR) absorbs. This region was varied from 610–730 to 690–730 nm, and the standard deviation of the calculated spectra minus the negative of the BR absorption spectrum times the PCR was determined. The smallest standard deviation along each search line defines potential L + M mixtures, and the smallest overall standard deviation was initially monitored to determine the appropriate tail region for L. This parameter leveled off beyond a certain wavelength for both the noise-free and the noisy data. Selection of the final region in which the L spectrum does not contribute was optimized as follows: For each region, the calculation described below was completed, and the resulting trial K, L, M_1, and M_2 spectra were used to fit the data matrix in the least squares sense, with the restriction of nonnegative intermediate concentrations. The best overall fit identifies the region in which only K and BR contribute and, consequently, the final LM line. This line, along with the nine points that determine it, is shown in Fig. 2*A*. Fig. 2*B* shows the corresponding nine difference spectra as well as the M_1 − BR spectrum.

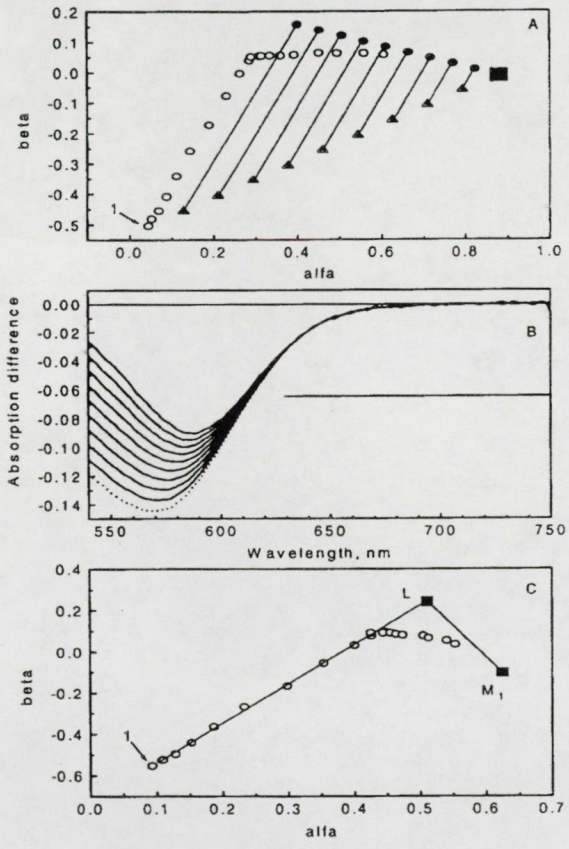


FIG. 2. Demonstration of self-modeling on noisy simulated data. (*A*) α-β projection of the stoichiometric plane fitted in the 540- to 750-nm region of the first 20 difference spectra augmented with the photocycling ratio times the negative of the BR spectrum. Open circles, input spectra, the first one marked as 1; solid triangles, initial nine points in the LM line search along the parallel lines; full circles, the nine points found on the LM line; solid square, pure M_1 − BR. (*B*) Difference spectra corresponding to the nine points on the LM line (solid lines) and the pure M_1 − BR difference spectrum (dotted line) in the 540- to 750-nm range. The horizontal bar represents the estimated interval where the L absorption is zero (630–750 nm). (*C*) α-β projection of the stoichiometric plane fitted to the first 20 difference spectra (full wavelength scale) augmented with the pure M_1 − BR spectrum. Open circles, input spectra, the first one marked as 1; solid squares, pure M_1 − BR and L − BR vertices. The LM and KL lines also are shown.

These truncated spectra were fitted with the spectral eigenvectors obtained from the SVD treatment of the first 20 columns of the original, full data matrix. Only the appropriate region (540–750 nm) of the full length (351–750 nm) SVD spectral eigenvectors was used in the least squares fit. The resulting combination coefficients multiplied by the SVD spectral eigenvectors (full length) yield the nine difference spectra defining the LM line as well as the pure M_1 − BR difference spectrum, now over the full spectral interval.

Linear regression of the nine combination coefficient triplets with the M vertex fixed provided the parameters of the LM line. The pure L − BR spectral point was sought by moving along this line away from the M vertex. The constraint used to define the L − BR point for the noisy data is the expectation that a spectral region (351–410 nm) exists where the L intermediate has approximately the same extinction as BR. The simulated input spectra were constructed in this way, based on low-temperature and room-temperature spectra determined earlier for L (23, 24). Thus, the vertex corresponding to the pure L − BR difference spectrum was identified as the point along the LM line that results in a difference spectrum whose average over the 351- to 410-nm interval is zero.

The criterion used to locate the LM line is not applicable in the search for the KL line because all mixtures of K and L have nonzero absorption throughout the entire spectral range. However, this line must contain the L vertex, and, because no contribution from M is expected at the earliest times, it must include the first (several) spectral points. Fig. 2*C* shows the first two combination coefficients of the first 20 data points (noisy matrix, full wavelength), the location of the pure M_1 − BR and L − BR vertices, the LM line, and the KL line, the latter determined by connecting the L vertex with the average of the first 3 spectral points. In fact, were it not for the noise, the first several spectral points alone could be used to locate the KL line. Then, the L vertex could be found at the intersection of the KL and LM lines, or, more generally, as an extrapolated intersection to time zero even if a little M already contributes to the earliest spectral points. Although this method worked for noise-free data, in the noisy case the intersecting points scatter too much (i.e., no clear progression is obtained with time), so the above method based on the L absorption in the blue region was preferred.

The pure K spectral point is determined by extrapolation of the early spectra to time zero. This was accomplished in two

steps. First, the integral of the initial four difference spectra of the data matrix in the 600- to 750-nm interval minus that of the pure L − BR spectrum yields areas proportional to the concentration of K. Extrapolation to the parameter value of K at time zero is achieved by assuming single exponential decay (see Fig. 3 *Inset*). In the second step, the K vertex is found by moving along the KL line beyond the first spectral point until the corresponding integral parameter reaches the extrapolated value. The first four difference spectra, the L − BR spectrum, and the extrapolated K − BR spectrum are plotted in Fig. 3.

A final SVD treatment of the entire data matrix augmented by the pure K − BR, L − BR, M_1 − BR, and M_2 − BR difference spectra was followed by determining the final "SP" in the four-dimensional space of this rank-4 matrix. This surface was calculated from the combination coefficients of the four pure intermediates by simple matrix inversion rather than by least squares fit. Fig. 4 shows the α, β, γ plot of the combination coefficients of all 34 spectra from the noisy data matrix as well as that of the pure intermediate difference spectra, defining a tetrahedron in three-dimensional space. The origin corresponds to BR, and the adherence of the first 20 points to the "SP" as well as the subsequent deviation from it as BR recovers is demonstrated in this three-dimensional projection of the four-dimensional space.

Comparison of the Input and Output Spectra and Kinetics. The pure absorption spectra of the intermediates are obtained by addition of the BR spectrum, scaled by the PCR, to the pure difference spectra corresponding to the vertices of the tetrahedron in Fig. 4. The recovered spectra for the noisy simulated data set are shown in Fig. 1*A*. The time evolution of the intermediates follows directly from the location of the spectral points within the tetrahedron in Fig. 4. It also can be computed by linear least squares fitting of the mixture spectra by the pure intermediate spectra, with the non-negativity constraint for the concentrations. The output kinetics for the noisy simulated data set are shown in Fig. 1*B*. The output pure spectra and kinetics obtained with the above procedure by using noise-free input spectra are visually indistinguishable from the pure spectra and kinetics used in the simulation (data not shown). The RMS noise content of the input difference spectra varies between 4.0×10^{-3} and 2.9×10^{-4} (higher noise in the earlier spectra). The RMS deviations between the noise-free input spectra and the output noisy spectra of the intermediates (all scaled by PCR) are 2.5×10^{-3}, 2.1×10^{-3}, 1.4×10^{-3}, and 2.0×10^{-4} for K, L, M_1, and M_2, respectively.

The output kinetics from the noisy data were fitted to the same reaction scheme used to generate the input data. The rate constants are listed in Table 1. The error of the fit is low, and all but the $M_1 \rightarrow$ BR rate constant agree very well with the input rates. With the noise-free data, all recovered rate

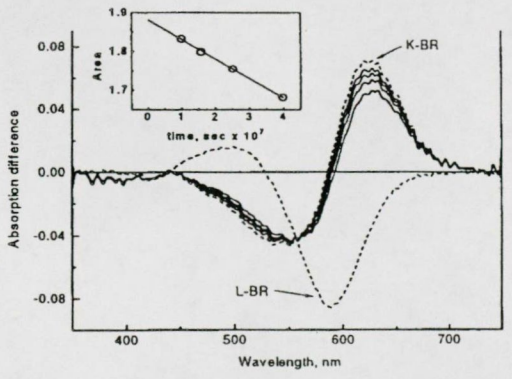


FIG. 3. The first four mixture difference spectra (solid line), the pure L − BR spectrum (dashed), and the extrapolated pure K − BR spectrum (dashed line). (*Inset*) The extrapolation to time zero of the logarithm of the integral parameter used to locate the K spectral point.

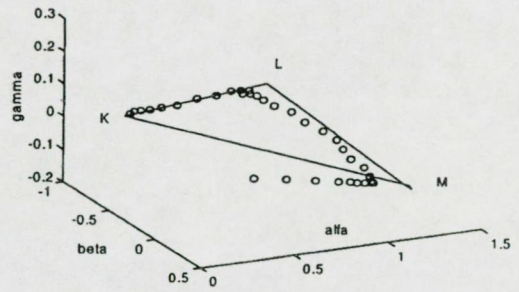


FIG. 4. α, β, γ plot of the combination coefficients of the entire noisy data matrix (open symbols) plus the pure K − BR, L − BR, M_1 − BR, and M_2 − BR spectra as vertices of the stoichiometric tetrahedron. The adherence of the data points to the stoichiometric surface is demonstrated. The M_1 and M_2 points are very close to each other because of their spectral similarity.

constants are essentially identical to those used in the simulation (data not shown). SVD-SM on a noisy data matrix generated similarly to the one discussed, but without the $M_1 \rightarrow$ BR step, resulted in the correct spectra and kinetics lacking the early BR recovery (data not shown).

DISCUSSION

Multichannel and single wavelength kinetic absorption measurements have been published in numerous articles on bacteriorhodopsin (14, 17–21, 24–29). Various strategies have been applied to obtain the ultimate information, the time dependence of the photocycle intermediates, which is essential to the elucidation of the mechanism and energetics of light energy conversion by this protein. Global model fitting on such data has generally returned ambiguous results, with spectra of the intermediates possessing unusual properties, such as more than one absorption band, incorrect baselines, etc. (25, 26).

L.Z. and J.K.L. have pursued the strategy of a model-independent determination of the pure intermediate spectra first, followed by calculation of the intermediate concentrations in the second step. The original trial-and-error method (27, 28) was improved by a grid search algorithm (17) and more recently by a Monte Carlo-based procedure to obtain the spectra (24). Pure intermediate spectra were found approximately by narrowing the limits imposed on various spectral parameters: for instance, the height, the width, and the negative value tolerance. Factor analysis (analogous to PCA) combined with similar spectral criteria applied simultaneously on visible and Fourier transform IR spectra was used by others to dissect the photocycle (29).

SVD was applied to chromoprotein spectra to estimate the number of spectrally independent components, to eliminate random noise, and to store spectral information in a compressed form (17, 30, 31). The kinetics vectors of the SVD output were fitted by sums of exponentials, yielding phenomenological rates and amplitudes. Such information can be used to obtain microscopic rate constants by the fitting of photocycle models. SVD alone tends to underestimate the number of components from noisy data, which is usually more accurately determined by the multiexponential fit, if the pure component spectra are not clearly distinguishable (or are not linearly independent) (32). However, the advantage of SVD is utilized here and in the following paper (16), where SVD is combined with the application of self-modeling: i.e., the search for the pure intermediate spectra in the SVD eigenvector space. Most of the assumed, empirical spectral criteria that were essential in earlier methods are rendered unnecessary by SVD-SM, which takes advantage of the stoichiometric behavior of the photocycle.

D.F.S. and J.S. have applied PCA with self-modeling in photochemistry in resolving absorption and fluorescence spectra of mixtures of unknown composition (4, 6–11). The original SM procedure (33) used no other constraint regarding the location of the pure forms on the combination coefficient normalization planes (lines) than the inner limits defined by the most extreme measured spectral points and the outer limits required by non-negative absorption. This procedure often results in a wide range of acceptable pure component spectra. Unique solutions can be obtained by applying additional constraints dictated by the known chemistry of each system. The introduction of the global Stern-Volmer constant optimization constraints for fluorescence spectra (9) is an illustration.

Adaptation of the procedure to BR requires several modifications. Normalization of spectra is abandoned because the input data are difference rather than absolute spectra. The stoichiometric condition is used instead. The advantage of relying on stoichiometric relationships is that they reveal as part of the analysis the time of the onset of BR recovery and the photocycling ratio. Self-modeling, performed on the stoichiometric plane, readily yields the LM side of the KLM triangle because only K contributes at the red edge of the spectra. The location of M on this line is based on its lack of absorption over most of the visible wavelength range. Because in the blue region of the visible spectrum there is no wavelength range in which any intermediate has zero absorption, two additional criteria are introduced outside the usual framework of SM to locate the remaining pure intermediate spectra: The absorption of intermediate L is considered the same as that of BR in the blue region (this constraint is not needed for the analysis of noise-free simulated spectra), and the spectrum of K is estimated by extrapolation to time zero. A single time-dependent exponential function is used in the extrapolation as the most reasonable choice.

Analysis of noise-free simulated data returns the input spectra almost exactly. Closer examination reveals a slight shift of the output M_1 spectrum toward M_2 and a minor discrepancy between the tails of the input and output L spectra (data not shown). Both are caused by imposing the three-component approximation to the first 18 mixture spectra. Although the stoichiometric condition holds up to the 18th spectrum, there is a small amount of M_2 present after the 15th spectrum. SVD of the noise-free input matrix reveals this, but, with the noisy matrix, the presence of trace M_2 is concealed. When SVD-SM analysis is performed on the noise-free matrix with only the first 15 mixture spectra included, the slight spectral discrepancies disappear. The same procedure does not succeed with the noisy data matrix because the level of noise and the small accumulation of M prevent the location of the proper SP when only 15 spectra are considered.

The small spectral discrepancies cause more visible deviations between input and output kinetics, the latter computed by non-negative least squares fitting the input mixture spectra with the output pure spectra. However, despite the noise level introduced here, the fit of the same reaction scheme to the output kinetics as the input photocycle model gives very good agreement, and with only minor differences in the rate constants. Only the $M_1 \rightarrow$ BR rate deviates by ≈25%, mostly because of the 3% underestimation of the PCR in the case of the noisy data.

The overall agreement of the input simulated spectra and kinetics with those recovered by SVD-SM shows that this approach, while avoiding the imposition of subjective spectral constraints, significantly narrows the range of potential solution spectra relative to earlier approaches based on the grid search and on the Monte-Carlo method. In the accompanying paper, we demonstrate that SVD-SM leads to more accurate intermediate spectra and kinetics in cases of real experimental data as well.

We thank Professor M. Kasha for suggesting this collaboration and Dr. G. Groma for providing the program (RATE 2.1) used in the fitting of kinetics models. This work was supported by grants from the National Scientific Research Fund of Hungary (OTKA T020470) to L.Z., from the National Institutes of Health (GM 29498) and the Department of Energy (DEFG03-86ER13525) to J.K.L., and from the National Science Foundation (CHE 9612316) to J.S.

1. Warner, I. M., Christian, G. D., Davidson, E. R. & Callis, J. B. (1977) *Anal. Chem.* **49**, 564.
2. Henry, E. R. & Hofrichter, J. (1992) *Methods Enzymol.* **210**, 129–192.
3. Aartsma, T. J., Gouterman, M., Jochum, C., Kwiram, A. L., Pepich, B. V. & Williams, L. D. (1982) *J. Am. Chem. Soc.* **104**, 6278–6283.
4. Saltiel, J. & Eaker, D. W. (1984) *J. Am. Chem. Soc.* **106**, 7624–7626.
5. Borgen, O. S. & Kowalski, B. R. (1985) *Anal. Chim. Acta* **174**, 1–16.
6. Saltiel, J., Sears, D. F., Mallory, F. B., Mallory, C. W. & Buser, C. A. (1986) *J. Am. Chem. Soc.* **108**, 1688–1689.
7. Sun, Y-P., Sears, D. F. & Saltiel, J. (1987) *Anal. Chem.* **59**, 2515–2519.
8. Saltiel, J., Choi, J.-O., Sears, D. F., Eaker, D. W., Mallory, F. B. & Mallory, C. W. (1994) *J. Phys. Chem.* **98**, 13162–13170.
9. Saltiel, J., Sears, D. F., Choi, J.-O., Sun, Y.-P. & Eaker, D. W. (1994) *J. Phys. Chem.* **98**, 35–46.
10. Saltiel, J., Choi, J.-O., Sears, D. F., Eaker, D. W., O'Shea, K. E. & Garcia, I. (1996) *J. Am. Chem. Soc.* **118**, 7478–7485.
11. Saltiel, J., Zhang, Y. & Sears, D. F. (1996) *J. Am. Chem. Soc.* **118**, 2811–2817.
12. Lanyi, J. K. (1993) *Biochim. Biophys. Acta* **1183**, 241–261.
13. Ebrey, T. G. (1993) in *Thermodynamics of Membranes, Receptors and Channels*, ed. Jackson, M. (CRC, Boca Raton, FL), pp. 353–387.
14. Lozier, R. H., Xie, A., Hofrichter, J. & Clore, G. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3610–3614.
15. Nagle, J. F. (1991) *Biophys. J.* **59**, 476–487.
16. Zimányi, L., Kulcsár, A., Lanyi, J. K., Sears, D. F., Jr., & Saltiel, J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4414–4419.
17. Zimányi, L. & Lanyi, J. K. (1993) *Biophys. J.* **64**, 240–251.
18. Váró, G. & Lanyi, J. K. (1990) *Biochemistry* **29**, 2241–2250.
19. Zimányi, L., Cao, Y., Chang, M., Ni, B., Needleman, R. & Lanyi, J. K. (1992) *Photochem. Photobiol.* **56**, 1049–1055.
20. Gergely, C., Ganea, C., Groma, G. & Váró, G. (1993) *Biophys. J.* **65**, 2478–2483.
21. Zimányi, L., Keszthelyi, L. & Lanyi, J. K. (1989) *Biochemistry* **28**, 5165–5172.
22. Nagle, J. F., Zimányi, L. & Lanyi, J. K. (1995) *Biophys. J.* **68**, 1490–1499.
23. Becher, B., Tokunaga, F. & Ebrey, T. G. (1978) *Biochemistry* **17**, 2293–2300.
24. Gergely, C., Zimányi, L. & Váró, G. (1997) *J. Phys. Chem. B* **101**, 9390–9395.
25. Nagle, J. F., Parodi, L. A. & Lozier, R. H. (1982) *Biophys. J.* **38**, 161–174.
26. Xie, A. H., Nagle, J. F. & Lozier, R. H. (1987) *Biophys. J.* **51**, 627–635.
27. Váró, G. & Lanyi, J. K. (1991) *Biochemistry* **30**, 5008–5015.
28. Váró, G. & Lanyi, J. K. (1991) *Biochemistry* **30**, 5016–5022.
29. Hessling, B., Souvignier, G. & Gerwert, K. (1993) *Biophys. J.* **65**, 1929–1941.
30. Hug, S. J., Lewis, J. W., Einterz, C. M., Thorgeirsson, T. E. & Kliger, D. S. (1990) *Biochemistry* **29**, 1475–1485.
31. Hoff, W. D., van Stokkum, I. H. M., van Ramesdonk, H. J., van Brederode, M. E., Brouwer, A. M., Fitch, J. C., Meyer, T. E., van Grondelle, R. & Hellingwerf, K. J. (1994) *Biophys. J.* **67**, 1691–1705.
32. Dioumaev, A. K. (1997) *Biophys. Chem.* **67**, 1–25.
33. Lawton, W. H. & Sylvestre, E. A. (1971) *Technometrics* **13**, 617.

Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 4414–4419, April 1999
Biophysics

Intermediate spectra and photocycle kinetics of the Asp96 → Asn mutant bacteriorhodopsin determined by singular value decomposition with self-modeling

LÁSZLÓ ZIMÁNYI*†, ÁGNES KULCSÁR*, JANOS K. LANYI‡, DONALD F. SEARS, JR.§, AND JACK SALTIEL§¶

*Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, Szeged, H-6701, Hungary; ‡Department of Physiology and Biophysics, University of California, Irvine, CA 92697; and §Department of Chemistry, Florida State University, Tallahassee, FL 32306-4390

Communicated by Michael Kasha, Florida State University, Tallahassee, FL, February 16, 1999 (received for review October 27, 1998)

ABSTRACT **Singular value decomposition with self-modeling is applied to resolve the intermediate spectra and kinetics of the Asp96 → Asn mutant bacteriorhodopsin. The search for the difference spectra of the intermediates is performed in eigenvector space on the stoichiometric plane. The analysis of data at pH values ranging from 4 to 8 and temperatures between 5 and 25°C reveals significant, early partial recovery of the initial state after photoexcitation. The derived spectra are not biased by assumed photocycles. The intermediate spectra derived in the initial step differ from spectra determined in prior analyses, which results in intermediate concentrations with improved stoichiometric properties. Increasingly more accurate photocycles follow with increasing assumed complexity, of which parallel models are favored, consistent with recent, independent experimental evidence.**

Application of singular value decomposition with self-modeling (SVD-SM) to the determination of the pure intermediate spectra and kinetics of a simulated bacteriorhodopsin photocycle was demonstrated in the preceding paper (1). SVD-SM is applied to real data in this paper. Bacteriorhodopsin (BR) functions as a light-driven proton pump in the cell membrane of *Halobacterium salinarium*. Photoexcitation results in isomerization of the all-*trans*-retinal chromophore and sequential structural changes of the protein moiety (photocycle), characterized in the visible range by the distinct spectra of the metastable intermediates J, K, L, M, N, and O. At the end of the photocycle, the initial state, BR, recovers. The net result is the transfer of a proton from the cytoplasmic to the extracellular side (for reviews, see refs. 2 and 3). The proton pathway in the extracellular half channel involves the Schiff base (covalently bonding the retinal and Lys-216), Asp-85, and the region of Glu-204 and Glu-194, with water molecules also playing an important role (4–13). In the cytoplasmic half channel, Asp-96 is the proton donor to the transiently deprotonated Schiff base (4, 5, 14, 15). Replacement of Asp-96 with asparagine decelerates this step by orders of magnitude, effectively preventing the accumulation of any intermediate after M, while retaining the proton pumping activity (16–18). The photocycle of the Asp96 → Asn (D96N) mutant has been the target of a number of investigations because of its relative simplicity (19–22).

Understanding the proton pump requires knowledge of the exact time evolution of the intermediates. This information can be obtained from multichannel kinetic absorption spectroscopy once the spectra of the pure intermediates are determined. Determination of the spectra is difficult, however, because of the strong temporal and spectral overlap of the intermediates. Global model fits are designed to yield the spectra and the kinetics simultaneously, but experience, in the case of BR, shows that optimization routines are plagued by local minima that hamper selection among assumed kinetics models (23).

Similar problems are general in spectroscopy and have been addressed by various algebraic methods. Principal component analysis (PCA) and SVD each yield the effective rank of the data matrix and an orthonormal spectral basis set of eigenvectors and corresponding combination coefficients with the number of significant components being equal to the effective rank (24, 25). Self-modeling originally was applied to search for the pure component absorption or fluorescence spectra in the eigenvector space by using the non-negativity criterion (26, 27). The known photophysics of a system provide additional criteria, such as the global Stern-Volmer plot optimization criterion, that allow selection of unique pure component spectra from the range permitted by the non-negativity constraint (28).

Multichannel spectroscopy on BR provides a data matrix that consists of difference rather than absolute spectra, and, because the exact fraction of the photoexcited BR molecules is unknown, these spectra cannot be converted directly to absolute spectra. PCA-SM procedures do not apply because difference spectra cannot be normalized. Instead, the search for the pure component spectra can be based on stoichiometric constraints. The details of the algorithm were elaborated and tested on simulated data resembling data measured on D96N BR (1). Treatment of noise-free simulated data recovered the known input intermediate spectra and kinetics almost exactly. With realistic noise included, the output spectra and kinetics are sufficiently close to the input spectra and kinetics to allow the fit of the same photocycle scheme as the input scheme, with very good rate constant recovery. The method reveals the correct stoichiometric behavior of the intermediates and detects the early partial recovery of the initial state included in the input kinetics.

The analysis of real data measured on D96N BR is described here. SVD confirms that these data reflect a problem with four major intermediates (K, L, M_1, and M_2). Determination of the stoichiometric plane (SP) reveals an early partial recovery of BR. Self-modeling is performed in the eigenvector space of data representing the first half of the photocycle, a rank-three matrix, and yields the spectra of the K, L, and M_1 intermediates. The M_2 spectrum is determined from direct addition of an appropriately scaled BR spectrum to one of the late difference spectra. Comparison of the pure intermediate spec-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: BR, bacteriorhodopsin; PCA, principal component analysis; PCR, photocycling ratio; SM, self-modeling; SP, stoichiometric plane; SVD, singular value decomposition.

†Parts of this work were done during L.Z.'s visits to the University of California, Irvine, and to Florida State University.

¶To whom reprint requests should be addressed. e-mail: jsaltiel@chemmail.chem.fsu.edu.

tra with published spectra obtained from the same data set with different methods (19, 20) reveals deceptively small shifts and amplitude changes, which result in substantial differences in the kinetics. Earlier reported deviations of the sums of the intermediate concentrations from the expected value of unity are diminished. The simplest photocycle schemes that fit these kinetics with acceptable accuracy, with the expected Arrhenius and pH dependencies of the rate constants, are either branched at the level of M substates, in accordance with the pH dependence of proton release, or involve two independent pathways. However, only the latter, parallel scheme seems to be consistent with new, photostationary state experimental evidence for the accumulation of pure M intermediate in this mutant (29).

MATERIALS AND METHODS

Experimental data analyzed in this work are described in previous publications (19, 20, 22). D96N site-directed mutant bacteriorhodopsin was purified as purple membranes as described (30). The membranes were embedded in polyacrylamide gel and were soaked in 100 mM NaCl and 50 mM phosphate buffer plus 2 mM, 200 μM, 20 μM, 2 μM, and no NaN_3 at pH 8, 7, 6, 5, and 4, respectively. Azide was added to catalyze the reprotonation of the Schiff base in the absence of the proton donor group, D96. The samples were light-adapted throughout the measurements. Time resolved difference spectra were measured at 5, 15, and 25°C after excitation with a nitrogen laser-pumped dye laser, using the gated optical multichannel analyzer setup described earlier (31). Thirty-five spectra were collected in each experiment, with logarithmically equidistant time base from 70 ns to 420 ms. The gate pulse width of the detector was increased in five steps with increasing delay while the number of averaged scans per spectrum was decreased. Nevertheless, the noise content of the early spectra is higher than that of the late spectra. The potential actinic effect of the measuring light was ruled out by taking similar data at decreased light intensity. This resulted in no change of the difference spectra apart from a higher noise content. Data analysis was performed by using the method described in the previous article (1). The program, RATE2.1, written by Géza Groma (Biological Research Center, Szeged, Hungary), was used to fit photocycle schemes to the kinetics of the intermediates.

RESULTS

Singular Value Decomposition. The effective ranks of the data matrices consisting of the time-resolved difference spectra measured on D96N BR at pH 8 at 5, 15, and 25°C and at pH 4, 5, 6, and 7 at 25°C were determined by SVD combined with the rotation procedure of Henry and Hofrichter (25). The latter is necessary to "pull" forward later eigenvectors based on their higher temporal autocorrelation and to reject earlier eigenvectors with higher singular values but no correlation. Experience with simulated data shows that the false ordering of some of the SVD components is a result of the uneven noise distribution in the data matrix and is removed by the rotation method (1). The rank of the data matrices is 4 at all temperatures and pH values studied. Accordingly, the original data matrices were reconstructed and thereby noise-filtered by using the first four SVD eigenvectors. For example, the SVD-reconstituted input difference spectra at pH 8, 5°C are shown in Fig. 1*A*. The absorption spectrum of light adapted D96N BR is shown in Fig. 1*B*.

The intermediate spectra were independently determined from the pH 8 data at the three temperatures. Pure M absorption spectra (later designated as M_2) were obtained by adding appropriately scaled BR absorption spectra to difference spectra 33, 30, and 30 at 5, 15, and 25°C, respectively, each

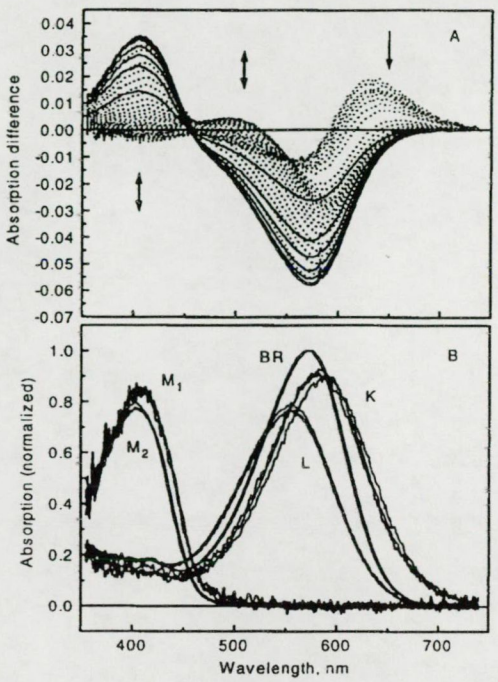


FIG. 1. (*A*) Time-resolved difference spectra measured in the 70-ns to 420-ms time domain on D96N mutant BR at pH 8, 5°C. The arrows indicate the time evolution of the signal, and the solid curves correspond to the final decay phase. The spectra are reconstructed from the first 4 SVD components. (*B*) Intermediate absorption spectra of D96N BR calculated for 5, 15, and 25°C at pH 8.

of which consists of essentially pure M_2, by requiring the best possible baseline in the >540-nm wavelength range.

The Stoichiometric Planes. The number of difference spectra before the recovery of the BR initial state was determined from truncated data matrices ($\lambda > 540$ nm), as in the case of the simulated data (1). The stoichiometric plane equation holds for the combination coefficients (elements of matrix A) up to the onset of the recovery of BR (Eq. 7 in ref. 1), where r is the rank of the data matrix, m is the total number of difference spectra, l is the number of spectra before BR recovery, $A_{j,k}$ is the combination coefficient of the jth difference spectrum relative to the kth SVD spectral eigenvector, and R_k is the parameters of the stoichiometric plane as determined by least squares fit:

$$\sum_{k=1}^{r} R_k A_{j,k} = 1, \quad j = 1, \ldots, l \leq m \qquad [1]$$

Table 1 shows the standard deviation from unity of the left-hand side of Eq. 1 fitted to the first 11, 12, . . . , 35 spectral points. At all three temperatures, there is an early increase of this deviation followed by a plateau before the final increase. The number of spectra before the early recovery is estimated as 23, 19, and 17 at 5, 15, and 25°C, respectively, and the parameters of the stoichiometric planes were calculated accordingly.

Self-Modeling. The photocycling ratio (PCR) was calculated as the average of two slightly different values obtained by two different methods. First, PCR was varied as a parameter until the best stoichiometric plane was found in the least squares sense that fits the first 23, 19, and 17 truncated difference spectra at 5, 15, and 25°C, respectively, augmented with the varied PCR scaled negative of the BR absorption spectrum, in the spectral range of 540–740 nm. Second, the dot product was calculated between the spectral eigenvectors from the SVD treatment of the first 23, 19, and 17 truncated difference

Table 1. The stoichiometric planes and BR recovery

No. of spectra	Standard deviation		
	5°C	15°C	25°C
11	0.0114	0.0231	0.0230
12	0.0151	0.0220	0.0222
13	0.0179	0.0211	0.0223
14	0.0198	0.0205	0.0217
15	0.0202	0.0200	0.0229
16	0.0200	0.0194	0.0233
17	0.0198	0.0190	**0.0237**
18	0.0193	0.0185	0.0248
19	0.0194	**0.0184**	0.0265
20	0.0189	0.0186	0.0280
21	0.0189	0.0197	0.0318
22	0.0187	0.0214	0.0347
23	**0.0186**	0.0248	0.0362
24	0.0200	0.0288	0.0376
25	0.0224	0.0312	0.0380
26	0.0269	0.0343	0.0382
27	0.0312	0.0375	0.0382
28	0.0348	0.0392	0.0382
29	0.0371	**0.0408**	0.0377
30	**0.0389**	0.0427	0.0376
31	0.0427	0.0452	**0.0380**
32	0.0461	0.0523	0.0446
33	0.0538	0.0667	0.0581
34	0.0683	0.0878	0.0849
35	0.1038	0.1347	0.1389

Standard deviation of the stoichiometric plane fits to the combination coefficients of the first 11, 12, . . . , 35 difference spectra. Data are for the D96N mutant bacteriorhodopsin at 5, 15, and 25°C, pH 8. Bold numbers represent the last points considered to be on the plane and the points at which the main phase of BR recovery starts.

spectra, respectively, and the BR absorption spectrum in the 540- to 740-nm interval, yielding its combination coefficients. Because, for any valid difference spectrum that belongs on the SP Eq. **1** holds, so it does for the −BR spectrum (equivalent in the above spectral range to M − BR) times the true PCR. Therefore, on introducing the combination coefficients of −BR into Eq. **1**, we obtain 1/PCR. The PCRs, estimated in this way, are 20.2, 20.2, and 21.4% at 5, 15, and 25°C, respectively. Previous analyses of the same data (19, 20, 22) assumed that the photocycling ratios, determined from the pure M − BR difference spectra that are expected in the millisecond time range, apply also to earlier times of the photocycle. It was noted in those papers that this assumption might not be valid. The present stoichiometric analysis establishes that it is not.

The LM lines were determined as for the simulated data by using the truncated data set (1). Systematic search on the SP along directions representing decreasing amounts of K leads to points on this line by monitoring the red spectral region, where no contribution by either M or L absorption is expected. This region, identified as 697–740 nm, yields pure intermediate spectra that give the most accurate fit to the mixture input spectra. Once the truncated difference spectra on the LM line were found, the spectral region was extended to the full range of 355–740 nm, and the spectrum of M, designated here as M_1, was obtained at each temperature. The extended spectra are calculated as linear combinations of the SVD eigenvectors over the full spectral range, with coefficients obtained by least squares fit of the truncated spectra with the truncated eigenvectors. The L spectrum was found by proceeding away from M along the LM line as for the simulated data until the average absorption of L in the 355- to 410-nm interval matched that of BR.

The KL line at each temperature was identified as the line connecting the vertex of L with the average of the points on the SP corresponding to the first three measured spectra. The vertex of pure K − BR was found along this line by monitoring the integral of the resulting spectrum in the 600- to 740-nm interval. This parameter was estimated by extrapolation to time zero by using the first four measured spectra and the pure L − BR spectrum (1).

The adherence of the early spectral points to the stoichiometric planes and the biphasic deviations from these thereafter, as the BR initial state recovers, are visible when the entire spectral matrices are represented in the combination coefficient space of the first three SVD components only (Fig. 2). The intermediate spectra are plotted in Fig. 1*B* together with the M_2 spectra, determined from the late difference spectra, and the BR spectrum. The M_2 spectrum is slightly blue-shifted relative to the M_1 spectrum and has a somewhat smaller amplitude. This is in accord with the small, but clearly visible, gradual blue shift of the positive band in the 400-nm region of the difference spectra in the millisecond range. There is no systematic temperature dependence of the intermediate spectra.

The K, L, M_1, and M_2 minus BR difference spectra were used in a nonnegative least squares fit to all 35 measured

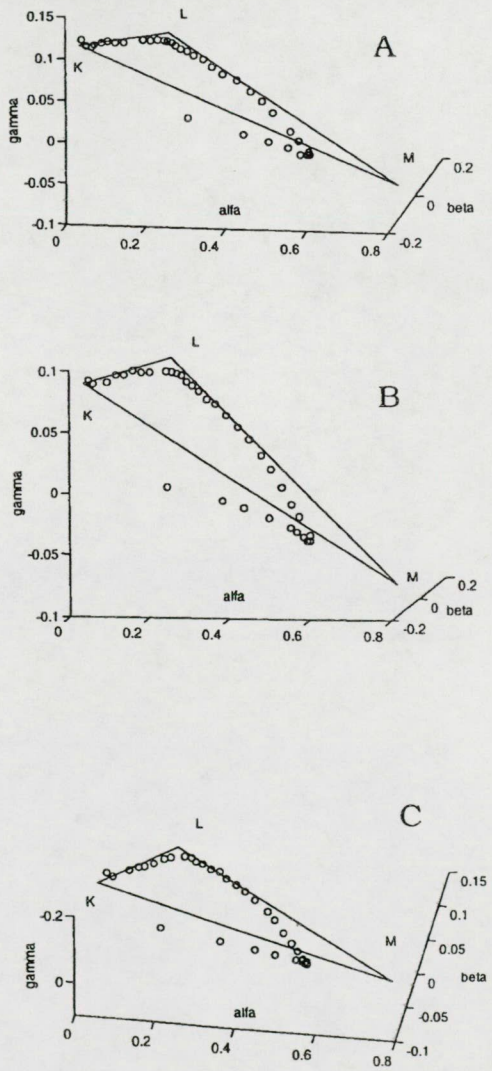


FIG. 2. Three-dimensional plots of the combination coefficients for D96N BR at 5°C (*A*), 15°C (*B*), and 25°C (*C*), pH 8, in the rank-3 approximation of the entire data matrix. Open circles, combination coefficients for the measured difference spectra. The stoichiometric plane is represented as the K, L, M_1 triangle.

difference spectra at the three temperatures. The resulting intermediate concentrations are plotted in Fig. 3 *A*, *B*, and *C* (symbols). The kinetics of BR were obtained as 1 minus the sum of the intermediate concentrations and show the early recovery phase, as already indicated by the fit of the stoichiometric plane. Before this phase, the deviations of the intermediate sums from unity, as reflected in the deviation of the BR concentration from zero, are improved relative to earlier analyses (see Fig. 6 in ref. 20). Performance of SM on the SP ensures optimized stoichiometric behavior by the intermediates. Especially noticeable in their early evolution is that M_1 appears later in this analysis. The spectra obtained in this work are compared with those from ref. 19 in Fig. 4. For this comparison, the spectra computed at the three temperatures were averaged to yield single difference (Fig. 4*A*) and absorption (Fig. 4*B*) spectral sets. A consequence of the earlier underestimation of the PCR is that the new K − BR and L − BR difference spectra are smaller in amplitude. This results in a shift of the respective absolute spectra so that the new K and L absorption bands are closer to the absorption band of the BR initial state.

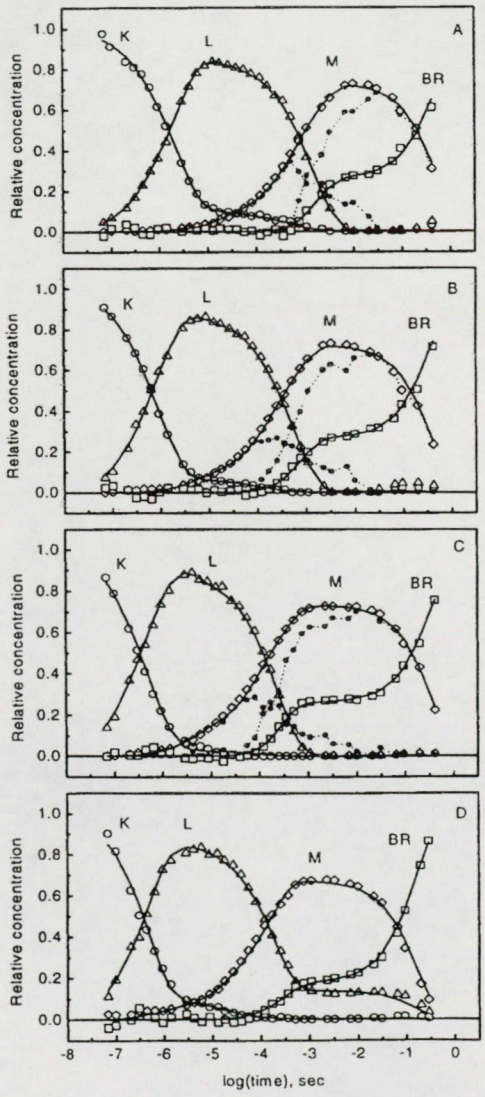


FIG. 3. Intermediate kinetics of D96N BR calculated for pH 8, 5°C (*A*), 15°C (*B*), and 25°C (*C*) and pH 4, 25°C (*D*). Symbols represent intermediate concentrations. In *A*–*C*, the spectrally distinct M_1 and M_2 are represented by black dots and dotted lines. Solid lines, fits by model 1 (see text).

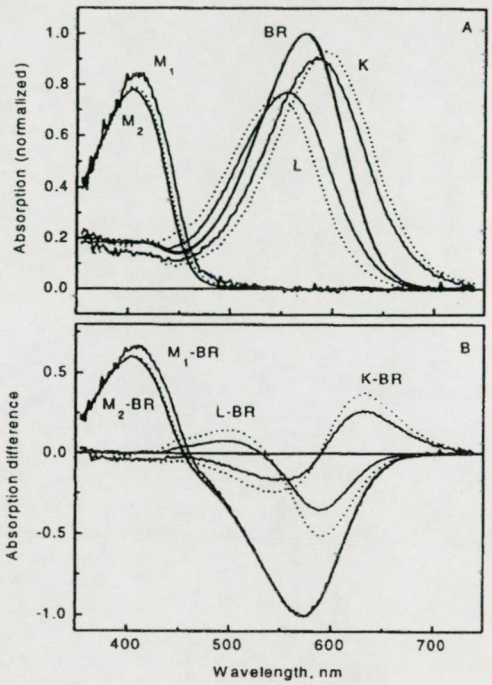


FIG. 4. Intermediate absorption spectra (*A*) and difference spectra (*B*) obtained by SVD-SM (solid lines) compared with those obtained by the grid search method from the same experimental data by Zimányi and Lanyi (19) (dashed lines). In ref. 19, only one M spectrum was resolved, which corresponds to M_2 in the present work.

The averaged intermediate difference spectra were used in a non-negative least squares fit to difference spectra measured at 25°C, pH 4, 5, 6, and 7. The resulting intermediate concentrations at pH 4 are shown in Fig. 3*D* (symbols). They confirm the earlier-reported (22) extended lifetime of L, which gradually diminishes with increasing pH (data not shown). New is the biphasic BR recovery which is as conspicuous in the pH 4–7 range as at pH 8. The sum of the intermediate concentrations before the BR recovery (reflected in the deviation of [BR] from zero) is improved relative to earlier analysis of the same data.

Fitting of Photocycle Models to the Intermediate Kinetics. It is tempting to seek reaction schemes that fit the kinetics in Fig. 3. A thorough analysis was not pursued because the low time resolution (35 points in time) of these data hampers identification of a unique, best model. However, the spectra of the pure intermediates determined here should be useful in fitting high time resolution single wavelength experimental data.

With the four intermediates, a minimal sequential model should consist of the $K \rightleftharpoons L \rightleftharpoons M_1 \rightleftharpoons M_2 \rightarrow BR$ reactions. Here, $M_1 \rightarrow M_2$ corresponds to the reprotonation switch coupled to extracellular proton release so that its reverse is negligible at pH 8 (32), and $M_2 \rightarrow BR$ coincides with the azide catalyzed cytoplasmic proton uptake (16, 18, 33). This model is obviously not capable of reproducing the biphasic BR recovery. Its simplest modification is a sequential BR-like form (BR_1, spectrally indistinguishable from BR) between M_2 and BR. With reversible reactions, except for the final $BR_1 \rightarrow BR$, this model can simultaneously reproduce the biphasic BR recovery and the non-negligible residual L as the pH is decreased. A second sequential L form improves the fit in the region of [M] rise, both in wild-type and in mutant BR (19, 34), and was interpreted as relaxation of the stress in the protein produced by the retinal isomerization (34). It results in a

substantially better fit in the submillisecond range in the present study as well.

Based on mechanistic considerations, the $M_1 \rightleftharpoons M_2$ transition should split into two steps: the reprotonation switch and the extracellular proton release, whose pK was estimated as 5.8–6.0 from the $[L]/[M_{total}]$ concentration ratio. There also should be a pathway to conclude the photocycle without the early proton release at pH lower than the above pK (22). These two internal steps have not yet been resolved, and their relationship is unknown. Hence, we attempted to fit the data with the following sequential, branched model to account for the low and high pH pathways:

$$K \rightleftarrows L_1 \rightleftarrows L_2 \rightleftarrows M_1 \rightleftarrows M_2 \rightleftarrows M_2' \rightleftarrows BR_1' \rightarrow BR \text{ (high pH)}$$
$$\Updownarrow$$
$$BR_1 \rightarrow BR \qquad \text{(low pH)}$$

The $M_1 \rightarrow M_2$ and the $M_2 \rightarrow M_2'$ reactions would correspond to the reprotonation switch and the (early) extracellular proton release by the terminal release group, respectively. The assumed pK = 6 of proton release was taken into account by fixing the $M_2 \rightarrow M_2'$ and $M_2 \leftarrow M_2'$ rate constants. Although this model gives satisfactory fits with reasonable temperature and pH dependencies of the rates, the inclusion of the relatively fast (as compared with the recovery of BR) equilibrium between M_2 and the postulated BR_1 is unavoidable in both branches. This raises serious problems with the sequential schemes in view of recent experimental evidence. In the D96N/D115N double mutant (29), as well as in D96N (J.K.L., unpublished results), photostationary states are established with continuous illumination at conditions similar to our pH 8 experiments (albeit in the absence of azide). The sample can be fully converted into M whereas the fitted rate constants for the above models predict a substantial contribution by BR_1.

Compatible with the photostationary accumulation of pure M is a parallel model with a rapid and a slow pathway. Good fits and meaningful temperature and pH dependencies are obtained with model 1:

$$K^r \rightleftarrows L_1^r \rightleftarrows M_1^r \rightleftarrows M_2^r \rightarrow BR \qquad \text{(rapid pathway)}$$

$$K^s \rightleftarrows L_1^s \rightleftarrows M_1^s \rightleftarrows M_2^s \rightarrow BR \qquad \text{(slow pathway)}$$

In Fig. 3, solid lines show the resulting fitted kinetics at pH 8 and 4. Table 2 lists all rate constants, the ratio between the two pathways, and the χ^2 parameter of the fit by model 1. The reverse of $M^r_1 \rightarrow M^r_2$ is found to be negligible at all pH and temperature values, suggesting a unidirectional reaction. The contribution of the rapid pathway is 18–27%, gradually increasing with increasing pH. The photostationary accumulation of pure M can be explained by postulating a dynamic equilibrium between the rapid and slow photocycles, which would eventually deplete the rapid pathway under continuous illumination. The proton-pumping activity of the D96N mutant is probably associated with the slow pathway. This ensures an efficiency of pumping commensurable with that of the wild type because the contribution of the slow cycle under these conditions is ≈80%. The expected pH dependence of the $M_1 \leftarrow M_2$ back reaction also is attributed to this pathway, reflecting the reverse of the extracellular proton release coupled to the reprotonation switch. Some of the other rate constants also exhibit systematic pH dependencies, possibly associated with the titration of groups close to the membrane surface.

DISCUSSION

In the preceding article (1), we introduced the method of SVD-SM to obtain pure intermediate spectra and kinetics from time-resolved multichannel absorption spectroscopic data on BR. The analysis of simulated data without noise returned the exact input data and a very good approximation of the input data with random noise included. We established that SVD-SM is capable of narrowing the range of potential solution spectra obtained earlier with a grid-search method, providing more confidence in its application to real experimental data. The method is applied here on the D96N mutant BR.

SVD of the difference spectra, measured on the D96N mutant BR, reveals that more than the three major intermediates, K, L, and M, are present. Application of the rotation procedure based on the temporal autocorrelation of the SVD kinetic vectors identified a fourth, clearly significant component. A small blue shift of the band corresponding to the M intermediate is evident in the input difference spectra, suggesting spectrally different early and late Ms, M_1 and M_2, consistent with published results (6). The blue shift of M_2 relative to the wild-type M was attributed to the altered hydrogen bonding of the deprotonated Schiff base in the cytoplasmic direction (21). Least squares fit of the intermediate spectra, found by self-modeling, to the input difference spectra indicates small discrepancies at the very end of the photocycle: The concentrations of K and L, which had decayed to zero, appear to increase again (Fig. 3). This suggests

Table 2. Logarithm of rate constants in s^{-1} from the fit of model 1 to the kinetics of intermediates determined by SVD-SM

	pH 8, 5°C	pH 8, 15°C	pH 8	pH 7	pH 6	pH 5	pH 4
$K^r \rightarrow L^r$	6.2	6.5	6.8	6.4	6.4	6.6	6.8
$K^r \leftarrow L^r$	6.0	6.0	5.9	6.1	6.2	6.7	6.9
$L^r \rightarrow M^r_1$	4.4	4.9	5.1	4.8	4.7	5.2	6.0
$L^r \leftarrow M^r_1$	4.5	5.1	5.1	4.6	4.5	4.6	6.0
$M^r_1 \rightarrow M^r_2$	3.9	4.4	4.4	4.3	4.2	4.6	5.2
$M^r_2 \rightarrow BR$	2.8	3.1	3.7	3.6	3.8	3.6	3.5
$K^s \rightarrow L^s$	5.6	5.9	6.2	6.2	6.2	6.3	6.3
$K^s \leftarrow L^s$	4.2	4.1	4.1	4.0	4.0	4.0	4.7
$L^s \rightarrow M^s_1$	2.8	3.2	3.6	3.6	3.6	3.6	3.6
$L^s \leftarrow M^s_1$	3.9	4.5	4.7	4.5	4.6	4.3	4.1
$M^s_1 \rightarrow M^s_2$	5.3	5.8	6.2	6.3	6.6	6.5	6.6
$M^s_1 \leftarrow M^s_2$	2.0	2.3	2.5	3.0	4.4	5.0	5.5
$M^s_2 \rightarrow BR$	0.3	0.4	0.4	0.4	0.4	0.6	0.9
r:s	26:74	27:73	26:74	25:75	26:74	21:79	18:82
$\chi^2(\times 10^{-2})$	2.3	3.1	2.0	2.3	3.0	2.0	2.0

Output rate constants were obtained by fitting the parallel photocycle scheme, model 1, to the intermediate kinetics derived from the fit of measured difference spectra by the pure intermediate spectra from the SVD-SM analysis. Unless otherwise shown, the temperature was 25°C.

formation of trace amounts of N and/or O. The small total accumulation of these late intermediates, ≤5% of PCR, does not allow determination of their spectra.

SVD-SM reveals a previously unnoticed early recovery phase of the initial BR state. In a former treatment of the same data, we assumed that the photocycling ratio, obtained by adding a properly normalized BR absorption spectrum to one of the late difference spectra to get a pure M spectrum (M_2 in this paper), is universal. This is equivalent to the assumption that no BR recovery has taken place at the time of the difference spectrum used. One can obtain seemingly acceptable intermediate spectra in this way, but, as noted earlier, a discrepancy remains (19, 20): The sum of the intermediate concentrations deviates systematically from unity. Attempts to compensate for this discrepancy involved normalization of all concentrations (20), or the L concentration only (19), by this sum. This normalization is rendered unnecessary in the present work, in which stoichiometric constraints guide the spectral search. Optimized concentration sums naturally follow.

Model fitting to the intermediate concentrations is not expected to be decisive because of the relatively low number of difference spectra (i.e., time points) available. It is important, however, that previous models must be modified substantially to accommodate the new finding of biphasic BR recovery. The simple photocycle model used to construct simulated data (1) would describe the intermediate kinetics in D96N BR at pH 8, where because of the virtual unidirectionality of the $M_1 \rightarrow M_2$ reaction, the K and L forms decay before the full accumulation of M. At lower pH, however, the simulated photocycle model is inadequate to fit the intermediate kinetics. The simplest extension is a sequential form between M and BR, which is spectrally indistinguishable from BR (designated BR_1). Sequential models, which include such an intermediate, fit the data in the present work well; however, they contradict recent observations that demonstrate that 100% M intermediate can accumulate under continuous illumination in the D96N mutant (29). No quadratic term in the light-saturation curve in these experiments was found; thus, it is improbable that reexcitation of BR_1 would convert it to M (L. S. Brown and J.K.L., unpublished work). Therefore, a BR-like form, in equilibrium with M, in amounts required by the above sequential models, is unlikely. Parallel models fit the present data equally well with the introduction of a rapid and a slow pathway. Assuming a dynamic equilibrium between the two pathways, the photostationary accumulation of pure M is explained by the gradual depletion of the rapid pathway. This alternative does not contradict the well established finding that, other than having a slower photocycle, D96N transports protons normally (16–18) if proton transport is linked to the slow but dominant pathway. On the basis of this fit, the spectrally distinct M forms are associated with the two pathways in the parallel model. The suggested heterogeneity and the ensuing parallel pathways are only slightly pH-dependent, unlike the reported heterogeneity in wild-type BR, arising at alkaline pH (35).

The present work validates the SVD-SM approach for the treatment of multichannel kinetic spectroscopic data. The derived, highly accurate intermediate spectra and kinetics for the D96N mutant BR are essential for the evaluation and refinement of proposed photocycle models. The previous assignment of the extracellular proton release to the $M_1 \rightarrow M_2$ step is now confined to the dominant slow pathway of the parallel photocycle model. Work in progress on the SVD-SM analysis of data for the wild-type BR indicates that the early partial recovery of BR, and therefore a parallel pathway of the type suggested in this paper, is not a feature of the wild-type photocycle.

We thank Dr. G. Groma for providing the program (RATE 2.1) used in the fitting of kinetic models. This work was supported by grants from the National Scientific Research Fund of Hungary (OTKA T020470) to L.Z., from the National Institutes of Health (GM 29498) and the Department of Energy (DEFG03-86ER13525) to J.K.L., and from the National Science Foundation (CHE 9612316) to J.S.

1. Zimányi, L., Kulcsár, Á., Lanyi, J. K., Sears, D. F., Jr., & Saltiel, J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4408–4413.
2. Lanyi, J. K. (1993) *Biochim. Biophys. Acta* **1183**, 241–261.
3. Ebrey, T. G. (1993) in *Thermodynamics of Membranes, Receptors and Channels*, ed. Jackson, M. (CRC, Boca Raton, FL), pp. 353–387.
4. Braiman, M. S., Mogi, T., Marti, T., Stern, L. J., Khorana, H. G. & Rothschild, K. J. (1988) *Biochemistry* **27**, 8516–8520.
5. Butt, H.-J., Fendler, K., Bamberg, E., Tittor, J. & Oesterhelt, D. (1989) *EMBO J.* **8**, 1657–1663.
6. Thorgeirsson, T. E., Milder, S. J., Miercke, L. J. W., Betlach, M. C., Shand, R. F., Stroud, R. M. & Kliger, D. S. (1991) *Biochemistry* **30**, 9133–9142.
7. Brown, L. S., Sasaki, J., Kandori, H., Maeda, A., Needleman, R. & Lanyi, J. K. (1995) *J. Biol. Chem.* **270**, 27122–27126.
8. Balashov, S. P., Imasheva, E. S., Govindjee, R. & Ebrey, T. G. (1996) *Biophys. J.* **70**, 473–481.
9. Richter, H. T., Brown, L. S., Needleman, R. & Lanyi, J. K. (1996) *Biochemistry* **35**, 4054–4062.
10. Balashov, S. P., Imasheva, E. S., Ebrey, T. G., Chen, N., Menick, D. R. & Crouch, R. K. (1997) *Biochemistry* **36**, 8671–8676.
11. Dioumaev, A. K., Richter, H. T., Brown, L. S., Tanio, M., Tuzi, S., Saito, H., Kimura, Y., Needleman, R. & Lanyi, J. K. (1998) *Biochemistry* **37**, 2496–2506.
12. Rammelsberg, R., Huhn, G., Lübben, M. & Gerwert, K. (1998) *Biochemistry* **37**, 5001–5009.
13. Luecke, H., Richter, H. T. & Lanyi, J. K. (1998) *Science* **280**, 1934–1937.
14. Gerwert, K., Hess, B., Soppa, J. & Oesterhelt, D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4943–4947.
15. Otto, H., Marti, T., Holz, M., Mogi, T., Lindau, M., Khorana, H. G. & Heyn, M. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9228–9232.
16. Tittor, J., Soell, C., Oesterhelt, D., Butt, H.-J. & Bamberg, E. (1989) *EMBO J.* **8**, 3477–3482.
17. Holz, M., Drachev, L. A., Mogi, T., Otto, H., Kaulen, A. D., Heyn, M. P., Skulachev, V. P. & Khorana, H. G. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2167–2171.
18. Otto, H., Marti, T., Holz, M., Mogi, T., Lindau, M., Khorana, H. G. & Heyn, M. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9228–9232.
19. Zimányi, L. & Lanyi, J. K. (1993) *Biophys. J.* **64**, 240–251.
20. Nagle, J. F., Zimányi, L. & Lanyi, J. K. (1995) *Biophys. J.* **68**, 1490–1499.
21. Zimányi, L., Cao, Y., Chang, M., Ni, B., Needleman, R. & Lanyi, J. K. (1992) *Photochem. Photobiol.* **56**, 1049–1055.
22. Zimányi, L., Váró, G., Chang, M., Ni, B., Needleman, R. & Lanyi, J. K. (1992) *Biochemistry* **31**, 8535–8543.
23. Nagle, J. F. (1991) *Biophys. J.* **59**, 476–487.
24. Warner, I. M., Christian, G. D., Davidson, E. R. & Callis, J. B. (1977) *Anal. Chem.* **49**, 564.
25. Henry, E. R. & Hofrichter, J. (1992) *Methods Enzymol.* **210**, 129–192.
26. Aartsma, T. J., Gouterman, M., Jochum, C., Kwiram, A. L., Pepich, B. V. & Williams, L. D. (1982) *J. Am. Chem. Soc.* **104**, 6278–6283.
27. Saltiel, J. & Eaker, D. W. (1984) *J. Am. Chem. Soc.* **106**, 7624–7626.
28. Saltiel, J., Sears, D. F., Choi, J.-O., Sun, Y.-P. & Eaker, D. W. (1994) *J. Phys. Chem.* **98**, 35–46.
29. Brown, L. S., Dioumaev, A. K., Needleman, R. & Lanyi, J. K. (1998) *Biophys. J.* **75**, 1455–1465.
30. Needleman, R., Chang, M., Ni, B., Váró, G., Fornes, J., White, S. H. & Lanyi, J. K. (1991) *J. Biol. Chem.* **266**, 11478–11484.
31. Zimányi, L., Keszthelyi, L. & Lanyi, J. K. (1989) *Biochemistry* **28**, 5165–5172.
32. Váró, G. & Lanyi, J. K. (1991) *Biochemistry* **30**, 5016–5022.
33. Cao, Y., Váró, G., Chang, M., Ni, B., Needleman, R. & Lanyi, J. K. (1991) *Biochemistry* **30**, 10972–10979.
34. Gergely, C., Ganea, C., Groma, G. & Váró, G. (1993) *Biophys. J.* **65**, 2478–2483.
35. Song, L., Logunov, S. L., Yang, D. & El-Sayed, M. A. (1994) *Biophys. J.* **67**, 2008–2012.

Dissecting the Photocycle of the Bacteriorhodopsin E204Q Mutant from Kinetic Multichannel Difference Spectra. Extension of the Method of Singular Value Decomposition with Self-Modeling to Five Components

Ágnes Kulcsár[†], Jack Saltiel[‡] and László Zimányi*[†]

Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, P.O.Box 521, Szeged, Hungary, H-6701, and Department of Chemistry, Florida State University, Tallahassee, FL 32306-4390 USA

Abstract:

Kinetic multichannel difference spectroscopy in the visible spectral range of the Glu204 → Gln (E204Q) site-directed mutant of bacteriorhodopsin revealed five spectrally distinct metastable intermediates, as for the wild type. Due to the perturbation of the extracellular proton release cluster, the late O intermediate accumulates in much higher amounts in this mutant, and the photocycle is not complicated by the *p*H-dependent branching observed in the wild type protein. This mutant is therefore more amenable than the wild type to the determination of the intermediate spectra with the method of singular value decomposition with self-modeling, developed recently for three components (Zimányi et al., *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4408-4413, 4414-4419). The method provides the most reliable spectra so far, defining the time evolution of the intermediates essential to the determination of the reaction scheme that describes the photocycle. The analysis confirms published results on this mutant by and large, but revises the locations of the L intermediates in the photocycle. In addition, it allows identification of the *p*H-dependent transitions of the photocycle, and indicates a different mechanism for the *p*H dependence of the yield and kinetics of the late O intermediate.

---

*Address correspondence to this author, email: zimanyi@nucleus.szbk.u-szeged.hu

[†]Hungarian Academy of Sciences

[‡]Florida State University

Introduction

The seven-helical transmembrane protein bacteriorhodopsin (BR) in the cell membrane of *Halobacterium salinarum* is a light-driven proton pump that functions as an alternative to respiration. The chromophore is an all-trans retinal bound covalently via a protonated Schiff-base to the ε-amino group of K216, whose isomerization to 13-cis after light absorption initiates a sequence of structural changes. As a result, a proton is released in the extracellular medium and another one is captured on the cytoplasmic side, contributing to the build-up of the transmembrane protonmotive force (for recent reviews see refs 1-3). Photoexcited BR returns to the initial state through consecutive metastable intermediates, which were identified first by their distinct visible absorption spectra, and later by other spectroscopic methods, such as resonance Raman and Fourier transform infrared spectroscopy.[4-6] A schematic description of the photocycle with the intermediates' estimated absorption maxima[7] (rather than the maxima of their difference spectra) and characteristic transition times is shown by the sequence:

$$\mathrm{BR_{570}} \xrightarrow{h\nu} \mathrm{J_{625}} \xrightarrow{5\ ps} \mathrm{K_{590}} \xrightarrow{2\ \mu s} \mathrm{L_{540}} \xrightarrow{50\mu s} \mathrm{M_{410}} \xrightarrow{1\ ms} \mathrm{N_{560}} \xrightarrow{3\ ms} \mathrm{O_{630}} \xrightarrow{5\ ms} \mathrm{BR_{570}},$$

where reversible reactions and spectrally similar substates have been omitted, and J corresponds to a vibrationally hot product whose decay yields K, the first metastable intermediate. The considerably blue shifted absorption of the M intermediate reflects transfer of a proton from the protonated Schiff base to D85,[8] both buried in the interior of the protein.

The 3D structure of BR is now available at 1.55 Å resolution,[9] and the structures of some of the photocycle intermediates have also been published.[10-14] As a result, bacteriorhodopsin is the best characterized membrane protein, and the molecular details of the proton pump are being unveiled with unparalleled accuracy.

The absorption spectra of the intermediates are broad, structureless, and overlap significantly with each other. The time evolution of the metastable intermediates covers at least 5 orders of magnitude (in wild type, at room temperature), in an overlaping fashion, so that the macroscopic sample consists of a mixture of several intermediates at all times after the initiation of the photocycle by a short laser flash. As a result, kinetic multichannel absorption spectroscopy yields a series of spectra with unknown composition of the pure intermediate spectra, which are *a*

priori also unknown. The mathematical problem at hand is to decompose the measured spectra to obtain the pure intermediate spectra, the intermediate kinetics, and finally, to determine the most probable and physically meaningful reaction scheme which accounts for the experimental data with acceptable accuracy. As the structural details about intermediates emerge, the accurate description of the photocycle should serve as the dynamic information for the thorough understanding of the proton pumping mechanism.

The simplest functional photocycle is observed in the Asp96 → Asn (D96N) mutant BR, where due to the removal of the proton source to the Schiff base, the reprotonation of the latter becomes rate limiting, and no intermediate after M accumulates in measurable amounts.[15-17] The method of singular value decomposition with self-modeling (SVD-SM) was developed and originally applied to the dissection of the photocycle of this mutant with 3 distinguishable components (K, L and M).[18,19] The advantage of SVD-SM was demonstrated in the highly accurate intermediate spectra that were determined in the initial step of the analysis, revealing hitherto hidden features in the time evolution of the intermediates.[19]

The extracellular proton release in wild type BR (WT) takes place from the cluster of E204, E194, R82 and one or more water molecules.[20] Early (10-100 μs) proton release is inhibited at *p*H below the (transient) *p*K = 5.8 of the proton release group, and takes place after the cytoplasmic proton uptake in the ms range (when the *p*K is further lowered.[21,22] Therefore, the wild type photocycle is complicated by a *p*H-dependent branching unless the *p*H is significantly higher than this *p*K.[22] Appropriate mutation of any of the contributing amino acids, such as the Glu204 → Gln (E204Q) replacement, inhibits proton release from the cluster altogether, and is expected to eliminate this complication. The photocycle slows down considerably, since in this mutant it is rate-limited by the much slower direct proton release from the buried D85, the primary proton acceptor of the Schiff-base. As a side-effect, the usually marginal last intermediate, O, accumulates to amounts comparable to the yield of the other intermediates.[23,24] This facilitates the accurate determination of the absorption spectra of all intermediates.

Extension of the SVD-SM method to five components leads to the accurate absorption spectra of the intermediates for the E204Q mutant BR. No *p*H or temperature dependence of the spectra was found in the *p*H range 5-9, and temperature range 15-29 °C. An unbranched

photocycle scheme fits the intermediate kinetics with two L and two M substates at all *p*H and temperature values studied. The model with the formerly suggested cul-de-sac position of L_2 is shown to be superior to the one with sequential L substates. Detailed analysis of the N ⇋ O → BR sequence of the photocycle reveals the energetic relations of these forms. An explanation for the anomalous temperature and *p*H dependence of the kinetics and yield of the O intermediate is provided.

Materials and Methods

Purple membranes containing the E204Q mutant bacteriorhodopsin were received as a generous gift from professors J. K. Lanyi and R. Needleman. The membranes were embedded in polyacrylamide gel in order to prevent aggregation and rotational diffusion and to reduce light scattering. Gel slabs of 4 mm x 9 mm were soaked several times in the desired solvent and placed in a sample cuvette with 9 mm pathlength for the measuring beam. The absorbance of the samples at 570 nm was ~ 0.7. Reference gels without purple membranes were used to determine the absorption spectrum of the non-excited sample. All measurements were carried out in 100 mM NaCl, 50 mM Bis-Tris propane buffer at *p*H 5, 7 or 9. Temperatures were maintained at 15 ± 0.2, 22 ± 0.2 and 29 ± 0.2 °C.

Light-induced difference spectra were measured on an optical multichannel analyzer consisting of a spectrograph (Jobin Yvon HR320) and a gated intensified diode array detector system (Spectroscopy Instruments GmbH, detector: IRY-512, controller: ST-120, gate pulse generator: PG-10). The actinic light pulse was obtained from an excimer laser pumped dye laser (JATE XEL604, rhodamine 6G) propagating perpendicularly to the white continuum measuring beam from a Tungsten halogen lamp (not polarized). The measuring light was chopped with a digital shutter (Uniblitz D122, Vincent Associates), which was open for 20 ms with appropriate delay after the laser pulse. Measurements with reduced light intensity demonstrated no actinic effect of the normal intensity measuring light. Control difference spectra taken at the beginning, during, and at the end of the entire spectral set showed no degradation of the sample.

Light intensity spectra with and without the laser pulse were simultaneously averaged in two separate files, and used to calculate time-resolved difference spectra due to the laser excitation of the sample. Delay times of the diode array detector's gate pulse were selected in the

500 ns – 1 s time domain in a logarithmically equidistant fashion. Each decade in time was covered by 7 delay times. The pulse width was chosen for optimal signal amplitude, without significantly lengthening the nominal delay time. This meant increasing the pulse width at longer delay times, from the initial 100 ns, in several steps.

The original data matrix consisted of the raw difference spectra as column vectors. The noise amplitude of the difference spectra varies roughly as the square root of the light intensity incident on the detector, which, in turn, has a spectral distribution. The total accumulated light per scan also varies linearly with the gate pulse length. As a result, the noise distribution of the raw data matrix is uneven along both the spectral and temporal dimensions. A new data matrix was constructed with more homogeneous noise distribution by normalizing with both the square root of the incident light intensity spectrum and the gate pulse width. SVD[25] was performed first on this corrected matrix, which guarantees the accurate ordering of the eigenvectors not only with diminishing significance but also with decreasing autocorrelation. After the reconstruction of the data matrix from the significant eigenvectors, the spectral and temporal corrections were reversed, resulting in the reconstructed data matrix with reduced noise.

New eigenvectors and the corresponding combination coefficients were calculated by a second SVD treatment, since the orthonormality of the original eigenvectors disappears with the reversed corrections. The abstract spectral eigenvectors and the combination coefficients were used in the SM procedure described earlier,[18,19] modified in this paper for >3 components. For details see Results.

Photocycle intermediate kinetics were fitted by the program Rate 2.1, written by Dr. Géza Groma. The N $\leftrightharpoons$ O $\rightarrow$ BR segment of the photocycle was fitted to the analytical solution of the corresponding differential equations using Matlab (The MathWorks, Inc.). The entire SVD-SM procedure was also programmed in Matlab.

Results

Initial treatment of the data. The data analysis is demonstrated on the experiment at *p*H 5, 15°C. Difference spectra on E204Q bacteriorhodopsin after laser pulse excitation, reconstructed from the first 5 SVD components, and normalized by the absorption of the sample at 570 nm in the initial state, are shown in Figure 1. The first 5 eigenvectors account for 99.99 %

of the total variance of the signal. The autocorrelations of the first 5 spectral eigenvectors were > 0.91, and those of the first 5 temporal eigenvectors were > 0.77. The autocorrelations of the 6^{th} spectral and temporal eigenvectors were 0.309 and -0.188, respectively. Hence we conclude from the singular values and the autocorrelations that the first 5 eigenvectors represent the data adequately. This is in accordance with the expectation of 5 spectrally distinct intermediates, and was similar for data matrices at other pH and temperature values. Figure 1A shows the first 20 difference spectra, which constitute a pure 3 component system, as judged by a separate SVD treatment on these spectra only. As expected already from visual inspection, only the K, L and M intermediates contribute to these 20 difference spectra. For comparison, the initial BR absorption spectrum is also plotted (dashed line) after adjustment by the photocycling ratio (p) determined in the following section. Addition of this BR spectrum to the difference spectra would directly yield the absolute spectra of the intermediate mixtures. Figure 1B shows the remaining spectra 21-40, which, in principle, may include contributions from all 5 intermediates.

SVD of the entire data matrix, **D**, provides the abstract spectral eigenvectors, **U**, the diagonal matrix of the singular values, **S**, and the abstract temporal eigenvectors, **V**:

$$\mathbf{D} = \mathbf{U}\,\mathbf{S}\,\mathbf{V}^{\mathbf{T}}, \qquad (1)$$

where only the first 5 columns of **U** and **V**, and the first 5 singular values are significant. The product

$$\mathbf{A} = (\mathbf{S}\,\mathbf{V}^{\mathbf{T}})^{\mathbf{T}} \qquad (2)$$

is a matrix whose elements are the combination coefficients in the equivalent principal component analysis. Any j-th column of the **D** matrix is the linear combination of the spectral eigenvectors in **U** with the combination coefficients in the j-th row of **A**. A stoichiometric relationship also holds for the rows of **A**.[18] Since the concentration sum of the intermediates is unity (with proper normalization) before the onset of the initial BR state's recovery, the rows of **A** must obey the relationship

$\mathbf{A}_j \mathbf{R} = 1, \qquad j <= k < 40,$ (3)

where k is the index of the last spectrum before BR recovery. Figure 2 shows the left-hand side of Eq. 3 calculated after determining the 5 component vector **R**, which minimizes the deviation of Eq. 3 from unity, with linear least squares procedure. The plot demonstrates that the recovery of BR does not start before the 32nd spectrum (k=32). Combination coefficients of the difference spectra corresponding to the pure intermediates must also obey Eq. 3, i.e. their 5-dimensional points must be located on the "stoichiometric surface" defined by Eq. 3 and **R**.

Determination of the photocycling ratio and the K, L and M spectra with self-modeling. The first 20 difference spectra shown in Figure 1A were treated as a 3-component system. The procedure developed for the D96N mutant BR was applied with minor modifications to compute the K, L and M spectra. A new, truncated matrix consisting of the difference spectra in the >540 nm range was augmented with the negative of the truncated BR absorption spectrum, multiplied by varying trial values of *p*. This latter spectrum is equivalent to the pure M-BR difference spectrum in this spectral interval, where M does not absorb and, therefore, it is expected to be one of the three pure contributions to the mixture difference spectra. In addition, the proper *p* will yield an M-BR spectrum whose combination coefficient point will be on the stoichiometric plane. After SVD, Eq. 3 was solved for each trial *p* value, and the smallest overall deviation of the left-hand side from unity determined the correct *p*, as 0.3.

The full range M-BR difference spectrum multiplied by *p* must be a linear combination of the spectral eigenvectors obtained from the SVD of the first 20 full difference spectra. The proper combination coefficients, by which the M-BR difference spectrum was calculated, was obtained by fitting the >540 nm portion of the $-p$×BR spectrum with the >540 nm part of the SVD eigenvectors.

Figure 3 shows the 3D plot of the combination coefficients (α, β and γ) for the first 20 difference spectra and the vertex of M, as provided by the above procedure. The LM side of the KLM triangle on the stoichiometric plane could be located by moving along lines parallel to the direction defined by the first 4 data points. Points on the LM line were identified by calculating the corresponding difference spectra and monitoring their onset in the red spectral interval.

Proper removal of the contribution by the K-BR spectrum means that there is a range where the resulting difference spectra (L+M mixtures) overlap with the –BR spectrum. This is equivalent to the assumption that there is a region where neither L, nor M absorbs, only K does. The vertex of M and several such points define the LM line, on which the vertex of L is obtained by moving away from M until the average of the difference spectra in the <410 nm range reaches zero. This is equivalent to the assumption that the average absorption of L in this region equals that of BR. Early difference spectra measured on wild type BR before the onset of the M intermediate support this assumption.[26]

The KL side of the KLM triangle had to be located differently than in the case of the D96N mutant. Since in E204Q the accumulation of the M intermediate is faster than in either the WT or the D96N mutant, even the earliest measured difference spectra contain a little M (see Figure 1). Therefore the first several data points in Figure 3 cannot lie on the KL line. The contribution by M could be removed by extrapolating the lines connecting the M vertex with the first 4 data points, until the absorbance in the <410 nm interval diminished (■). The slightly lower absorbance of K relative to BR in this region was ignored by this assumption. The K vertex was located along this line while monitoring the resulting absolute spectrum, obtained by adding the BR spectrum to the difference spectra represented by points on the KL line. K was selected when its maximal amplitude reached 0.86 times the amplitude of the BR spectrum. This value is between the 0.83 and 0.9 values obtained for WT BR with the Monte Carlo method[7] and for D96N with SVD-SM,[19] respectively. The insets in Figure 3 show the difference spectra associated with the K, L and M vertices (the pure K-BR, L-BR and M-BR difference spectra) and, as an illustration, the spectrum associated with the 6th data point, composed of all three intermediates.

Determination of the N and O spectra with self-modeling. The last several data points in Figure 3 lie almost on the LM side of the triangle. This indicates what was also proven by fitting the first 20 difference spectra with the K-BR, L-BR and M-BR difference spectra, that K diminishes completely before the second half of the photocycle represented in Figure 1B. Therefore, only L, M, N and O may contribute to difference spectra No. 21-32 in Figure 4A. Due to the unique absorption of M in the blue region, the M-BR difference spectrum could be

subtracted. Unlike in the case of the KL line, the average of the difference spectra in the <410 nm interval was not assumed to be zero after the proper subtraction of the contribution by M. The inset in Figure 4 shows that a late difference spectrum (No. 37, dashed line) is negative in this region. By adding an appropriately scaled BR spectrum (dotted line) to it, we obtained the solid line, which demonstrates that the 37^{th} spectrum is dominated by O. This enabled an estimate of the ratio between the average (negative) absorption in the blue region and the maximum of the difference spectrum at 640 nm. The ratio was –0.03. Thus, the M-BR difference spectrum was subtracted from each mixture spectrum in Figure 4A until the ratio of the remaining spectrum's average below 410 nm to its amplitude at 640 nm became –0.03. The resulting spectra were then normalized to compensate for the missing amount of M, resulting in the spectra in Figure 4B. These are assumed to be composed of L, N and O only, and should still obey the stoichiometric relationship in Eq. 3. SVD of the matrix plotted in Figure 4B confirmed that it is a 3 component system, and the stoichiometric plane was also determined. The adherence of the data points to the stoichiometric plane is demonstrated in Figure 5. Since the L spectrum has already been determined, the corresponding vertex of the LNO triangle is known. The LN side of the triangle was located similarly to the LM side of the KLM triangle in Figure 3. Based on previous knowledge about the intermediate spectra we assumed that only the O intermediate absorbs in the >700 nm spectral range. By moving on the stoichiometric plane along lines parallel to the direction defined by data points 32...28, several points were located on the LN line, based on the disappearance of absorption in this spectral range. Because of the uncertainty of the absorption maxima of the calculated spectra due to noise, the first moment of the (absolute) spectra along the LN line served as the criterion for locating the N vertex. There is a certain arbitrariness in the selection of the value of the first moment, but the subsequent determination of the O spectrum, which depends on the accuracy of the N spectrum (on the location of the N vertex) helped in removing much of the uncertainty associated with the N spectrum.

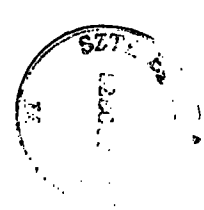
Once the N vertex is located, there is an inner limit to the NO line, set by the last several data points, since all data points must fall within the LNO triangle. The assumption that L has entirely decayed by spectrum 32 is equivalent to the assumption that data point No. 32 lies on the NO side (solid NO line in Figure 5). This was proved in retrospect by calculating with a slightly inclined NO' side (dashed line), and using the resulting O' spectrum in fitting the data. Although

the O' spectrum was only slightly different from the O spectrum, the so obtained intermediate kinetics proved to be unrealistic: the concentration of L first decayed and then started to increase again.

The vertex of O was located by setting the maximum of the resulting O spectrum to 1.1 times that of the BR spectrum. This is different from the WT spectrum calculated by the Monte Carlo method (Table I). There the amplitude of O was 98% of BR. Trial O spectra were calculated along the NO line with amplitudes less than, or greater than 1.1 times the BR amplitude. The former (including the factor of 0.98) yielded spectra with a shoulder on the blue side, the latter provided unrealistically tall and narrow spectra. There is of course some ambiguity regarding the selected amplitude of 1.1, nevertheless it can be stated that the spectrum of O is substantially different in the E204Q mutant and the WT. Insets in Figure 5 show the difference spectra of the pure L, N and O intermediates, and for comparison a selected mixture spectrum corresponding to the 26^{th} data point.

The intermediate spectra were calculated at all 3 *p*H values and all 3 temperatures with SVD-SM in a similar manner. The N spectra were blue shifted by ca. 2 nm at *p*H 9, otherwise all spectra proved to be *p*H and temperature-independent. The *p*H and temperature average of the difference and absolute spectra are plotted in Figure 6. Parameters of the absolute spectra in the E204Q mutant are compared with WT in Table I. The positions of the absorption maxima are within 2 nm from the WT, except for N, which is 10 nm blue shifted. There are more deviations in the maximal amplitudes, especially for M and O. The half widths of the spectra are similar, except for the O spectrum due to the slightly different shape of the WT O spectrum obtained with the Monte Carlo method (see Figure 5 in ref. 7).

Kinetics of the intermediates. The difference spectra of the pure intermediates (Figure 6A) were used to fit the spectra in Figure 1 with a nonnegative least squares routine. The first 18 spectra were fitted with the K, L and M difference spectra alone, and spectra No. 22-40 with L, M, N and O, in accordance with the findings that the first 20 spectra constitute a 3 component system, and that the K intermediate decays before the second 20 time points. Spectra No. 19-21 were fitted with all 5 intermediate spectra, thereby allowing marginal amounts of K, as well as N and O. The resulting intermediate kinetics are plotted in Figure 7 (symbols). The characteristic



features of the E204Q intermediate kinetics as compared to WT are the early rise of M, the high yield of O and the slow recovery of BR. The kinetics were obtained similarly at the other temperature and *p*H values (not shown). The fit of the measured spectra with the intermediate spectra were excellent in each case, and the sum of the intermediate concentrations before BR recovery yielded values close to unity, as reflected by the calculated BR concentration in Figure 7, obtained as 1 minus this sum.

Similar procedure with the alternative O' spectrum resulted in a rise of the amount of L after its partial decay, and in some cases residual L even after the total decay of M. Because of these anomalies the O' spectrum was rejected. This justified in retrospect the selection of the NO line during self-modeling as the line connecting the N vertex with the outermost experimental point in Figure 5.

Various photocycle schemes were tested with increasing complexity to fit the intermediate kinetics. Schemes with a single M or a single L intermediate fit the data poorly, as expected. The simplest scheme with two L and two M intermediates is the sequential Scheme I:

$$BR \rightarrow K \leftrightharpoons L_1 \leftrightharpoons L_2 \leftrightharpoons M_1 \leftrightharpoons M_2 \leftrightharpoons N \leftrightharpoons O \rightarrow BR \qquad \text{(S-I)}$$

The fitted kinetics, with $\chi^2 = 0.0852$, are shown as dashed lines in Figure 7. There are systematic deviations from the experimental data especially at the maximum of the M concentration and in the decaying phase of the L concentration. Similar discrepancies were obtained at the other temperature and *p*H values. A better fit with $\chi^2 = 0.0424$ was obtained using Scheme II (solid lines), which contains L_2 in a cul-de-sac as proposed earlier for the D96N mutant,[17] and which gave a similarly good fit to D96N data as Scheme I in another study[27]:

$$\begin{array}{l} BR \rightarrow K \leftrightharpoons L_1 \leftrightharpoons M_1 \leftrightharpoons M_2 \leftrightharpoons N \leftrightharpoons O \rightarrow BR \\ \qquad\qquad \Updownarrow \\ \qquad\qquad L_2 \end{array} \qquad \text{(S-II)}$$

The $M_1 \leftarrow M_2$ back reaction turned out to be more than 3 orders of magnitude slower than the $M_1 \rightarrow M_2$ forward reaction, and removing it from the scheme did not affect the other rate constants.

In the same manner, the $M_1 \leftarrow M_2$ back reaction was negligible at all *p*H and temperature values, as found for the WT photocycle at alkaline *p*H.

The rate constants obtained with Scheme II display Arrhenius behavior. There is no pH dependence of the rate constants up to intermediate N, except for a non-systematic dependence of the $L_2 \rightarrow M_1$ transition (not shown). This rate constant could be forced to be *p*H-independent without compromising the quality of the fit, however, this required the reintroduction of the $M_1 \leftarrow M_2$ back reaction, whose value, although still orders of magnitudes lower than the corresponding forward reaction, depended non-monotonously on *p*H. Further complicating the photocycle scheme with additional intermediate substates and pathways could probably remove these uncertainties, however, the possibilities are numerous and may not yield conclusive evidence. Instead, we concentrate on the systematic *p*H dependence at all 3 temperatures of the $N \rightarrow O$ and the $N \leftarrow O$ rates, which play a key role in the kinetics of the O intermediate.

Factors affecting the kinetics of the O intermediate. Perhaps the most enigmatic in the bacteriorhodopsin photocycle is the time-dependence and yield of the O form. Ottolenghi, Sheves and coworkers[28,29] reported that the titration of 3 amino acids may control the rates of the rise and decay as well as the maximal accumulation of O. This work was based on the *p*H dependence of phenomenological rate constants, thereby avoiding model-dependent conclusions. Balashov, Ebrey and coworkers,[30-32] however, fitted the measured O kinetics both with sums of exponentials to obtain phenomenological rates, and with the model $M \leftrightharpoons N \leftrightharpoons O \rightarrow BR$, which provides the *p*H dependence of the model-dependent intrinsic rate constants. Their conclusion was that the rate of O-rise depends on the titration of D96 in the N intermediate, as the proton uptake by D96 on the cytoplasmic side may be coupled to the subsequent reisomerization of the retinal. The decay of O to BR was suggested to depend on the titration of the proton release cluster on the cytoplasmic side in the O form. According to their model, above this *p*K proton release from this cluster is followed by its reprotonation by D85, a fast process, whereas below this *p*K proton release from the cluster is inhibited, and release must proceed directly from D85, a slower process. This *p*H-dependence should be absent in the E204Q mutant, and the decay of O should be slow, in accordance with our observations.

In the present work we obtained not only the kinetics of O, but also those of the precursors to it, N and M. Figure 8A (symbols) shows the kinetics of O (the same as in Figure 7). It was fitted with Scheme III:

$$N \leftrightharpoons O \rightarrow BR \qquad \text{(S-III)}$$

The fit to the O kinetics (solid line) is acceptable, however, it assumes that the precursor to O follows the kinetics represented by the dashed line. The rate constants at *p*H 5, 15 °C were: 34 s^{-1}, 0.3 s^{-1} and 8.8 s^{-1} for N → O, N ← O and O → BR, respectively. The same O kinetics, actually with a better fit, could be obtained if the precursor to O, the N intermediate was taken into account correctly (Figure 8B). The corresponding rate constants were 86 s^{-1}, 11 s^{-1} and 8.7 s^{-1}, considerably different from the previous values. The accurate determination of the rate constants in this segment of the photocycle requires the knowledge of the kinetics of not only O (measured usually at 630 nm), but the time evolution of its precursors, too. The latter can only be determined reliably from multichannel spectroscopy, with the global analysis of the data matrix by methods such as SVD-SM.

Figure 9 shows that the rates of both the N → O and N ← O transitions decrease with increasing *p*H. This behaviour can be explained by assuming that the transition state barrier between these two intermediates is increasing with increasing *p*H. Since the O → BR rate is *p*H-independent, the free energy of the O state is probably also *p*H-independent, although, due to the unidirectionality of this reaction, we cannot estimate the barrier for the O ← BR transition. From the Eyring plot of the corresponding rate constants (not shown) we estimated the relative free energies of N and O and the barrier heights as a function of *p*H (Table 2). Although there is a little shift of the free energy content of N meaning its stabilization with increasing *p*H, the main factor controlling the rise, decay and yield of O as a function of pH and temperature is the *p*H-dependence of the barrier height between the N and the O intermediates. If this barrier height is controlled by the titration of an amino acid residue, its *p*K should be close to 8, based on the similar pH-dependence of the N → O and N ← O rates (Figure 9).

Discussion

The proton pumping activity of bacteriorhodopsin is accompanied by structural changes which are reflected in spectral changes both in the UV-visible and the infrared region. The elusive nature of the photocycle stems from the overlap of the lifetimes of the successive intermediates and from the spectral overlap of their visible absorption. Intensive research has demonstrated that there are five metastable intermediates with characteristic visible spectra (excluding the ultrafast "hot" J form and its precursors) but, in addition, there exist kinetically distinct substates of the intermediates with no or little spectral differences. The wild type photocycle is also complicated by a *p*H-dependent branching at the level of the M intermediate,[22] and there may also be branching from the N intermediate directly to BR.[26] It is therefore a non-trivial task to determine the correct intermediate kinetics from absorption spectroscopy. The D96N mutant provided a simpler system with 3 main intermediates only (K, L and M), and we have recently developed the method of singular value decomposition with self-modeling (SVD-SM) to obtain accurate intermediate spectra for this three-component system, which defined their time evolutions and helped in the selection of a viable photocycle model. The superior power of SVD-SM over earlier analytical approaches is reflected in that it revealed for the first time biphasic recovery of BR in the D96N mutant.

The E204Q mutant lacks early proton release due to the inhibition of the extracellular proton release cluster. Here we expect no *p*H-dependent branching, and published data demonstrate that the yield of the O intermediate is much higher than in WT. The photocycle is otherwise similar to WT, with the usual five intermediates (K, L, M, N and O) following each other in sequence. We used this mutant to further develop the SVD-SM method, and to study its photocycle kinetics, especially the elusive N-O-BR segment, as a prelude to applying the method to the more difficult WT BR problem.

SVD-SM combines abstract spectral eigenvectors to compute the unknown spectra of the intermediates. The stoichiometric constraint reduces the dimension of the search by one, by restricting the search for the intermediate spectra to the corresponding stoichiometric surface. The E204Q mutant turned out to be an optimal target for SVD-SM not only because of the aforementioned reasons, but also because according to the stoichiometric analysis, it lacks the complication of early partial BR recovery found in the D96N mutant. Another advantage of

E204Q became evident during the analysis: the photocycle could be divided into two time domains so that the first domain consisted only of the K, L and M intermediates, as in D96N, whereas the second domain lacked the contribution by K, representing a 4, rather than 5 component system. This is not the case in WT, where at *p*H below neutral some K intermediate remains during the onset of N .[33]

The spectrum of M is the easiest to determine and is the most accurate due to its substantially blue-shifted maximum and zero absorption over much of the visible region. Other intermediate spectra carry some ambiguity, as to their exact location, maximal absorption and half-width. This is the result of the arbitrary choice of the parameter values used as criteria in finding these spectra on the one hand, and of the noise content of the data on the other. Slightly different spectra could be obtained in our analysis if we allowed deviation of the L and the K spectra from that of BR in the blue region, or a different height for K. There is some uncertainty in the choice of the position of the N spectrum, measured by its first moment. Finally, the height of the O spectrum could also be selected somewhat differently. Inspection of the calculated spectra during the self-modeling procedure rules out suspiciously narrow or broad candidates, or those carrying signs of "contamination" by other intermediates, for instance mixing of N into the O spectrum. Nevertheless, a narrow range of spectra around the selected ones would be equally acceptable. One strong criterion is the condition that all measured data points be within the triangle (or polyhedron of higher dimensions) defined by the pure intermediate vertices, in order to avoid negative intermediate concentrations. It is reassuring that the intermediate spectra obtained from the 9 separate experiments independently were almost identical, and their average (the final spectra in Figure 6) fitted the experimental difference spectra accurately. In addition, the fits yielded nonnegative concentration values, smooth intermediate kinetics and no systematic deviation of the concentration sums from unity before the onset of BR recovery. It was also clear that small changes in the intermediate spectra can result in dramatic changes in the kinetics. Especially sensitive is the time-dependence of the L and N intermediates, whose spectra are rather similar. This limits the reliability of the intermediate concentrations and the results obtained by fitting photocycle schemes. Therefore, we have not pursued a global kinetic and thermodynamic analysis in the present work, and focused on a few robust features instead.

Due to the genetic alteration of the proton release cluster the photocycle kinetics change in the 1-100 μs time domain compared to WT. The altered L and M kinetics enabled us to compare two previously suggested schemes which differ in the location of the two L substates. Although the sequential and the cul-de-sac models (S-I and S-II, respectively) gave similarly good fits for WT and D96N, the E204Q mutant clearly distinguishes between them in favor of S-II. L_2 in equilibrium with M_1 must differ from L_1 in some respect, which cannot be the protonation state of the proton release cluster. One possibility would be the change of accessibility of the Schiff base region during the lifetime of M_1. High resolution structures for an early M (M_1) in E204Q[14] and a late M (M_2 or M_N) in D96N[12] indicate a gradual rearrangement of the Schiff base from the original extracellular orientation to the fully relaxed cytoplasmic orientation. The hydrogen bond between the Schiff base and D85 via a water molecule in BR and, presumably, in L_1, is replaced by a hydrogen bond connecting the Schiff base to T89 instead. These, and other structural motions involving water molecules would permit a L_2 form different from L_1, in equilibrium with the early, but already partially reoriented M intermediate reported in E204Q. This equilibrium, however, must be fully shifted towards deprotonated Schiff base by the formation of M_2, since the alternative model with L_2 in a cul-de-sac equilibrium with M_2 (rather than M_1) gives unacceptable fits.

Another interesting issue is the yield and kinetics of the O intermediate, and its strong dependence on *p*H and temperature. The E204Q mutant is especially suitable to study this problem due to the high yield of O and its slow recovery. The O intermediate in WT has been reported to contain the *all-trans*-retinyl moiety (the only such intermediate besides BR),[34] with D85 still protonated,[35] and with a still deprotonated release cluster in WT. The O in E204Q also contains all-trans retinal and protonated D85,[36] but there cannot be a proton deficiency in the release cluster, as in WT at acidic *p*H. This protonation (charge and/or hydrogen bonding) difference may be reflected in the higher amplitude of the O spectrum in E204Q.

Accurate determination of the intrinsic rate constants leading to and away from the O intermediate require the knowledge of the kinetics of the precursor to O. As long as one assumes that O is formed from N and decays to BR, allowing a back-reaction to N, the kinetics of N and O are sufficient to obtain these rates, independent of the reactions leading to N, since the differential equation describing the kinetics of O involves the time-dependent amount of N and

the relevant rate constants only. Multichannel spectroscopy with the method of SVD-SM reliably provides the N and O kinetics. The analysis showed that both the N → O and the N ← O rates decrease with increasing *p*H, while others are *p*H-independent. The estimated *p*K for the control of both rates is around 8. The simultaneous *p*H-dependencies of the forward and reverse rates suggest that the barrier height between the two intermediates is *p*H-dependent, rather than the free energy levels of the intermediates themselves. It appears that the reisomerization of the retinal, the key structural change in the N → O transition, may be catalysed by a group with this *p*K, so that protonation of this residue would accelerate reisomerization. It is known that retinyl isomerization is affected by the charge environment of the protein. For instance, trans-to-cis isomerization after light absorption is also controlled by the charge of the retinyl environment: in several mutants and in the acid forms of BR (with protonated D85) the isomerization is slower than in the native pigment.[37,38]

It has been suggested[31] that at *p*H lower than the transient *p*K of the proton release cluster in the O form of WT, proton release from D85 takes place directly to the membrane surface, without the involvement of the release cluster. This is a slow process, consistent with the O decay in the E204Q mutant. This second *p*K plays of course no role in the lifetime and yield of O in E204Q, rendering their *p*H-dependence less complicated. Our results with the E204Q mutant are in qualitative agreement with previous results on this mutant.[23,24] In addition, detailed analysis of the O kinetics enabled us to propose a single reason for the complex behaviour of O as a function of *p*H and temperature. Complications are expected in WT due to the alternative deprotonation routes of D85 during the decay of O and recovery of BR, depending on the protonation state of the release cluster.

Acknowledgment

This work was supported by the National Scientific Research Fund of Hungary (OTKA T020470). Work at Florida State University was supported by NSF Grant CHE 9985895.

References

(1) Heberle, J. *Biochim. Biophys. Acta* **2000**, *1458*, 135-147.

(2) Stoeckenius, W. *Protein Sci.* **1999**, *8*, 447-459.

(3) Lanyi, J.K. *J. Biol. Chem.* **1997**, *272*, 31209-31212.

(4) Becher, B.; Tokunaga, F.; Ebrey, T.G. *Biochemistry* **1978**, *17*, 2293-2300.

(5) Diller, R. ; Stockburger, M. *Biochemistry* **1988**, *27*, 7641-7651.

(6) Gerwert, K. *Biochim. Biophys. Acta* **1992**, *1101*, 147-153.

(7) Gergely, C.; Zimányi, L.; Váró, G. *J. Phys. Chem. B* **1997**, *101*, 9390-9395.

(8) Braiman, M.S.; Mogi, T.; Marti, T.; Stern, L.J.; Khorana, H.G.; Rothschild, K.J. *Biochemistry* **1988**, *27*, 8516-8520.

(9) Luecke, H.; Schobert, B.; Richter, H.-T., Cartailler, J-P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899-911.

(10) Subramaniam, S.; Lindahl, M., Bullough, P.; Faruqi, A.R.; Tittor, J.; Oesterhelt, D., Brown, L.S.; Lanyi, J.K.; Henderson, R. *J. Mol. Biol.* **1999**, *287*, 145-161.

(11) Edman, K.; Nollert, P.; Royant, A.; Belrhali, H.; Pebay-Peyroula, E.; Hajdu, J.; Neutze, R.; Landau, E. M. *Nature* **1999**, *401*, 822-826.

(12) Luecke, H.; Schobert, B.; Richter, H.-T.; Cartailler, J-P.; Lanyi, J.K. *Science* **1999**, *286*, 255-260.

(13) Sass, H.J.; Büldt, G.; Gessenich, R.; Hehn, D.; Neff, D.; Schlesinger, R.; Berendzen, J.; Ormos, P. *Nature* **2000**, *in press*

(14) Luecke, H.; Schobert, B.; Cartailler, J-P.; Richter, H.-T.; Rosengarth, A.; Needleman, R.; Lanyi, J.K. *J. Mol. Biol.* **2000**, *300*, 1237-1255.

(15) Butt, H.-J.; Fendler, K.; Bamberg, E.; Tittor, J.; Oesterhelt, D. *EMBO J.* **1989**, *8*, 1657-1663.

(16) Tittor, J.; Soell, C.; Oesterhelt, D.; Butt, H.-J.; Bamberg, E. *EMBO J.* **1989**, *8*, 3477-3482

(17) Zimányi, L.; Lanyi, J.K. *Biophys. J.* **1993**, *64*, 240-251.

(18) Zimányi, L.; Kulcsár, Á.; Lanyi, J.K.; Sears, D.F.; Saltiel, J. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4408-4413.

(19) Zimányi, L.; Kulcsár, Á.; Lanyi, J.K.; Sears, D.F.; Saltiel, J. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4414-4419.

(20) Rammelsberg, R.; Huhn, G.; Lübben, M.; Gerwert, K. *Biochemistry* **1998**, *37*, 5001-5009.

(21) Váró, G.; Lanyi, J.K. *Biochemistry* **1990**, *29*, 6858-6865.

(22) Zimányi, L.; Váró, G.; Chang, M.; Ni, B.; Needleman, R.; Lanyi, J.K. *Biochemistry* **1992**, *31*, 8535-8543.

(23) Brown, L.S.; Sasaki, J.; Kandori, H.; Maeda, A.; Needleman, R.; Lanyi, J.K. *J. Biol. Chem.* **1995**, *270*, 27122-27126.

(24) Richter, H.-T.; Brown, L.S.; Needleman, R.; Lanyi, J.K. *Biochemistry* **1996**, *35*, 4054-4062.

(25) Henry, E.R.; Hofrichter, J., *Methods. Enzymol.* **1992**, *210*, 129-192.

(26) Váró, G.; Lanyi, J.K. *Biochemistry* **1991**, *30*, 5008-5015.

(27) Gergely, C.; Ganea, C.; Groma, G.; Váró, G. *Biophys. J.* **1993**, *65*, 2478-2483.

(28) Bressler, S.; Friedman, N.; Li, Q.; Ottolenghi, M.; Saha, C.; Sheves, M. *Biochemistry* **1999**, *38*, 2018-2025.

(29) Li, Q.; Bressler, S.; Ovrutsky, D.; Ottolenghi, M.; Friedman, N.; Sheves, M. *Biophys. J.* **2000**, *78*, 354-362.

(30) Misra, S.; Govindjee, R.; Ebrey, T.G.; Chen; N.; Ma, J.-X.; Crouch, R.K. *Biochemistry* **1997**, *36*, 4875-4883.

(31) Balashov, S.P.; Lu, M.; Imasheva, E.S.; Govindjee, R.; Ebrey, T.G.; Othersen, B.; Chen, Y.; Crouch, R.K.; Menick, D.R. *Biochemistry* **1999**, *38*, 2026-2039.

(32) Lu, M.; Balashov, S.P.; Ebrey, T.G.; Chen, N.; Chen, Y.; Menick, D.R.; Crouch, R.K. *Biochemistry* **2000**, *39*, 2325-2331.

(33) Ludmann, K.; Gergely, C.; Váró, G. *Biophys. J.* **1998**, *75*, 3110-3119.

(34) Smith, S.O.; Pardoen, J.A.; Mulder, P.P.J.; Curry, B.; Lugtenburg, J.; Mathies, R.A. *Biochemistry* **1983**, *22*, 6141-6148.

(35) Braiman, M.S.; Bousché, O.; Rothschild, K.J. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2388-2392.

(36) Kandori, H.; Yamazaki, Y.; Hatanaka, M.; Needleman, R.; Brown, L.S.; Richter, H.-T.; Lanyi, J.K.; Maeda, A. *Biochemistry* **1997**, *36*, 5134-5141.

(37) Logunov, S.L.; El-Sayed, M.A.; Lanyi, J.K. *Biophys. J.* **1996**, *71*, 1545-1553.

(38) Song, L.; El-Sayed, M.A.; Lanyi, J.K. *J. Phys. Chem.* **1996**, *100*, 10479-10481.

Figure Captions

Figure 1. Flash-induced difference spectra of E204Q bacteriorhodopsin at *p*H = 5, *T* = 15 °C. A: spectra No. 1-20, time domain: 520 ns – 370 μs, and the absorption spectrum of the initial form (dashed line). B: spectra No. 21-40, time domain: 520 μs – 800 ms. Arrows indicate the time-evolution of the signal. The spectra are reconstituted using the 5 most significant SVD vectors.

Figure 2. The stoichiometric behavior of the data from Figure 1. The left-hand side of eq 3 is plotted versus time. Data points represent the difference spectra in Figure 1.

Figure 3. Demonstration of self-modeling for the first half of the photocycle. Three dimensional plot of the combination coefficients in matrix A (•) representing the original difference spectra in Figure 1A. The points which determine the K-L line and the vertices corresponding to the pure difference spectra of the K, L and M intermediates (■) outline the stoichiometric plane. Insets: The difference spectra corresponding to the pure intermediates and a mixture difference spectrum (spectrum No. 6).

Figure 4. A: Difference spectra No. 21-32 from Figure 1B, which serve as the basis to determine the spectra of the pure N and O intermediates. B: Spectra from A after subtraction of the M intermediate. The resulting spectra are assumed to be contributed by the L, N and O intermediates. For details of the subtraction, see text. Inset: Difference spectrum No. 37 (dashed line), properly normalized, so that the addition of the BR absorption spectrum (dotted line) yields an absorption spectrum (solid line), which is dominated by the O intermediate.

Figure 5. Demonstration of self-modeling on the difference spectra in Figure 4B. Three dimensional plot of the combination coefficients in matrix A (•) and the vertices corresponding to the pure difference spectra of the L, N and O intermediates (■) outlining the stoichiometric plane. Insets: The difference spectra corresponding to the pure intermediates and a mixture difference spectrum (No. 26).

Figure 6. Difference spectra (A) and absolute spectra (B) of the photocycle intermediates K, L, M, N and O, as determined by SVD-SM, and the spectrum of the initial state BR. The spectra are the average of spectra determined from 9 independent experiments on E204Q BR at *p*H 5, 7 and 9, T = 15, 22 and 29 °C.

Figure 7. Kinetics of the photocycle intermediates of E204Q BR at *p*H 5, T = 15 °C (symbols). Dashed lines represent the fit by the photocycle scheme S-I, solid lines the fit by S-II (see text).

Figure 8. Fit of the kinetics of the O intermediate (symbols) separately, by the simplified scheme S-III (A). The assumed kinetics of the precursor to O (N, or the equilibrated M+N) is shown by the dashed line. B: Fit of the kinetics of the O intermediate by S-II, which correctly accounts for the time evolution of the precursor to O, the N intermediate. Both methods yield accurate O kinetics, but with different rate constants (see text).

Figure 9. Dependence of the rate constants on *p*H and temperature, obtained from the fit of the photocycle scheme S-II to the intermediates kinetics.

Table 1. Parameters of the Intermediate Spectra of the E204Q Mutant[a] and of Wild-Type Bacteriorhodopsin[b]

	E204Q			WT		
	λ_{max} (nm)	A_{max} (rel. units)	FWHM (cm^{-1})	λ_{max} (nm)	A_{max} (rel. units)	FWHM (cm^{-1})
BR	568	1	21.1×10^3	570	1	20.4×10^3
K	588	0.85	22.6×10^3	586	0.83	22.3×10^3
L	542	0.79	24.1×10^3	544	0.77	24.3×10^3
M	410	0.88	33.6×10^3	409	0.77	33.3×10^3
N	552	0.79	23.0×10^3	562	0.73	22.6×10^3
O	630	1.1	20.9×10^3	629	0.98	14.6×10^3

[a]This work. [b]Reference 7.

Table 2. Free Energy of the N Intermediate Relative to the O Intermediate, and Activation Free Energies of the N ⇋ O → BR Segment of the Photocycle Estimated at *p*H 5, 7 and 9

	ΔG (kJ/mol)			$\Delta G^{\ddagger}$ (kJ/mol)		
	*p*H 5	*p*H 7	*p*H 9	*p*H 5	*p*H 7	*p*H 9
N	5.2	4.2	1.6			
N → O				59.4	61.1	67.3
N ← O				64.6	65.3	68.9
O → BR				69.6	69.8	69.6

Figure 1

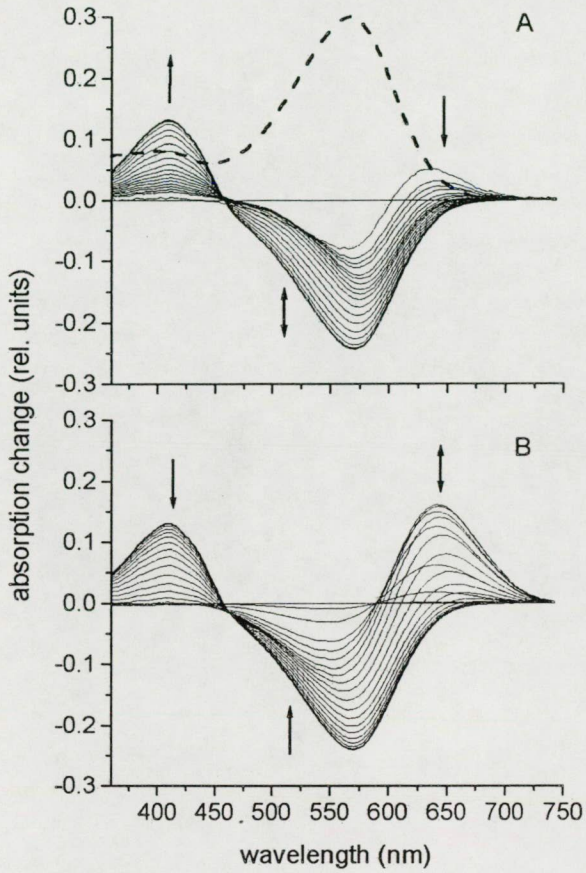


Figure 2

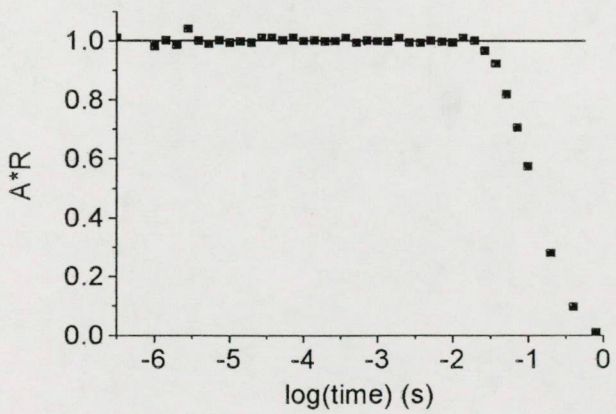


Figure 3

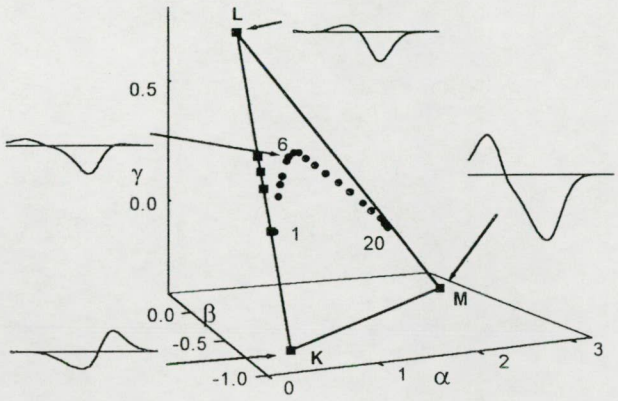


Figure 4

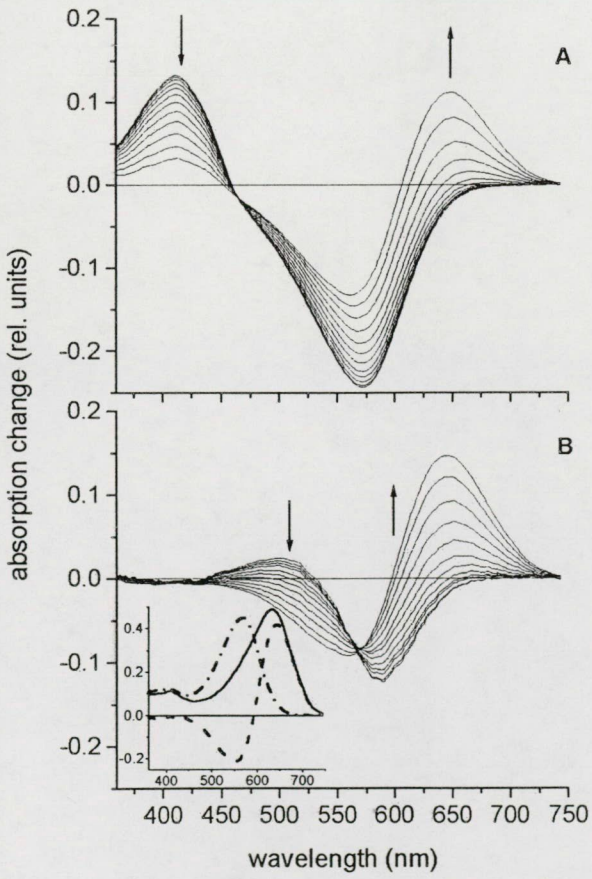


Figure 5

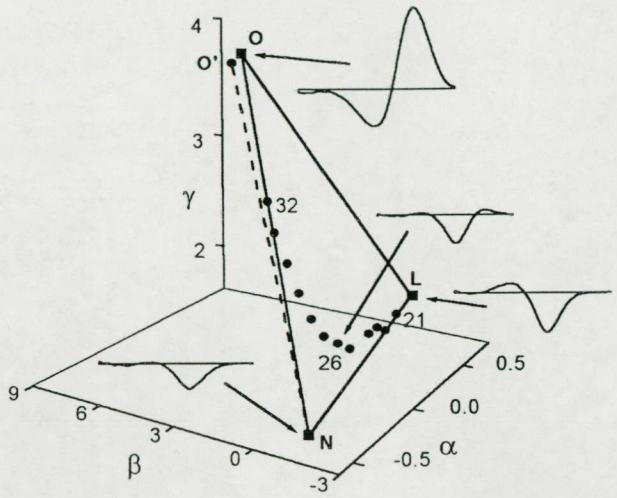


Figure 6

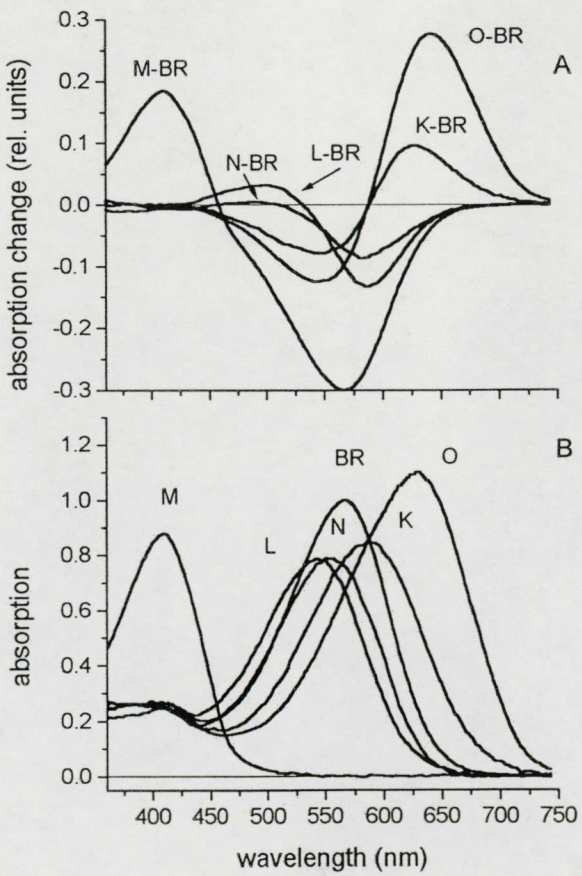


Figure 7

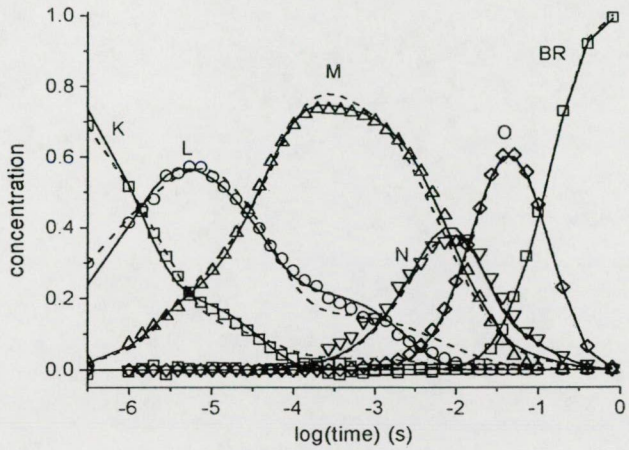


Figure 8

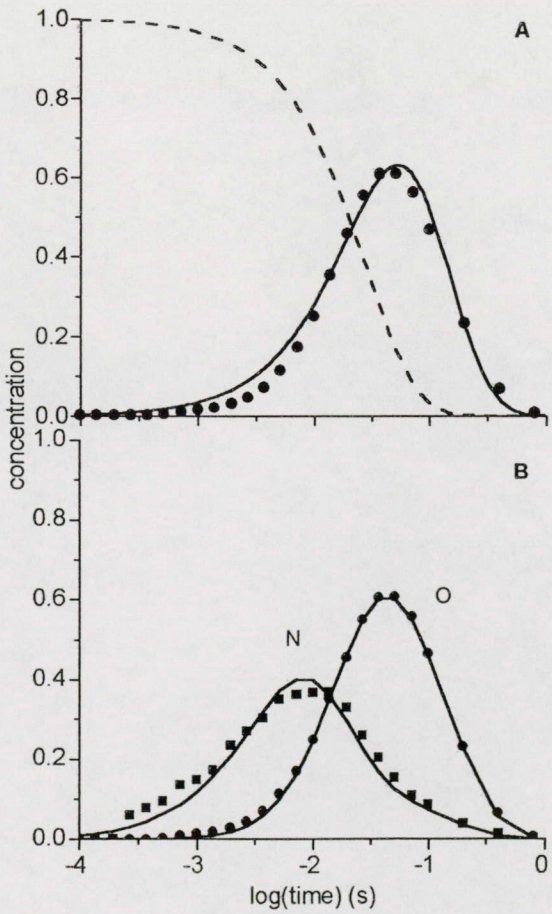
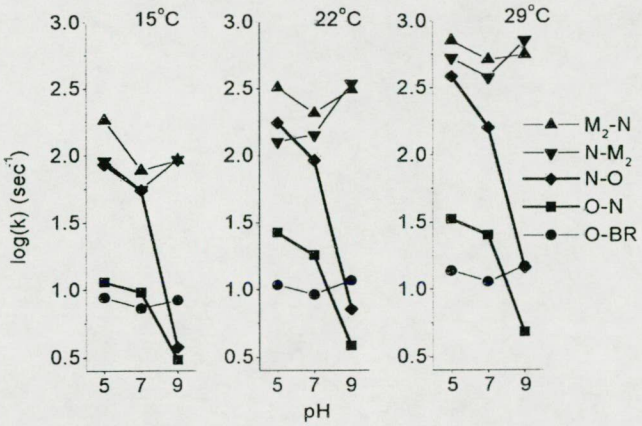


Figure 9



Characterization of the Proton-Transporting Photocycle of Pharaonis Halorhodopsin

Ágnes Kulcsár,* Géza I. Groma,* Janos K. Lanyi,† and György Váró*
*Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, Szeged, H-6701, Hungary; and †Department of Physiology & Biophysics, University of California, Irvine, California 92697 USA

ABSTRACT The photocycle of pharaonis halorhodopsin was investigated in the presence of 100 mM NaN_3 and 1 M Na_2SO_4. Recent observations established that the replacement of the chloride ion with azide transforms the photocycle from a chloride-transporting one into a proton-transporting one. Kinetic analysis proves that the photocycle is very similar to that of bacteriorhodopsin. After K and L, intermediate M appears, which is missing from the chloride-transporting photocycle. In this intermediate the retinal Schiff base deprotonates. The rise of M in halorhodopsin is in the microsecond range, but occurs later than in bacteriorhodopsin, and its decay is more accentuated multiphasic. Intermediate N cannot be detected, but a large amount of O accumulates. The multiphasic character of the last step of the photocycle could be explained by the existence of a HR' state, as in the chloride photocycle. Upon replacement of chloride ion with azide, the fast electric signal changes its sign from positive to negative, and becomes similar to that detected in bacteriorhodopsin. The photocycle is enthalpy-driven, as is the chloride photocycle of halorhodopsin. These observations suggest that, while the basic charge translocation steps become identical to those in bacteriorhodopsin, the storage and utilization of energy during the photocycle remains unchanged by exchanging chloride with azide.

INTRODUCTION

Pharaonis halorhodopsin (pHR) is a small integral membrane protein in the haloalkaliphile *Natronobacterium pharaonis* (Bivin and Stoeckenius, 1986). It belongs to the family of retinal proteins, like bacteriorhodopsin (BR) and salinarum halorhodopsin (sHR) from *Halobacterium salinarum*. Upon photoexcitation it transports chloride ion into the cell (Duschl et al., 1990). Amino acid sequence comparison reveals a 25% homology of pHR with the proton transporter BR and 66% with the chloride transporter sHR (Lanyi et al., 1990). The extent of identity is greater in the putative seven transmembrane helices, and the residues in the retinal binding pocket are particularly highly conserved. The important exceptions are the proton acceptor Asp-85 and proton donor Asp-96 in BR, which are replaced by Thr and Ala, respectively, in both halorhodopsins. Lack of the proton acceptor ensures that the Schiff base remains protonated during the photocycle while chloride is transported through the membrane.

No structural information exists about pHR, but its sequential identity and functional similarity with sHR makes it probable that it has a similar structure, i.e., seven transmembrane helices surrounding the retinal binding pocket (Havelka et al., 1995). The retinal is bound to Lys-256 via a protonated Schiff base in all-*trans* or 13-*cis* configuration. Contrary to BR and sHR, its retinal isomer composition is independent of illumination, containing 85% all-*trans* and 15% 13-*cis* retinal (Váró et al., 1995a; Zimányi and Lanyi, 1997). Although the protein suspension is not homogeneous in retinal conformation, there is no evidence that the 13-*cis*-containing protein is photoactive (Váró et al., 1995a).

Spectral titration of pHR with sodium chloride yields a chloride-binding constant of 1 mM, characterized by a 13-nm blue shift of the absorption peak (Scharf and Engelhard, 1994; Váró et al., 1995a). Other halide ions also bind to the protein with similar spectral shifts but different binding constant (Scharf and Engelhard, 1994). Raman and FTIR spectroscopic studies proved the existence of an interaction between the retinal Schiff base and halide ion (Walter and Braiman, 1994; Gerscher et al., 1997). Titration with azide causes almost the same spectral change, with a binding constant of 10 mM, and chloride and azide ions compete for the same binding site (Váró et al., 1995a). In salt-free solution the protein denaturates. For this reason, during titration the sodium content of the solution is kept constant by supplementing the titrated salt with sodium sulfate, which itself does not bind to the protein (Scharf and Engelhard, 1994; Váró et al., 1995a).

Azide was used as a substitute for the proton donor in the study of the D96N BR mutant (Tittor et al. 1989; Zimányi and Lanyi, 1993). In several cases azide could also replace the role of the proton acceptor Asp-85 (Tittor et al., 1994; Dickopf et al., 1995). In wild type BR azide has an observable effect on the decay of intermediate M only at very high concentrations (Ormos et al., 1997). In sHR the presence of azide resulted a side-reaction of the photocycle, yielding deprotonation of the Schiff base (Hegemann et al., 1985). In pHR azide could replace both the proton donor and acceptor in chloride-free solution and lead to a proton transporting photocycle (Váró et al., 1996).

Received for publication 28 March 2000 and in final form 21 July 2000.
Address reprint requests to Dr. Gyorgy Varo, Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, Szeged, H-6701, Hungary. Tel.: 36-62-432232; Fax: 36-62-43331333; E-mail: varo@nucleus.szbk.u-szeged.hu.

0006-3495/00/11/2705/09 $2.00

The chloride-transporting photocycle of halorhodopsin has intermediates analogous to the BR photocycle, but without an M intermediate containing a deprotonated Schiff base (Váró et al., 1995a,b). The only important difference between the photocycle of the two halorhodopsins is that in the photocycle of sHR the intermediate O was not observed (Váró et al., 1995a), whereas in the photocycle of pHR this intermediate accumulates in rather large amounts (Váró et al., 1995b). In the photocycle of sHR the observed red-shifted species, which earlier was assigned as intermediate O, belongs to the non-transporting 13-*cis* photocycle (Váró et al., 1995a). There is no evidence to decide if intermediate O is really missing from the photocycle, or for kinetic reasons does not accumulate. Based on the multiexponential decay of the pHR photocycle, the HR′ state was introduced as the last intermediate, with essentially the same spectrum as HR (Váró et al., 1995a). When chloride is replaced with azide, the M intermediate appears in the photocycle of pHR (Váró et al., 1996). There are other conditions, when an M-like intermediate appears in the halorhodopsin photocycle, such as double flash excitation (Bamberg et al., 1993) or prolonged illumination (Hegemann et al. 1985), when both azide and chloride are present in the sample. These M-like intermediates have a very long lifetime and transport protons across the membrane only when the illumination is with both green and blue light. From the temperature dependence of the photocycle, the free energy, enthalpy and entropy changes of the reactions were determined. The pH dependence of the energetics of BR photocycle is complex, but the driving force of the reactions is mostly the entropy decrease (Ludmann et al., 1998b). In contrast, the photocycle reactions of pHR are enthalpy-driven (Váró et al., 1995b).

It has been suggested that there must be an analogy between proton transport in BR and chloride transport in HR (Oesterhelt and Tittor, 1989; Oesterhelt et al., 1992; Haupts et al., 1997; Brown et al., 1998; Váró, 2000). This implies the possibility that either protein could be converted to the ion specificity of the other. When Thr replaced the proton acceptor Asp-85, the function of BR changed from proton to chloride translocation (Sasaki et al., 1995). The same effect could be achieved by neutralizing the proton acceptor at low pH. The pK_a of Asp-85 is approximately 2.6 (Balashov et al., 1991; Balashov et al., 1995). At pH below 2, with chloride ion present, BR transports chloride through the membrane (Dér et al., 1989, 1991).

The electric response signal of a light activated protein can be measured on anisotropic sample consisting of oriented membranes. These electric signals reveal details about the ion-transporting steps in the photocycle. Oriented samples can be obtained by incorporating the protein containing membrane fragments into a bilayer lipid membrane (Bamberg et al., 1981, 1984), or by oriented attachment of purple membranes to a lipid-impregnated filter (Drachev et al., 1984) or thin Teflon films (Holz et al., 1988). Another possibility is to apply an external electric field to the membrane suspension (Keszthelyi and Ormos, 1983; Keszthelyi and Ormos, 1989). This orientation can be fixed if the membranes are polymerized in an acrylamide gel, when the external electric field is applied (Dér et al., 1985a,b). The measured photoelectric response signals show a strong correlation with the photocycle (Keszthelyi and Ormos, 1989). The electrogenicity of an intermediate (by definition the change of the electric dipole moment of the protein in that intermediate relative to the unexcited state) characterized the magnitude of the charge shifts inside the protein (Trissl, 1990). The sign of the electrogenicity was considered positive, when the change of the electric dipole corresponds to the shift of a positive charge toward the extracellular side of the membrane. The electrogenicity of the intermediates in the BR photocycle and in the chloride photocycle of pHR were determined (Ludmann et al., 1998a, 2000). These help understand the process of charge transfer through the membrane.

In the present work the photocycle of pHR was investigated in the presence of azide and absence of chloride. The replacement of chloride ion with azide makes possible the deprotonation of the Schiff base, resulting in an M-like intermediate. Measurements on cell envelope vesicles proved that proton is transported through the membrane (Váró et al., 1996). Time-resolved spectroscopic, absorption kinetic and electric signal measurements could reveal details about the changes of the photocycle. Using time-resolved spectroscopy the spectra of photocycle intermediates can be determined (Nagle et al., 1995; Gergely et al., 1997; Zimányi et al., 1999). Absorption kinetic measurements at several wavelengths help to determine the photocycle model and calculate the microscopic rate constants (Váró et al., 1995b; Ludmann et al., 1998b). By electric signal measurements the charge motions inside the protein can be studied, which give information about the motion of the transported ion through the membrane (Ludmann et al., 1998a, 2000; Dér et al., 1999). Applying all these techniques, the similarities and differences between the azide photocycle of pHR and that of BR were investigated.

MATERIALS AND METHODS

Halorhodopsin-containing membrane suspension was prepared from *Halobacterium salinarum* strain L33, in which the *Natronobacterium pharaonis hop* structural gene and the novobiocin resistance gene for selection were introduced. This resulted in a greatly enhanced production of pHR. The process of preparation of the suspension was described earlier (Váró et al., 1995a). All spectroscopic and absorption kinetic measurements were performed on membranes encased in polyacrylamide gels, as described before (Váró et al., 1995c) on a sample of optical density about 1 at 570 nm. Electric signal measurements were carried out on oriented gel samples, prepared according to the procedure described earlier (Dér et al., 1985b). Before measurements the samples were exhaustively washed in a solution of 1 M Na_2SO_4, 50 mM (2-[*N*-morpholino]ethanesulfonic acid) MES, 100 mM NaN_3, pH 6. In the case of electric signal measurement, the concentration of Na_2SO_4 and MES was reduced to half to lower the

conductivity of the sample. This change in the salt concentrations did not introduce any observable alteration in the kinetics.

Laser excitation was performed with a frequency-doubled Nd-YAG laser (Continuum, Surelite I-10,λ = 532 nm). During the measurement the sample was kept in a temperature-controlled sample holder. Time-resolved spectroscopy with a gated optical multichannel analyzer provided difference spectra at various time points of the photocycle (Zimányi et al., 1989). The spectra of intermediates were calculated from these difference spectra, after noise reduction with singular value decomposition (SVD) (Golub and Kahan, 1992; Váró and Lanyi, 1991; Gergely et al., 1997). Absorption kinetic signals were recorded at five wavelengths (410, 500, 570, 610, and 650 nm) and five temperatures between 10 and 30°C, in a time interval from 100 ns to 1 s, using a transient recorder card with 16 MB memory (National Instruments, NI-DAQ PCI-5102), controlled by a program developed in our Institute. Each measurement involved averaging of 100–200 signals. At the end of the measurement, the linear time base was converted to a logarithmic one by averaging in the time interval between logarithmically equidistant points, which improved the signal to noise ratio. Model fitting was performed with the RATE program, as described elsewhere (Ludmann et al., 1998b). From the temperature dependence of the rate constants the energetic picture of the reactions were calculated, using the EYRING program (Ludmann et al., 1998b). Electric signals were measured on the earlier described set-up (Gergely et al., 1993), with the modification that a very low-noise home-made amplifier was used. The time resolution of the system was about 100 ns.

RESULTS

The replacement of the 1 M NaCl with 100 mM NaN_3 and 1 M Na_2SO_4 produces dramatic changes in the absorption kinetic signals (Fig. 1, *A* and *B*). The most important change is the absorption increase at 410 nm, characteristic of the intermediate M containing deprotonated Schiff base. In the presence of chloride ion the dominant absorption maximum was at 500 nm, which characterizes the L intermediate that persists over the entire photocycle (Fig. 1 *A*). In azide the absorption at 500 nm has a maximum at around 10 μs and later decreases to negative values (Fig. 1 *B*). This implies that in the photocycle with azide the intermediate L disappears in the millisecond range. The absorption kinetic signals measured at five wavelengths are very similar to that measured on BR (Fig. 1 *C*). The resemblance of the kinetic traces measured in pHR with azide and BR suggests that the photocycles are similar. To analyze the similarities and differences a thorough study was carried out, as described below.

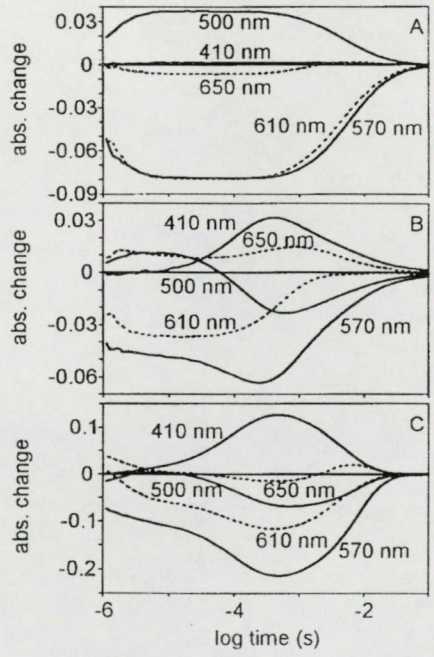


FIGURE 1 Absorption kinetic signals measured at five wavelengths on pharaonis halorhodopsin (*A*) and bacteriorhodopsin (*C*) in 1 M NaCl, 50 mM MES, pH 6, 20°C, and on pharaonis halorhodopsin (*B*) in 100 mM NaN_3, 1 M Na_2SO_4, 50 mM MES, pH 6, 20°C.

The difference spectra measured with an optical multichannel analyzer were submitted to SVD analysis. The autocorrelation function of the basis spectra and their time dependence showed a long-range structure in the first four components. The other components contained only noise. The weight factors of the four components were 83.6, 36.2, 16.4, and 2.2. All the following weight factors were much smaller. The analysis therefore suggests the existence of at least four independent spectral components. Noise-filtered difference spectra were reconstructed from these four SVD components (Fig. 2) and used in the calculation of the spectra of the intermediates. The difference spectra taken at the early time points show the existence of a red-shifted component around 650 nm, presumably the K intermediate. Later a peak around 500 nm appears, which can be attributed to intermediate L. Its decay is accompanied by the appearance of a strongly blue-shifted peak at around 410 nm, characteristic of intermediate M. The red-shifted peak does not disappear completely, but when the absorption at 410 nm approaches its maximum, a downshift of the red part of the difference spectrum can be observed, accompanied by an absorption increase, which suggests the appearance of another intermediate. Finally all the absorption changes disappear, showing that the pHR returns to its unexcited state.

Because the intermediates strongly overlap both spectrally and in time, finding a unique set of spectra is seriously hampered (Gergely et al., 1997). A control in the process of calculation was a reasonable model fit with RATE to the five absorption kinetic signals, when the extinction coefficients of the intermediates were taken from the calculated spectra. Based on the kinetic similarities between the studied photocycle and that belonging to BR (Fig. 1, *B* and *C*), the search was started with the same number of intermediates and with a model containing reversible reactions. When five intermediate spectra were calculated (K, L, M, N, and O; data not shown) the spectrum of N was very broad and its peak coincided with that of HR. No matter how the starting parameters of the fit to the absorption kinetic signals were taken, the fitting procedure resulted in no N

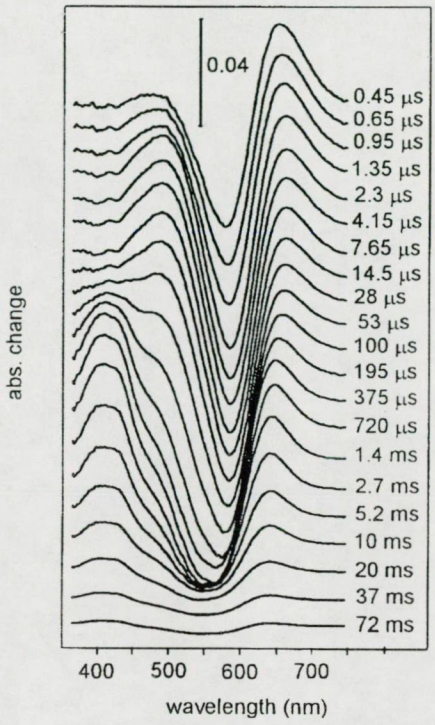


FIGURE 2 SVD-filtered, time-resolved difference spectra measured on pharaonis halorhodopsin at the indicated times after the photoexcitation. Measuring conditions: 1M Na_2SO_4, 100 mM NaN_3, 50 mM MES, pH 6 and 20°C.

intermediate in the photocycle. This led us to the conclusion that only the spectra of K, L, M, and O intermediates are present in the set of difference spectra. Using this information the spectra of the intermediates were calculated (Fig. 3).

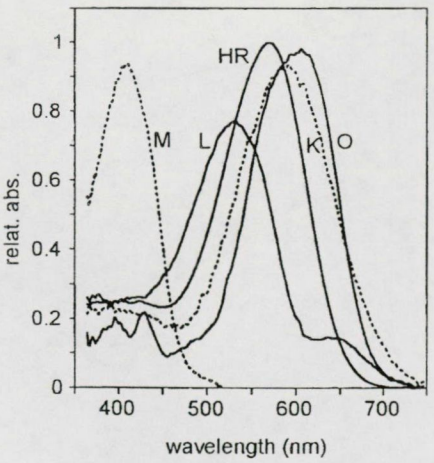


FIGURE 3 The spectra of intermediates calculated from the difference spectra shown in Fig. 2. The small secondary peak in the spectrum of L can be attributed to the 5–10% sulfate photocycle present after the photoexcitation.

Based on the resemblance of the azide photocycle with that of BR, but with a multiphasic decaying part (similar to pHR) and no intermediate N, the initially considered model was:

$$K \Leftrightarrow L \Leftrightarrow M \Leftrightarrow O \Leftrightarrow HR' \Rightarrow HR$$

Although the fit to the absorption kinetic signals (Fig. 4 *A*) had a relatively large error (±10%), it gave a general scheme about the time course of the concentrations of the intermediates (Fig. 4 *B*). The error could be attributed to the above-mentioned presence of the sulfate photocycle. In the starting model all the transitions were reversible, except the last one, ensuring that the photocycle terminates in a stable product. The fit resulted an M to L back-transition infinitely slow, leading to a model with a unidirectional step between these two intermediates:

$$K \Leftrightarrow L \Rightarrow M \Leftrightarrow O \Leftrightarrow HR' \Rightarrow HR$$

This results in the disappearance of intermediates K and L in the second part of the photocycle. Large amount of O appears, in strong equilibrium with intermediate M, which does not accumulate as much as in BR. The HR′ intermediate had to be considered, based on the very long, multiexponential decay of the absorption kinetic signals, and introduced earlier for a similar reason in the chloride photocycle

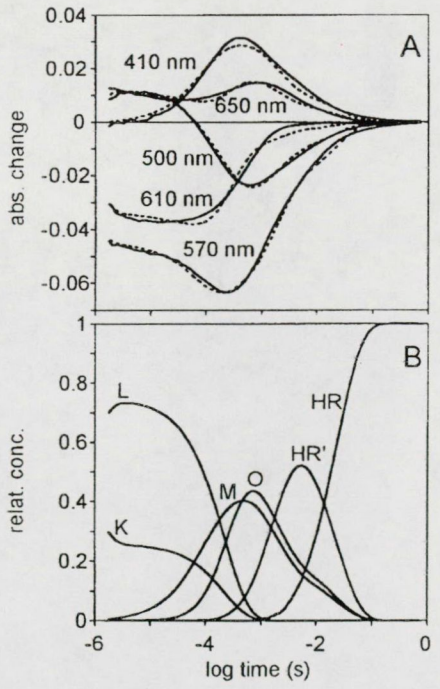


FIGURE 4 (*A*) the fit of the photocycle model (*broken lines*) to the absorption kinetic signals (*continuous lines*). (*B*) time evolution of the concentration of the intermediates. Measuring conditions were the same as in Fig. 2. The error of the fit was ±10%.

of pHR (Váró et al., 1995a,b). To determine the energetic scheme of the photocycle, the temperature dependence of the absorption kinetic signals were measured between 10°C and 30°C (Fig. 5). Below 10°C it was impossible to make this measurement: the Na_2SO_4 precipitated, and the sample became opaque and scattered the measuring light. Despite the large error in the model fit, the Eyring plot of the rate constants showed a reasonable linear dependence (Fig. 6), indicating that the model is correct.

The fit of the data with two parallel photocycles was also investigated. The temperature dependence of the rate constants gave no linear Eyring plot, eliminating this model from further consideration, similar to the case described earlier for BR (Ludmann et al., 1998b).

The free energy, enthalpy, and entropy of the photocycle transitions were calculated as previously (Váró et al., 1995a; Ludmann et al., 1998b) using the Eyring formula:

$$\ln k = -\frac{\Delta H^*}{R \cdot T} + \frac{\Delta S^*}{R} + \ln\frac{k_B \cdot T}{h}$$

$$\Delta G^* = \Delta H^* - T \cdot \Delta S^*$$

where k is the rate constant, T is the temperature expressed in °K, R is the gas constant, k_B is the Boltzmann constant, and h is the Planck's constant. From the calculated transition energies the energetic scheme of the whole photocycle was drawn (Fig. 7). All the energy levels are relative to K (or M in the second part). The exact position of the energy level of K (or M), relative to unexcited HR is not known, because the quantum efficiency and the energy stored in K were not determined. As the reverse transition M to L does not exist (it is unmeasurably slow, compared to the others), all that is known is that the free energy level of M is much lower than that of L, but no information exists about its position in enthalpy and entropy diagrams. Although M is placed at the same level as K, the dashed line separates the two parts, indicating that they are not related. The enthalpy and entropy diagrams show great similarity to that of chloride transporting pHR (Váró et al., 1995b).

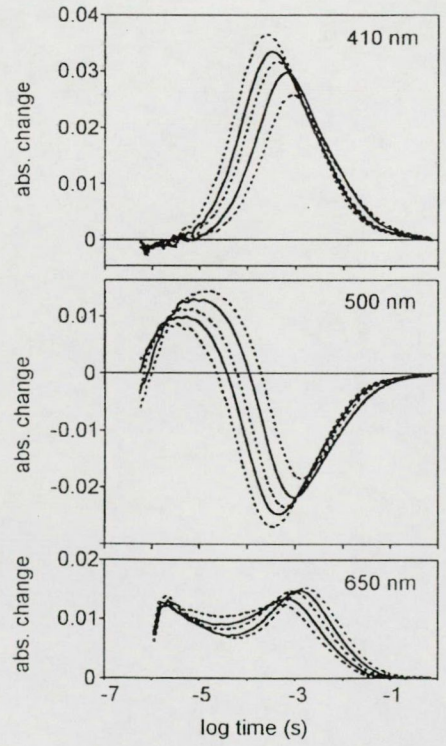


FIGURE 5 Absorption kinetic signals measured on pharaonis halorhodopsin at different temperatures. From right to left the temperatures are 10, 15, 20, 25, and 30°C. The other measuring conditions were the same as in Fig. 2.

The light-induced electric response signal measurements provide information about the charge motions inside the membrane. The electric signals of pHR were much smaller than that of BR. The membranes containing pHR form almost closed vesicles leading to a low degree of orientation by external electric field applied in the course of the sample preparation. The samples were measured in three different bathing solutions and compared with each other and with that of BR (Fig. 8). The fast component of the current signal measured in the presence of chloride is positive (in the same direction as that of the net charge transport) (curve A), in sulfate is negative (curve B), whereas in azide it is negative and crosses the zero line (curve C), becoming very similar to that of BR (curve BR).

DISCUSSION

The changes in the absorption kinetic signals when the chloride is exchanged with azide (Fig. 1) and the resemblance of these traces to that of BR raises the following questions: What kind of processes inside the protein were changed? Which steps are common in the proton and chloride translocation and which are different? Which steps determine the specificity of the transported ion? To clarify these questions, a thorough investigation of the azide pumping photocycle of pHR was carried out and compared to the two related photocycles: the proton transporting photocycle of the BR and the chloride transporting photocycle of the pHR.

It was established before that the binding constant of azide to pHR is 10 mM (Váró et al., 1996). In the bathing solution of the sample 1 M Na_2SO_4 and 100 mM azide was present, leading to a non-homogeneous sample containing 90–95% azide-bound protein. Because the spectrum of the azide-free protein is shifted with about 13 nm toward the red, the laser excitation at 532 nm is less effective upon it than upon the azide-bound pHR. Maximum 10% of the excited protein goes through the so-called sulfate photocycle (Váró et al., 1995a). This photocycle consists of only one intermediate, with its absorption maximum around 670

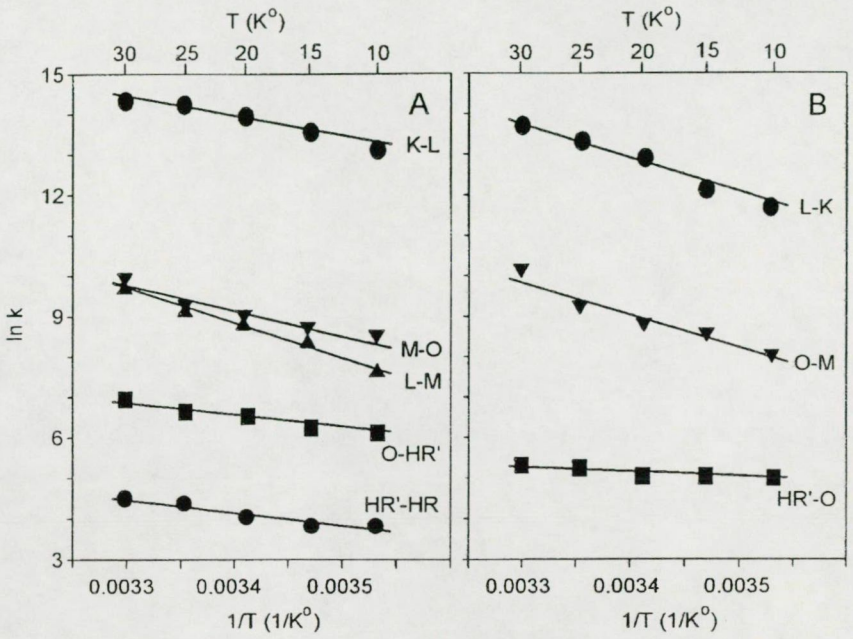


FIGURE 6 The Eyring diagram of the rate constants calculated from the model with reversible reactions. The lines are the fits from which the thermodynamic parameters were calculated. (*A*) Forward reactions of the model. (*B*) Back reactions. The error of the fit is ±5%.

nm, and decays in the 10 μs to millisecond time domain. The broadening in the red side of the spectrum of intermediate K and the small extra peak in L (Fig. 3) were attributed to this photocycle. Not having exact information about the quantum efficiencies of the two protein forms, which contribute to the signal, it would be very difficult to separate the two photocycles. The 10% sulfate photocycle is considered as an error in the fitting procedure, when the extinction coefficients are taken from the spectra of intermediates at

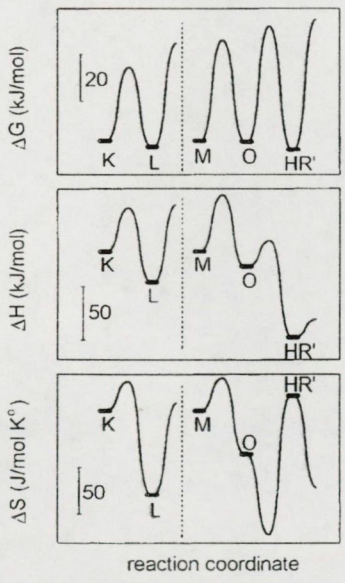


FIGURE 7 The free energy, enthalpy, and entropy diagram of the pharaonis halorhodopsin azide photocycle calculated form the fits shown in Fig. 6. The transition between intermediates L and M is unidirectional, which leaves undetermined the position of the energy level corresponding to M, represented by the separating broken lines between the two intermediates.

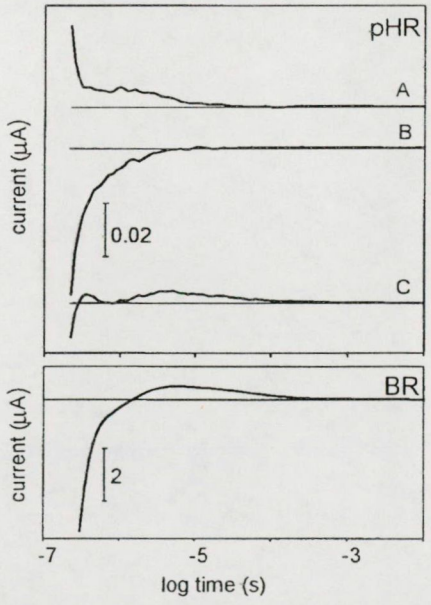


FIGURE 8 Light-induced electric signals of pharaonis halorhodopsin (*upper panel*) and bacteriorhodopsin (*lower panel*) measured on oriented gel samples. The measuring conditions are as follows: Line A and BR were measured in 1M NaCl, line B in 0.5 M Na_2SO_4, and line C in 0.5 M Na_2SO_4 and 100 mM NaN_3. All solutions had 25 mM MES, pH 6, 20°C.

the measured wavelengths. It is known that the sulfate photocycle is completed in several ms and affects only the red end of the intermediate spectra. Nevertheless, when the absorption kinetic traces are fitted with the RATE program, the errors are distributed uniformly on all wavelengths (Fig. 4 *A*).

The time-dependent concentration changes of the intermediates of the azide photocycle of pHR were compared (Fig. 4 *B*) with the chloride photocycle (Váró et al., 1995a) and that of BR photocycle (Ludmann et al., 1998b). The K to L transition is as fast as in chloride, but the equilibrium between the two intermediates is shifted to a lower K concentration and these intermediates disappear from the photocycle in the ms time region, similar to that in BR. In the microsecond range, but later than in BR, the M intermediate appears, indicating that the Schiff base deprotonates. The rise of M is not as clearly multiphasic as in BR. The model fit did not demand the introduction of two M forms, but this does not exclude the existence of them. In both BR and pHR, intermediate M reaches its maximum concentration in about the same time. The elimination of intermediate N from the photocycle model does not mean that it is completely excluded, only that it can not be observed. Our measurements can not decide, whether N is absent from this photocycle or only due to kinetic reasons does not accumulate. Instead of N a large amount of O is accumulated. The equilibrium constant of the M to O transition is almost 1, resulting in a equal amount of intermediates M and O, decaying through HR′ to HR.

All the transitions in the second part of the photocycle have weak temperature dependence, reflected in the absorption kinetic measurements over the millisecond time range (Fig. 5). The only exception is in the 650 nm measurement, characteristic for intermediate O, where a more accentuated change can be observed. With increasing temperature the amount of intermediate O decreases, contrary to the BR photocycle (Fig. 5, panel 650 nm). This can be attributed to the fact that the slope of the reverse transition O to M is larger than that corresponding to M to O and all the other transitions after intermediate O (Fig. 6). The free energy scheme, calculated from the Eyring plots, show two unidirectional reactions, between intermediates L and M, and the final decay of HR′. A reaction between two intermediates was considered unidirectional, when the forward reaction is at least 100 times faster than the corresponding back reaction. This corresponds to at least −11 kJ/mol free energy difference between the two states. The BR photocycle measured over pH 6 had also two unidirectional reactions (Ludmann et al., 1998b). The free energy scheme of both chloride transporting photocycle of pHR and the proton-transporting photocycle of BR are similar. Great differences exist in the enthalpy and entropy changes, however. While the reactions in the photocycle of BR are entropy-driven, because the entropic energy ($-T\Delta S$) decreases in almost every reaction (Ludmann et al., 1998b), the transitions in the chloride photocycle of pHR are enthalpy-driven (Váró et al., 1995a). The transitions in the azide photocycle of pHR are also enthalpy-driven. Except the reaction between intermediates L and M, where there is no information about the relation between the energy levels, all the other enthalpy levels look very similar to the chloride photocycle. The change in the function of the protein did not change the way the energy is stored and used during the photocycle. From the thermodynamic study can be concluded that in BR the conformational changes of the protein drive the photocycle, resulting in decreasing entropy. In pHR the driving force results mainly from changes in the electrostatic interactions inside the protein.

The electric signals are characteristic to the charge translocation steps in BR and pHR (Ludmann et al., 1998a, 2000). The fast component of the current signal in BR and pHR in different conditions were compared (Fig. 8). In BR the electrogenicity of intermediate K was negative, the L also negative and about two times larger, resulting in a fast negative signal, which crosses the zero line, becoming positive in several microseconds (Fig. 8, BR). The positive current corresponds to the rise of the M intermediate (Ludmann et al., 1998a). In the chloride photocycle of pHR the electrogenicity of K was negative, that corresponding to L was positive and about three times larger, resulting in a positive signal (Fig. 8 *A*) (Kalaidzidis et al., 1998; Ludmann et al., 2000). With pHR in sulfate, where a K-like intermediate dominates the photocycle (Váró et al., 1995a), the fast component was negative and it decayed to zero, without any positive component (Fig. 8 *B*) (Ludmann et al. 2000). When azide was added, the fast component of the current signal was also negative and later crossed the zero line (Fig. 8 *C*). As the electric signal was very small, and the fast component was close to the time resolution of the systems (100 ns), it cannot be excluded that the first positive peak is an overshoot, after the negative component. The second positive peak, in the microsecond time domain, corresponds to the rise of the intermediate M. The appearance of the fast negative component and the positive component corresponding to intermediate M makes the photoelectric response signal of the azide photocycle very similar to that of BR (Fig. 8).

The combination of electrogenicities corresponding to intermediates K and L yields the fast part of the electric response signal. Depending on their sign and amplitude, the result can be either positive or negative. The K intermediate, which corresponds to the retinal isomerization in all photocycles, has negative electrogenicity (Ludmann et al., 1998a; Ludmann et al. 2000). This process seems to be similar in both proteins at all conditions. The electric signal of the L intermediate corresponds to a local relaxation of the protein around the retinal chromophore. This relaxation is sensitive to the presence of the proton acceptor. It is positive in the absence of the proton acceptor and decreases or becomes

negative when the acceptor is present. All the other transporting steps are positive (Ludmann et al., 1998a, 2000).

Comparing the proton-transporting photocycle of pHR in the presence of azide to that of BR, it is shown here that kinetically they are very similar, with the very characteristic intermediate M containing unprotonated Schiff base. The process of charge translocation, manifested in the electric signal, is also very similar in the azide photocycle of pHR and BR. On the other hand, energetically the azide photocycle remains analogous to that of the chloride transporting one. It is driven by enthalpy decrease during the transitions between the intermediates. The transformation of the chloride-transporting photocycle to the proton-transporting one by simply replacing chloride with azide, which takes the role of both the proton donor and acceptor in the protein and the great similarity of it to that of BR, proves that the two very different ions are transported via the same basic steps.

We are grateful to Prof. L. Keszthelyi and Dr. L. Zimányi for helpful discussions. This work was supported by grants from the National Science Research Fund of Hungary (OTKA T022066, T020470) and the Research Fund of the Hungarian Academy of Sciences (AKP 97-71 3, 3/52).

REFERENCES

Balashov, S. P., R. Govindjee, and T. G. Ebrey. 1991. Red shift of the purple membrane absorption band and the deprotonation of tyrosine residues at high pH: origin of the parallel photocycles of transbacteriorhodopsin. *Biophys. J.* 60:475–490.

Balashov, S. P., R. Govindjee, E. S. Imasheva, S. Misra, T. G. Ebrey, Y. Feng, R. K. Crouch, and D. R. Menick. 1995. The two pK_a's of aspartate-85 and control of thermal isomerization and proton release in the arginine-82 to lysine mutant of bacteriorhodopsin. *Biochemistry.* 34:8820–8834.

Bamberg, E., N. A. Dencher, A. Fahr, and M. P. Heyn. 1981. Transmembranous incorporation of photoelectrically active bacteriorhodopsin in planar lipid bilayers. *Proc. Natl. Acad. Sci. USA.* 78:7502–7506.

Bamberg, E., P. Hegemann, and D. Oesterhelt. 1984. Reconstitution of the light-driven electrogenic ion pump halorhodopsin into black lipid membranes. *Biochim. Biophys. Acta.* 773:53–60.

Bamberg, E., J. Tittor, and D. Oesterhelt. 1993. Light-driven proton or chloride pumping by halorhodopsin. *Proc. Natl. Acad. Sci. USA.* 90: 639–643.

Bivin, D. B., and W. Stoeckenius. 1986. Photoactive retinal pigments in haloalkalophilic bacteria. *J. Gen. Microbiol.* 132:2167–2177.

Brown, L. S., A. K. Dioumaev, R. Needleman, and J. K. Lanyi. 1998. Local-access model for proton transfer in bacteriorhodopsin. *Biochemistry.* 37:3982–3993.

Dér, A., K. Fendler, L. Keszthelyi, and E. Bamberg. 1985a. Primary charge separation in halorhodopsin. *FEBS Lett.* 187:233–236.

Dér, A., P. Hargittai, and J. Simon. 1985b. Time-resolved photoelectric and absorption signals from oriented purple membranes immobilized in gel. *J. Biochem. Biophys. Methods.* 10(5–6):295–300.

Dér, A., L. Oroszi, A. Kulcsár, L. Zimányi, R. Tóth-Boconádi, L. Keszthelyi, W. Stoeckenius, and P. Ormos. 1999. Interpretation of the spatial charge displacements in bacteriorhodopsin in terms of structural changes during the photocycle. *Proc. Natl. Acad. Sci. USA.* 96:2776–2781.

Dér, A., S. Száraz, R. Tóth-Boconádi, Z. Tokaji, L. Keszthelyi, and W. Stoeckenius. 1991. Alternative translocation of protons and halide ions by bacteriorhodopsin. *Proc. Natl. Acad. Sci. USA.* 88:4751–4755.

Dér, A., R. Tóth-Boconádi, and L. Keszthelyi. 1989. Bacteriorhodopsin as a possible chloride pump. *FEBS Lett.* 259:24–26.

Dickopf, S., U. Alexiev, M. P. Krebs, H. Otto, R. Mollaaghababa, H. G. Khorana, and M. P. Heyn. 1995. Proton transport by a bacteriorhodopsin mutant aspartic acid-85->asparagine, initiated in the unprotonated Schiff base state. *Proc. Natl. Acad. Sci. USA.* 92:11519–11523.

Drachev, L. A., A. D. Kaulen, and V. P. Skulachev. 1984. Correlation of photochemical cycle, H^+ release and uptake, and electrical events in bacteriorhodopsin. *FEBS Lett.* 178:331–335.

Duschl, A., J. K. Lanyi, and L. Zimányi. 1990. Properties and photochemistry of a halorhodopsin from the haloalkalophile, Natronobacterium pharaonis. *J. Biol. Chem.* 265:1261–1267.

Gergely, C., C. Ganea, G. I. Groma, and G. Váró. 1993. Study of the photocycle and charge motions of the bacteriorhodopsin mutant D96N. *Biophys. J.* 65:2478–2483.

Gergely, C., L. Zimányi, and G. Váró. 1997. Bacteriorhodopsin intermediate spectra determined over a wide pH range. *J. Phys. Chem. B.* 101:9390–9395.

Gerscher, S., M. Mylrajan, P. Hildebrandt, M. H. Baron, R. Müller, and M. Engelhard. 1997. Chromophore-anion interactions in halorhodopsin from *Natronobacterium pharaonis* probed by time-resolved resonance Raman spectroscopy. *Biochemistry.* 36:11012–11020.

Golub, G., and W. Kahan. 1992. Calculating the singular values and pseudo-inverse of a matrix. *SIAM J. Num. Anal.* 2:205–224.

Haupts, U., J. Tittor, E. Bamberg, and D. Oesterhelt. 1997. General concept for ion translocation by halobacterial retinal proteins: the isomerization/switch/transfer (IST) model. *Biochemistry.* 36:2–7.

Havelka, W. A., R. Henderson, and D. Oesterhelt. 1995. Three-dimensional structure of halorhodopsin at 7 Å resolution. *J. Mol. Biol.* 247:726–738.

Hegemann, P., D. Oesterhelt, and M. Steiner. 1985. The photocycle of the chloride pump halorhodopsin. I. Azide catalyzed deprotonation of the chromophore is a side reaction of photocycle intermediates inactivating the pump. *EMBO J.* 4:2347–2350.

Holz, M., M. Lindau, and M. P. Heyn. 1988. Distributed kinetics of the charge movements in bacteriorhodopsin: evidence for conformational substates. *Biophys. J.* 53:623–633.

Kalaidzidis, IV, Y. L. Kalaidzidis, and A. D. Kaulen. 1998. Flash-induced voltage changes in halorhodopsin from *Natronobacterium pharaonis*. *FEBS Lett.* 427:59–63.

Keszthelyi, L., and P. Ormos. 1983. Displacement current on purple membrane fragments oriented in a suspension. *Biophys. Chem.* 18: 397–405.

Keszthelyi, L., and P. Ormos. 1989. Protein electric response signals from dielectrically polarized systems. *J. Membr. Biol.* 109:193–200.

Lanyi, J. K., A. Duschl, G. W. Hatfield, K. M. May, and D. Oesterhelt. 1990. The primary structure of a halorhodopsin from Natronobacterium pharaonis: structural, functional and evolutionary implications for bacterial rhodopsins and halorhodopsins. *J. Biol. Chem.* 265:1253–1260.

Ludmann, K., C. Gergely, A. Dér, and G. Váró. 1998a. Electric signals during the bacteriorhodopsin photocycle, determined over a wide pH range. *Biophys. J.* 75:3120–3126.

Ludmann, K., C. Gergely, and G. Váró. 1998b. Kinetic and thermodynamic study of the bacteriorhodopsin photocycle over a wide pH range. *Biophys. J.* 75:3110–3119.

Ludmann, K., G. Ibron, J. K. Lanyi, and G. Váró. 2000. Charge motions during the photocycle of pharaonis halorhodopsin. *Biophys. J.* 78: 959–966.

Nagle, J. F., L. Zimányi, and J. K. Lanyi. 1995. Testing BR photocycle kinetics. *Biophys. J.* 68:1490–1499.

Oesterhelt, D., and J. Tittor. 1989. Two pumps, one principle: light-driven ion transport in halobacteria. *Trends Biochem. Sci.* 14:57–61.

Oesterhelt, D., J. Tittor, and E. Bamberg. 1992. A unifying concept for ion translocation by retinal proteins. *J. Bioenerg. Biomembr.* 24:181–191.

Ormos, P., C. Gergely, S. Kruska, S. Száraz, and Z. Tokaji. 1997. The effect of azide on the photocycle of bacteriorhodopsin. *J. Photochem. Photobiol. B Biol.* 40:111–119.

Sasaki, J., L. S. Brown, Y.-S. Chon, H. Kandori, A. Maeda, R. Needleman, and J. K. Lanyi. 1995. Conversion of bacteriorhodopsin into a chloride ion pump. *Science.* 269:73–75.

Scharf, B., and M. Engelhard. 1994. Blue halorhodopsin from *Natronobacterium pharaonis*: wavelength regulation by anions. *Biochemistry.* 33: 6387–6393.

Tittor, J., C. Soell, D. Oesterhelt, H.-J. Butt, and E. Bamberg. 1989. A defective proton pump, point-mutated bacteriorhodopsin Asp96>Asn is fully reactivated by azide. *EMBO J.* 8:3477–3482.

Tittor, J., M. Wahl, U. Schweiger, and D. Oesterhelt. 1994. Specific acceleration of de- and reprotonation steps by azide in mutated bacteriorhodopsins. *Biochim. Biophys. Acta.* 1187:191–197.

Trissl, H. W. 1990. Photoelectric measurements of purple membranes. *Photochem. Photobiol.* 51:793–818.

Váró, G. 2000. Analogies between halorhodopsin and bacteriorhodopsin. *Biochim. Biophys. Acta.* 1460:220–229.

Váró, G., L. S. Brown, R. Needleman, and J. K. Lanyi. 1996. Proton transport by halorhodopsin. *Biochemistry.* 35:6604–6611.

Váró, G., L. S. Brown, N. Sasaki, H. Kandori, A. Maeda, R. Needleman, and J. K. Lanyi. 1995a. Light-driven chloride ion transport by halorhodopsin from *Natronobacterium pharaonis*. I. The photochemical Cycle. *Biochemistry.* 34:14490–14499.

Váró, G., and J. K. Lanyi. 1991. Kinetic and spectroscopic evidence for an irreversible step between deprotonation and reprotonation of the Schiff base in the bacteriorhodopsin photocycle. *Biochemistry.* 30:5008–5015.

Váró, G., R. Needleman, and J. K. Lanyi. 1995b. Light-driven chloride ion transport by Halorhodopsin from *Natronobacterium pharaonis* 2. Chloride release and uptake, protein conformation change, and thermodynamics. *Biochemistry.* 34:14500–14507.

Váró, G., L. Zimányi, X. Fan, L. Sun, R. Needleman, and J. K. Lanyi. 1995c. Photocycle of halorhodopsin from *Halobacterium salinarium*. *Biophys. J.* 68:2062–2072.

Walter, T. J., and M. S. Braiman. 1994. Anion-protein interaction during halorhodopsin pumping: halide binding at the protonated Schiff base. *Biochemistry.* 33:1724–1733.

Zimányi, L., L. Keszthelyi, and J. K. Lanyi. 1989. Transient spectroscopy of bacterial rhodopsins with optical multichannel analyser. 1. Comparison of the photocycles of bacteriorhodopsin and halorhodopsin. *Biochemistry.* 28:5165–5172.

Zimányi, L., A. Kulcsár, J. K. Lanyi, D. F. Sears, Jr., and J. Saltiel. 1999. Singular value decomposition with self-modeling applied to determine bacteriorhodopsin intermediate spectra: analysis of simulated data. *Proc. Natl. Acad. Sci. USA.* 96:4408–4413.

Zimányi, L., and J. K. Lanyi. 1993. Deriving the intermediate spectra and photocycle kinetics from time-resolved difference spectra of bacteriorhodopsin: the simpler case of the recombinant D96N protein. *Biophys. J.* 64:240–251.

Zimányi, L., and J. K. Lanyi. 1997. Fourier transform Raman study of retinal isomeric composition and equilibration in halorhodopsin. *J. Phys. Chem. B.* 101:1930–1933.

Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 2776–2781, March 1999
Biophysics

Interpretation of the spatial charge displacements in bacteriorhodopsin in terms of structural changes during the photocycle

ANDRÁS DÉR*†‡, LÁSZLÓ OROSZI*, ÁGNES KULCSÁR*, LÁSZLÓ ZIMÁNYI*, RUDOLF TÓTH-BOCONÁDI*†, LAJOS KESZTHELYI*, WALTHER STOECKENIUS§, AND PÁL ORMOS*

*Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, P.O. Box 521, Szeged H-6701, Hungary; †Grenoble High Magnetic Field Laboratory, 25 Avenue des Martyrs, BP 166, 38042 Grenoble Cedex 9, France; and §Department of Chemistry, University of California Santa Cruz, Santa Cruz, CA, 95064

Contributed by Walther Stoeckenius, January 4, 1999

ABSTRACT **We have recently introduced a method, made possible by an improved orienting technique using a combination of electric and magnetic fields, that allows the three-dimensional detection of the intramolecular charge displacements during the photocycle of bacteriorhodopsin. This method generates electric asymmetry, a prerequisite for the detection of electric signal on the macroscopic sample, in all three spatial dimensions. Purple membrane fragments containing bacteriorhodopsin were oriented so that their permanent electric dipole moment vectors were perpendicular to the membrane plane and pointed in the same direction. The resulting cylindrical symmetry was broken by photoselection, *i. e.*, by flash excitation with low intensity linearly polarized light. From the measured electric signals, the three-dimensional motion of the electric charge center in the bacteriorhodopsin molecules was calculated for the first 400 μs. Simultaneous absorption kinetic recording provided the time-dependent concentrations of the intermediates. Combining the two sets of data, we determined the discrete dipole moments of intermediates up to M. When compared with the results of current molecular dynamics calculations, the data provided a decisive experimental test for selecting the optimal theoretical model for the proton transport and should eventually lead to a full description of the mechanism of the bacteriorhodopsin proton pump.**

Bacteriorhodopsin (bR) is an integral protein in the plasma membrane of *Halobacterium salinarum* (1). Upon light absorption, bR transports protons across the membrane, converting the photon energy into the energy of a proton electrochemical gradient (for recent reviews, see refs. 2 and 3). bR is a single small protein and is the simplest known active ion pump and biological light energy transducer. Consequently, it is a prototype system for studying the basic steps and rules of biological energy transduction. bR consists of the 23-kDa protein moiety and a retinal chromophore attached to the side chain of Lys-216 via a protonated Schiff base. The retinal is located about halfway through the membrane and sustains an angle of ≈70° with the membrane normal. Trimers of bR form two-dimensional crystalline patches in the plasma membrane. These purple membrane fragments (pm) can be isolated as sheets of about 0.5-μm diameter. The structure of the bR molecule is known to ≈3-Å resolution from electron crystallography (4, 5), and thanks to recent successes in preparing three-dimensional crystals of bR, to 2.3-Å resolution from x-ray crystallography (6, 7). The protein is the paradigm of the larger family of seven-helix receptors found throughout the living world from Archaea to mammals.

Light absorption by retinal initiates a chain of reactions called the photocycle. This photocycle is usually represented as a sequence of spectroscopically characterized intermediates (8): bR—K—L—M—N—O—bR.

In the excited state the retinal undergoes a cis/trans isomerization around the C=C(13,14) double bond, which is completed in the bR–K transition. The primary form of energy storage is presumably charge separation and/or mechanical strain caused by conformational changes. This energy is converted in several thermally activated steps through the subsequent metastable intermediate states, and a proton is transported while the system returns to the initial state. Understanding the mechanism will require a detailed description of the molecular events occurring during the cycle.

Photoelectric methods are widely used to follow the charge motions in ion pumps. Light-induced currents carry information about the transported charge and the protein conformational changes accompanying the function. Various experimental techniques have been developed to measure the charge movement with high time resolution. One major family of techniques involves incorporation of pm in an oriented fashion into liposomes (9) or attaching them to planar bimolecular lipid membranes (10, 11) that separate two aqueous compartments. A different approach uses pm suspension, where the permanent electric dipole moment, parallel to the membrane normal, orients the pm in a dc electric field of moderate strength (12). If a suspension of oriented pm is excited by a laser flash, the intramolecular charge displacement can be measured as a displacement current by electrodes immersed in the electrolyte (13). When the orientation is "frozen" in a gel, the system remains stable and, in contrast to suspensions, it enables long, repetitive data acquisition in the absence of an external orienting electric field (14). Because the sample is well defined and a quantitative analysis of the photoelectric signal is possible, one can determine the charge displacements during each step of the photocycle (15). This technique has the important advantage that spectroscopic changes can be recorded on the same sample.

Fundamental for all photoelectric methods is asymmetry of the studied object. Bacteriorhodopsin must be oriented in both the membrane and suspension methods to measure a net current flow macroscopically. However, in all model systems, bR is oriented only in the direction of the membrane normal. Around this axis the bR molecules are arranged rotationally randomly. Because of this cylindrical symmetry, only one spatial component of the charge displacement parallel to the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: bR, bacteriorhodopsin; pm, purple membrane fragments.

‡To whom reprint requests should be addressed. e-mail: derandra@everx.szbk.u-szeged.hu.

membrane normal can be followed: all other components cancel. For the full description of the charge displacement it is, however, necessary to determine all three spatial components of the photocurrent, and this requires nearly perfect orientation, which can be achieved by the combination of electric and magnetic fields (16). We have recently shown that it is feasible to measure the lateral components of the photocurrent (17) if the cylindrical symmetry of the system is broken by photoselection. The sample of oriented pm is excited by low-intensity light of appropriate linear polarization so that bR molecules oriented in one direction, are preferentially excited, and the in-plane component of the photocurrent, which contains information about the lateral charge motion, can be recorded.

Here we describe a further elaboration of the method, which determines the three spatial components of the charge displacement quantitatively and correlates the charge displacements to absorption kinetics of the photocycle. The results are discussed and related to predictions from molecular dynamics calculations (18–20).

MATERIALS AND METHODS

Sample Preparation. The experiments described here require a suspension of pm with nearly perfect orientation, which cannot be achieved with a dc electric field alone, because in a high electric field two effects counteract orientation: (*i*) perturbation by an induced dipole moment, perpendicular to the permanent dipole moment and (*ii*) migration and a resulting shear of the membrane fragments because of the net electric charge of pm (12, 21). A high magnetic field will align the membranes perfectly, but because the alignment is caused by an induced magnetic dipole, it cannot orient them facing in the same direction. Therefore, orientation by the combination of electric and magnetic fields was introduced (16). Here the initial orientation, achieved by short exposure to a moderate dc electric field, is improved by the subsequent application of a magnetic field while the oriented membranes are being immobilized in the polymerizing gel. The procedure aligns the membrane sheets with their permanent dipoles parallel to the field, pointing in the same direction, and is not impaired by the disturbing effects present during the application of a strong electric field.

Pm fragments prepared by the standard procedure (1) were oriented and immobilized in gel as follows (16). Suspensions containing 50 μM pm, 15% acrilamide, 0.8% *N,N'*-methylenebisacrilamide, 1% tetramethylethylenediamine (TEMED), and 0.3% ammonium persulfate were filled into cuvettes of 20 × 15 × 10 mm fitted with platinized Pt electrodes. An electric field of 20 V/cm was first applied to orient the pm, and the sample was then positioned in a 17.5-T magnetic field (Max Planck Institute High Magnetic Field Laboratory, Grenoble, France). Samples were also prepared with magnetic ordering only. In this case the pm are aligned rather than oriented, i.e., the membrane sheets are parallel, but their sidedness (overall pumping direction) is random. Alignment of the membranes in the suspension and gel was checked by magnetic birefringence measurements (22) and were found to be close to perfect. Further technical details of the sample preparation can be found in ref. 16.

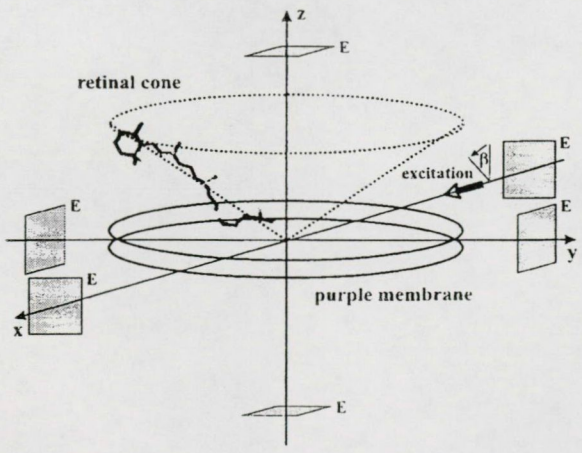


FIG. 1. Geometry of a pm oriented with its membrane normal coincident with the laboratory fixed *z* axis. The individual retinal chromophores' transition dipoles form a cone about this axis. The actinic laser pulse propagates along the laboratory fixed *x* axis with its polarization vector making an angle β with the *z* axis. The rectangular symbols represent the measuring electrodes (E).

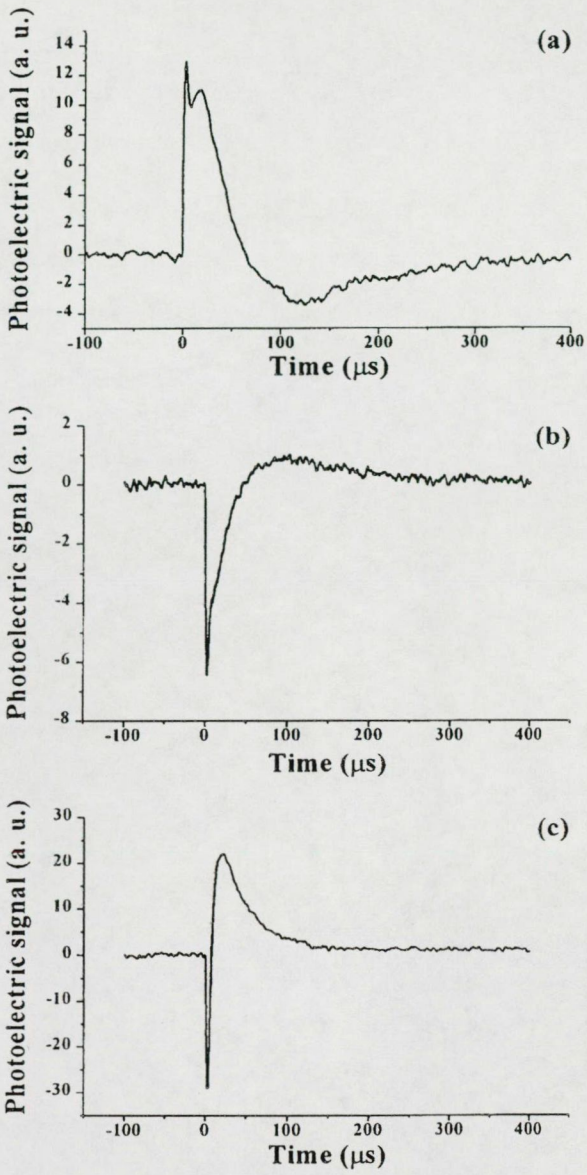


FIG. 2. Time-dependent photocurrents measured on pm oriented with the electric-magnetic field orientation method, in the three perpendicular laboratory fixed spatial directions. The signals were detected with electrode pairs oriented along the corresponding Cartesian axes. (*a*) $I_x(t)$. (*b*) $I_y(t)$. (*c*) $I_z(t)$.

Measuring Techniques. Cubes (1 × 1 × 1 cm) were cut out of gels with $OD_{575} = 2$ and were placed in a square measuring cell with three pairs of platinized Pt electrodes aligned along the *x*, *y*, and *z* axes (Fig. 1). Distance between the electrodes

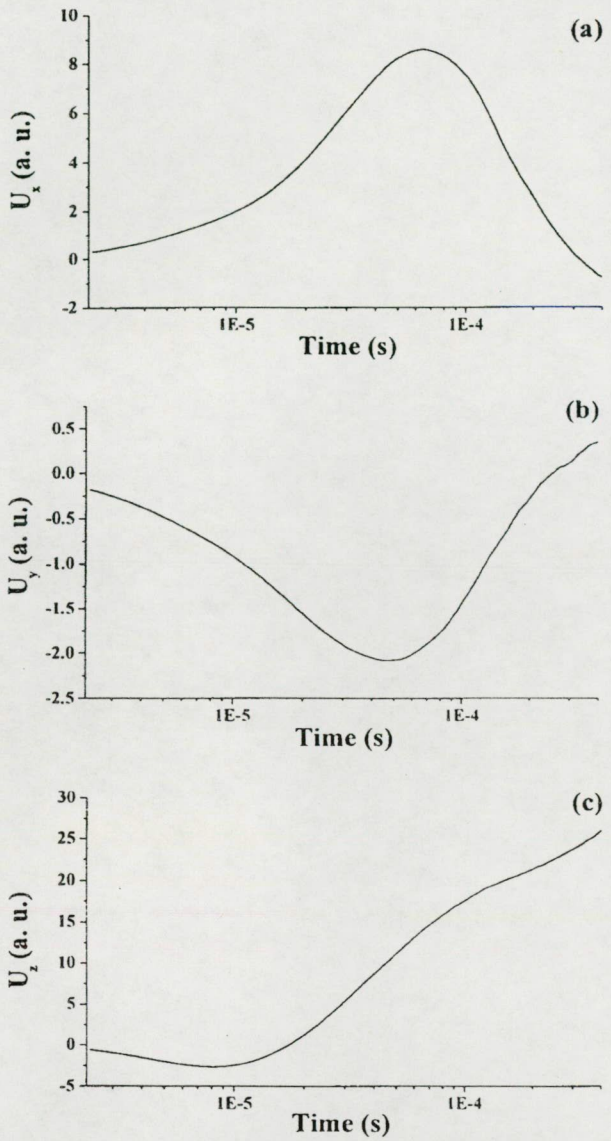


FIG. 3. Integral of the photoelectric signals in Fig. 2. $U_x(t)$ (*a*), $U_y(t)$ (*b*), and $U_z(t)$ (*c*) are linearly dependent on the net electric dipole moment of the photocycle intermediates (see text).

was 2 cm. Special care was taken to shield the electrodes from the exciting light to avoid artifacts caused by light scattering.

Fig. 1 shows the geometry of optical properties of a purple membrane oriented with its permanent dipole moment along the *z* direction. Samples were excited by polarized, 10-ns laser flashes of 100-μJ energy, from an excimer- laser-pumped dye laser (Rhodamine 6G), propagating in the *x* direction. The plane of polarization sustained angel $\beta = 45^{\circ}$ with the *z* axis. The signals were detected with the electrode pairs, amplified by voltage amplifiers, and stored in a transient recorder (LeCroy 9310L).

Absorption kinetics were measured on the same samples used in the electric measurements, under the same conditions (pH 6, 23°C). Spectral changes following the laser excitation were detected at logarithmically equidistant time delays from 300 ns to 400 μs, in the 350–750 nm wavelength range, with a gated Optical Multichannel Analyser (Spectroscopy Instruments, Gilching, Germany). Singular value decomposition provided the effective rank of the data matrix consisting of the time-resolved difference spectra, and the matrix was reconstructed and the noise filtered by using the significant singular value decomposition components only. The absorption spectra of the intermediates were derived from the difference spectra by using the Monte Carlo search algorithm described earlier (23). The concentration changes of the intermediates were calculated from the nonnegative linear least-squares fit of the pure intermediate spectra to the difference spectra. These time-resolved concentration changes were fitted by a linear photocycle scheme with reversible reactions to obtain the final intermediate concentrations, $c_i(t)$, by using the RATE 2.1 program written by Géza Groma. Further mathematical analysis of the data was carried out with the MATLAB program package.

Principles of the Method. Excitation with linearly polarized light allows measuring the in-plane components of the electric signal by breaking the intrinsic rotational symmetry of the sample around the *z* axis. We denote the three Cartesian component vectors of the directly detected electric signal by I_x, I_y and I_z, respectively.

Presuming perfect orientation and by using plausible assumptions (17, 24), the following matrix formula can be derived.

$$I_k(t) \propto \sum_j f_{jk} \sum_i \mu_i^j \frac{d}{dt} c_i(t) \qquad [1]$$

where j,k = *x*, *y*, or *z*, c_i is the concentration of the *i*th intermediate state, and the f_{jk} matrix element gives the weight of the μ_i^j dipole component relative to the *k* direction in the laboratory fixed-coordinate system. Note that the undefined proportionality constant in Eq. **1** depends on the ionic relaxation around the pm (15). Because of the membrane anisotropy and other reasons discussed in ref. 17, it must be different in the different Cartesian directions. By using the assignments of Fig. 1, f_{jk} can be written in a diagonal form:

$$[f_{jk}] = \begin{bmatrix} -\sin\beta\cos\beta & 0 & 0 \\ 0 & -\sin\beta\cos\beta & 0 \\ 0 & 0 & f_{zz}(\beta) \end{bmatrix} \qquad [2]$$

i.e., μ_i^j's are separated in this case in such a way that they contribute only to the corresponding $I_j(t)$'s. It is also easily seen from Eq. **2** that $I_x(t)$ and $I_y(t)$ have their maxima at $\beta = \pm 45^{\circ}$. The analytical form of $f_{zz}(\beta)$ has also been obtained. It (and, consequently I_z) has a minimum at $\beta = 0^{\circ}$ and a maximum at $\beta = 90^{\circ}$.

If the orientation is not perfect, $[f_{jk}]$ cannot be written in a diagonal form, that is, the measured $I_y(t)$ will be "contaminated" by μ_i^z's, as already discussed in ref. 16. By using symmetry arguments, it was also pointed out there that, in case of a magnetically aligned sample, only the μ_i^x components contribute to the measured signal.

To determine the corresponding μ_i^j values, the linear combination of $c_i(t)$'s was fitted to the time integral functions of $I_k(t)$'s ($U_k(t) = \int_0^t I_k(t)dt$).

RESULTS

Fig. 2 *a–c* show traces detected in different spatial dimensions, $I_x(t)$, $I_y(t)$, and $I_z(t)$, representing the Cartesian components of the photoelectric current measured on a sample containing pm oriented with the combination of electric and magnetic fields. As we pointed out earlier (ref. 24; *Materials and Methods*), the time-integral function of the photocurrent is proportional to the potential difference generated by the charge displacements during the pumping process, and it is properly scaled in terms of the dipole moment changes accompanying the photocycle of bR. Therefore, the integrated curves, $U_x(t)$, $U_y(t)$ and $U_z(t)$ were considered in the subsequent analysis (Fig. 3). Fig. 4

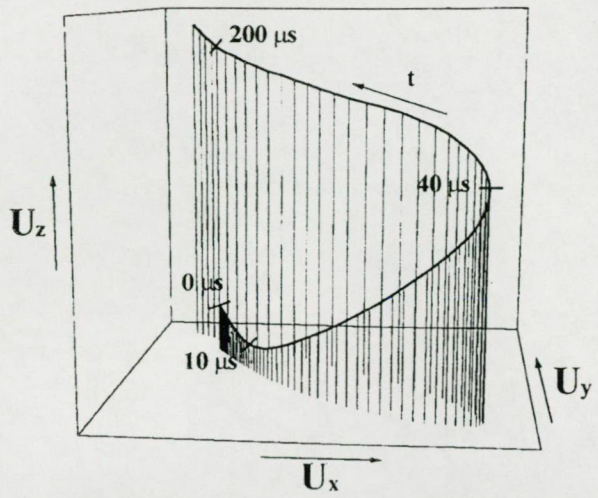


FIG. 4. Three-dimensional representation of the resultant integrated photocurrent, D(t), whose Cartesian components, $U_x(t)$, $U_y(t)$, and $U_z(t)$ are depicted in Fig. 3.

shows a three-dimensional trajectory defined by the resulting dipole moment vector $D(t) = [U_x(t),U_y(t),U_z(t)]$ in the first 400 μs of the photocycle.

From the definition of the electric dipole moment, it follows that this trajectory also characterizes the motion of the electric charge center of bR molecules during the photocycle. For a single molecule, however, we assume a discrete stationary value for the dipole moment of an intermediate (denoted by μ_i, where *i* stands for the actual photocycle intermediate). The measured D(t) is a smooth function related to the whole sample, which always contains a mixture of intermediates. The corresponding μ_i's, therefore, can be determined, if we know the time course of the concentration of the photocycle intermediates $[c_i(t)]$ in the macroscopic sample (see *Materials and Methods*). To determine the latter, absorption kinetics were measured on the same sample, under the same conditions, in addition to the photoelectric experiments. Difference spectra detected at various time delays after the exciting flash are shown in Fig. 5. A rank of 3 for the data matrix constructed from the time-resolved difference spectra was obtained by singular value decomposition. This establishes that three spectrally distinct intermediates, K, L, and M, are present in the time window of the measurement (300 ns–400 μs), i.e., no

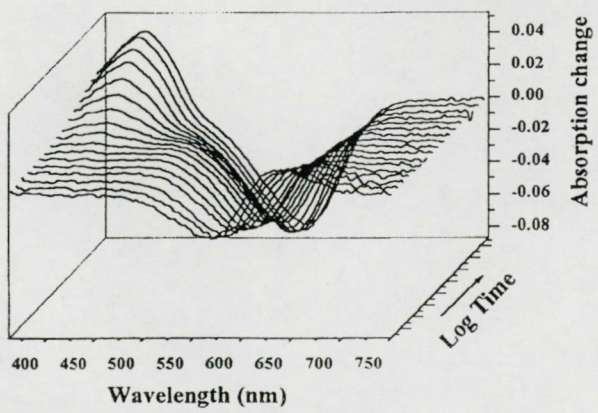


FIG. 5. Difference spectra in the visible wavelength range measured at logarithmically equidistant time delays after the laser flash excitation of electric–magnetic field orientation method-oriented pm. The evolution of the spectra is characteristic of the early transformation of the K intermediate to L and the subsequent accumulation of the blue-absorbing M intermediate.

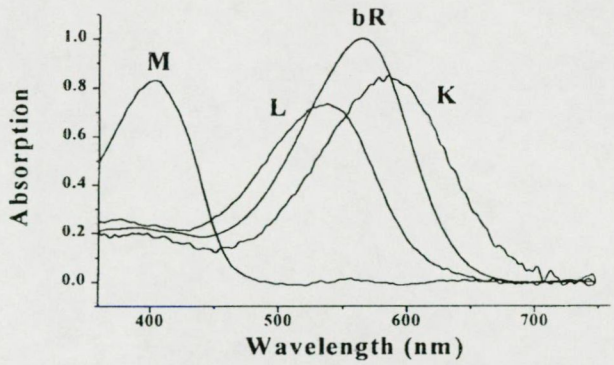


FIG. 6. Calculated absorption spectra of the photocycle intermediates K, L, and M, as well as that of the initial bR form. Spectra were obtained from the difference spectra in Fig. 5 with the Monte Carlo-based search algorithm. Substates of L and M are spectrally indistinguishable.

indication for the subsequent intermediates N and O was found. The absorption spectra of the K, L, and M intermediates (Fig. 6), as well as the corresponding time dependent concentrations (Fig. 7), were determined according to ref. 23, and the widely considered and tested photocycle model $bR \Rightarrow K \rightleftharpoons L_1 \rightleftharpoons L_2 \rightleftharpoons M_1 \rightleftharpoons M_2$ was used to calculate the intermediate concentrations up to 400 μs.

The result of a multilinear regression fit of D(t) with linear combinations of $c_i(t)$'s yielded the μ_i values shown in Fig. 8. As for the interpretation of the results in the molecular coordinate system, a positive μ_i^z corresponds to the displacement of a positive charge in the direction of the overall proton pumping. A positive μ_i^y results from the displacement of a positive charge parallel to the projection of the retinal transition dipole moment onto the membrane plane, roughly from helix E toward helix A. Finally, a positive μ_i^x corresponds to the displacement of a positive charge perpendicularly to both μ_i^y and μ_i^z, approximately from the direction of helix C to helix F.

The μ_i's determined in this way may contain errors from several sources. First, the correct photocycle model cannot be uniquely determined in most cases unless certain assumptions are made, such as: first-order kinetics for the elementary reactions (25), plausible restrictions for the calculated absorption spectra of the intermediates (23, 26) and the photocyle scheme (24, 27), etc. Another problem arises when the criterion of linear independency for the set of $c_i(t)$'s is poorly satisfied. In addition to the measuring errors, these

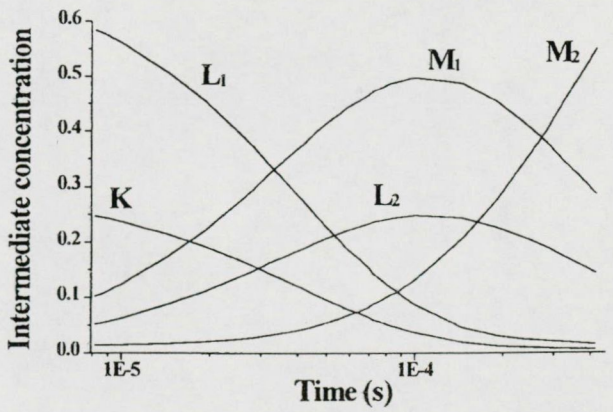


FIG. 7. Intermediate kinetics obtained after fitting the linear, reversible photocycle scheme (see text) to the intermediate concentrations calculated by the nonnegative least-squares fit of the pure intermediate spectra to the time-dependent difference spectra.

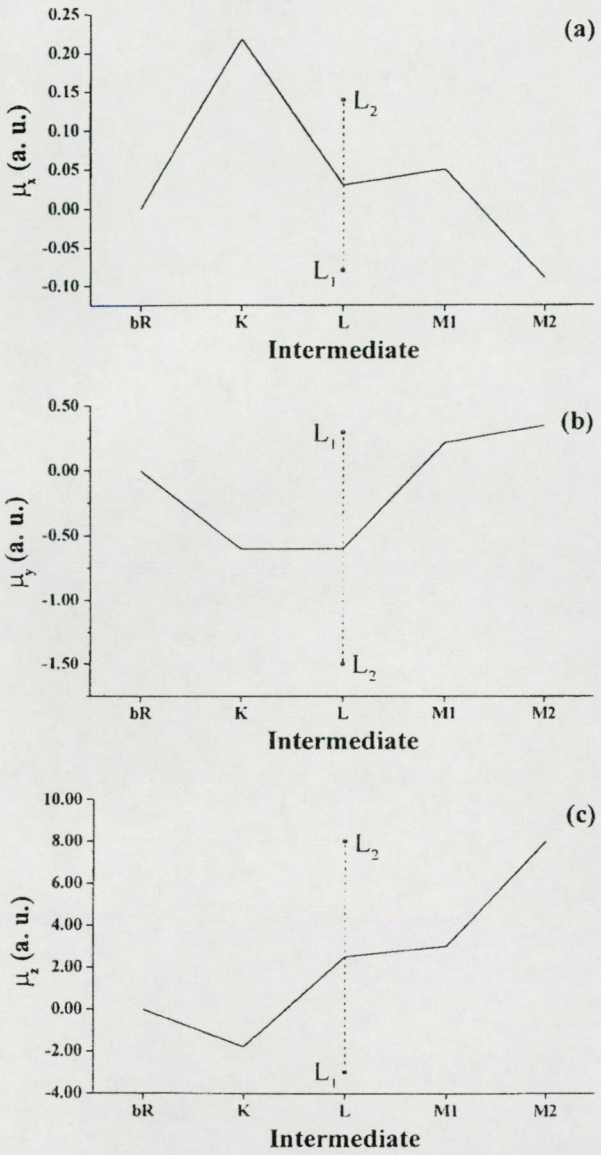


FIG. 8. Dipole moment changes of the photocycle intermediates relative to the initial state obtained from the three-dimensional photoelectric measurements. The individual dipole moments for the two L substates provided by our analysis were averaged to yield a single value comparable to the corresponding value from molecular dynamics calculations.

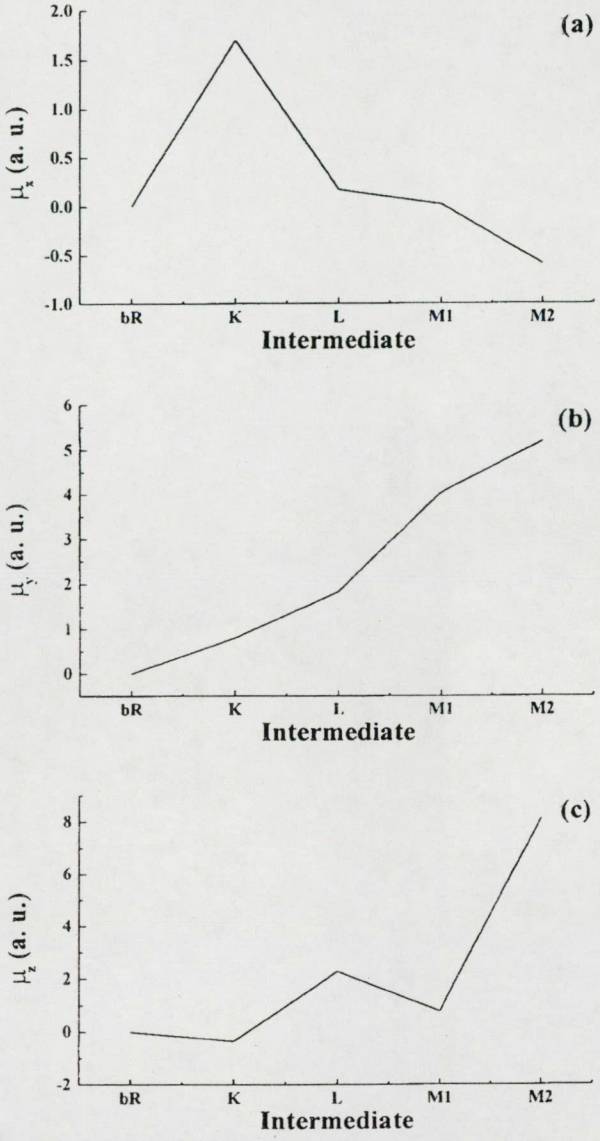


FIG. 9. Dipole moment changes of the photocycle intermediates relative to the initial state computed from the atomic coordinates and partial charges provided by molecular dynamics calculations (refs. 17 and 18; K. Schulten, personal communication).

are the major sources contributing to the uncertainty of the μ_i values.

In the simplest photocyle model that could account for the absorption kinetic changes in the time window of the measurements (i.e., the "first half of the photocycle"), a second L-like intermediate (L_2) in a rapid equilibrium with M_1 had to be assumed. The model with a single L yielded a considerably poorer fit, similar to some earlier observations (24, 26, 29). As a consequence of the rapid equilibrium, the time courses of L_2 and M_1 are nearly identical, which is reflected in high and mutually dependent error levels of the corresponding calculated dipole moments. Nevertheless, the results are consistent with earlier experimental data, obtained under different circumstances (13, 28), as well as with molecular dynamics calculations for the early events (up to the K intermediate) of the photocycle (18).

Table 1. Calculated electric dipole moment changes of the photocycle intermediates relative to bR

	μ_K			μ_L			μ_{M1}			μ_{M2}		
Group	*x*	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
1	1.70	0.80	−0.35	0.17	1.83	2.27	0.02	4.03	0.78	−0.59	5.19	8.13
2	—	—	—	5.07	5.28	0.10	1.41	7.85	−0.27	3.69	4.00	6.59
				0.56	14.0	11.5						

Data taken from the molecular dynamics structures provided by Schulten (Group 1) and Scharnagl *et al.* (Group 2). Data from Group 2 involved no K intermediate but two L intermediates.

DISCUSSION

As indicated in *Materials and Methods*, the experimentally determined μ_i^k values contain an undefined proportionality constant that is different in the three spatial directions. Consequently, the Cartesian components of a μ_i vector cannot be directly compared. However, the related μ_i^x, μ_i^y, and μ_i^z components do contain important information about charge rearrangements inside the protein during the photocycle; therefore, their comparison with theoretically predicted data is of primary interest.

The later part of the photocycle still lies beyond the time window of traditional molecular dynamics theory. Based on the high-resolution structure of bR determined by diffraction experiments, several attempts have been made by using semiempirical or classical force field calculations and the method of simulated annealing to compute the further progression of the photocycle (18–20, 30). Taking into account some known empirical parameters (p*K*'s, Fourier-transformed IR data, etc.) and by using plausible assumptions, structures for the L and M intermediates have also been published. We have calculated the μ_i values from two corresponding structures by using the atomic coordinates and partial charge values, which include the surrounding water molecules in addition to the atoms of bR (C. Scharnagl and K. Schulten, personal communication). Different molecular dynamics calculations yielded slightly different intermediate structures. Comparing the dipole moments (Table 1), one can conclude that the μ_i parameters represent a sensitive measure of these structural differences. The publications of Schulten's group (18, 19) contain a more complete data set, including the K intermediate, even though they considered only one L intermediate in the photocycle. To make our results directly comparable with their data, a single value (μ_L) stands for the experimentally determined dipole moment of the L intemediates in Fig. 8, representing the arithmetic average of μ_{L1} and μ_{L2}. Considering the numerous factors that may contribute to uncertainties of both the experimental and the theoretical values related to charge motions in the *x*–*z* plane, we find a noteworthy correlation between the normalized μ_i values determined by the two entirely different methods (Figs. 8 and 9).

Preliminary analysis regarding the assignment of components of the electric signal to rearrangements of particular charged and dipolar groups inside the bacteriorhodopsin molecule suggests that, while proton motion makes a major contribution to the signals measured in the *z* direction, the signals measured in the *x* and *y* directions are apparently dominated by the reorientation and redistribution of water molecules (unpublished data).

These results imply that there is a good chance to find an improved theoretical model that could account for all of the experimental results (including the observed charge displacements in the *y* direction), for which our method will provide the decisive test. Establishing such a model for the functional description of an ion pump protein at atomic level should have a strong impact on protein research in general.

The authors thank Profs. Klaus Schulten and Christina Scharnagl for sending us the results of their molecular dynamics calculations. We also thank Dr. G. Rikken for his technical help during sample preparation in the Grenoble High Magnetic Field Laboratory and Dr. G. Groma for providing the program (RATE 2.1) used in the fitting of kinetic models. This work was supported by grants from the Hungarian Scientific Research Fund (OTKA T29814, OTKA T017017, and OTKA T020470) and by a North Atlantic Treaty Organization Collaborative Research Grant (CRG 930419).

1. Oesterhelt, D. & Stoeckenius, W. (1974) *Methods Enzymol.* **31**, 667-678.
2. Lanyi, J. K. (1993) *Biochim. Biophys. Acta* **1183**, 241–261.
3. Ebrey, T. G. (1993) in *Thermodynamics of Membranes, Receptors and Channels*, ed. Jackson, M. (CRC, Boca Raton, FL), pp. 353–387.
4. Kimura, Y., Vassylyev, D. G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T. & Fujuyoshi, Y. (1997) *Nature (London)* **389**, 206–210.
5. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. (1996) *J. Mol. Biol.* **259**, 393–421.
6. Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. (1997) *Science* **277**, 1676–1681.
7. Luecke, H., Richter, H. T. & Lanyi, J. K. (1998) *Science* **280**, 1934–1937.
8. Lozier, R. H., Bogomolni, R. A. & Stoeckenius, W. (1975) *Biophys. J.* **15**, 955–962.
9. Skulachev, V. P. (1979) *Methods Enzymol.* **55**, 586–603.
10. Dancsházy, Z. & Karvaly, B. (1976) *FEBS Lett.* **72**, 136–138.
11. Bamberg, E., Dencher, N. A., Fahr, A. & Heyn, M. P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7502–7506.
12. Keszthelyi, L. (1980) *Biochim. Biophys. Acta* **598**, 429–436.
13. Keszthelyi, L. & Ormos, P. (1980) *FEBS Lett.* **109**, 189–194.
14. Dér, A., Hargittai, P. & Simon, J. (1985) *J. Biochem. Biophys. Methods* **10**, 295–300.
15. Keszthelyi, L. & Ormos, P. (1989) *J. Membr. Biol.* **109**, 193–200.
16. Dér, A., Tóth-Boconádi, R., Keszthelyi, L., Kramer, H. & Stoeckenius, W. (1995) *FEBS Lett.* **377**, 419–420.
17. Dér, A. & Ormos, P. (1995) *Biophys. Chem.* **56**, 159–163.
18. Humphrey, W., Xu, D., Sheves, M. & Schulten, K. (1995) *J. Phys. Chem.* **99**, 14549–14560.
19. Xu, D., Sheves, M. & Schulten, K. (1995) *Biophys. J.* **69**, 2745–2760.
20. Scharnagl, C., Hettenkofer, J. & Fischer, S. F. (1994) *Int. J. Quantum Chem.* **21**, 33–56.
21. Barabás, K., Dér, A., Dancsházy, Z., Marden, M., Ormos, P. & Keszthelyi, L. (1983) *Biophys. J.* **43**, 5–11.
22. Maret, G. & Weill, G. (1983) *Biopolymers* **22**, 2727–2744.
23. Gergely, C., Zimányi, L. & Váró, G. (1997) *J. Phys. Chem.* **101**, 9390–9395.
24. Dér, A., Tóth-Boconádi, R. & Száraz, S. (1992) *Colloq. INSERM* **221**, 197–200.
25. Nagle, J. F. (1991) *Biophys. J.* **54**, 476–487.
26. Zimányi, L. & Lanyi, J. K. (1993) *Biophys. J.* **64**, 240–251.
27. Ludmann, K., Gergely, C. & Váró, G. (1998) *Biophys. J.*, in press.
28. Ludmann, K., Gergely, C., Dér, A. & Váró, G. (1998) *Biophys. J.*, in press.
29. Gergely, C., Ganea, C., Groma, G. & Váró, G. (1993) *Biophys. J.* **65**, 2478–2483.
30. Engels, M., Gerwert, K. & Bashford, D. (1995) *Biophys. Chem.* **56**, 95–104.