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*,[†]

Contribution from the 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, Florida 32306-4390

Received August 15, 2000. Revised Manuscript Received January 25, 2001

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 pH-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. U.S.A.* **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 pH-dependent transitions of the photocycle, and offers an alternative mechanism for the pH dependence of the yield and kinetics of the late O intermediate.

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 proton-motive 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 FTIR spectroscopy.[4−6] A schematic description of the photocycle with the intermediates' estimated absorption maxima[7] and characteristic transition times is shown by the sequence:

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

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, 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 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

* Address correspondence to this author. E-mail: zimanyi@nucleus.szbk.u-szeged.hu.
[†] Hungarian Academy of Sciences.
[‡] Florida State University.

(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**, *406*, 649−653.
(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.

10.1021/ja0030414 CCC: $20.00
Published on Web 03/14/2001

of several intermediates at all times after the initiation of the photocycle by a 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 task is to decompose the measured spectra to obtain the pure intermediate spectra and the intermediate kinetics and, finally, to determine the most probable and physically meaningful reaction scheme that 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 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 pH below the (transient) pK = 5.8 of the proton release group, and takes place after the cytoplasmic proton uptake in the millisecond range, when the p*K* is further lowered.[21,22] Therefore, the WT photocycle is complicated by a pH-dependent branching unless the pH is significantly higher than this pK.[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 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]

Extension of the SVD-SM method to five components leads to the accurate absorption spectra of the intermediates for the E204Q mutant BR. No pH or temperature dependence of the spectra was found in the pH 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 pH 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. Analysis of the N ⇌ O → BR sequence of the photocycle reveals the energetic relations of these forms. An explanation for the anomalous temperature and pH 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 to prevent aggregation and rotational diffusion and to reduce light scattering. Gel slabs of 4 mm × 9 mm were soaked several times in the desired solvent and placed in a sample cuvette with 9 mm path length 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 nonexcited sample. All measurements were carried out in 100 mM NaCl, 50 mM Bis-Tris propane buffer at pH 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 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 to 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.

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 ⇌ O → BR segment of the photocycle was fitted to the analytical solution of the corresponding differential equations with Matlab (The MathWorks, Inc.). The entire SVD-SM procedure was also programmed in Matlab.

(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. U.S.A.* **1999**, *96*, 4408−4413.
(19) Zimányi, L.; Kulcsár, Á.; Lanyi, J. K.; Sears, D. F.; Saltiel, J. *Proc. Natl. Acad. Sci. U.S.A.* **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.

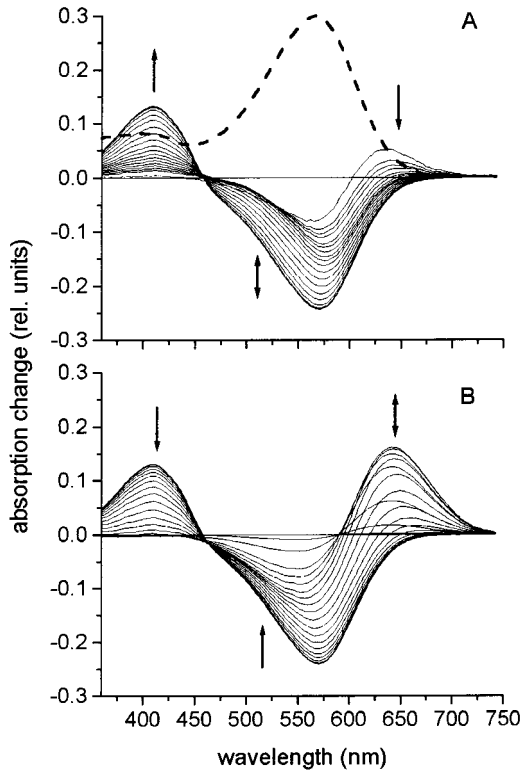

Figure 1. Flash-induced difference spectra of E204Q bacteriorhodopsin at pH 5, $T = 15$ °C: (A) spectra Nos. 1−20, time domain 520 ns to 370 μs, and the absorption spectrum of the initial form (dashed line); and (B) spectra Nos. 21−40, time domain 520 μs to 800 ms. Arrows indicate the time evolution of the signal. The spectra are reconstituted by using the five most significant SVD vectors.

Results

Initial Treatment of the Data. The data analysis is demonstrated on the experiment at pH 5 and 15 °C. Difference spectra on E204Q bacteriorhodopsin after laser pulse excitation, reconstructed from the first five SVD components, and normalized by the absorption of the sample at 570 nm in the initial state, are shown in Figure 1. The first five eigenvectors account for 99.99% of the total variance of the signal. The autocorrelations of the first five spectral eigenvectors were >0.91, and those of the first five temporal eigenvectors were >0.77. The autocorrelations of the 6th spectral and temporal eigenvectors were 0.309 and −0.188, respectively. Hence we conclude from the singular values and the autocorrelations that the first five eigenvectors represent the data adequately. This is in accordance with the expectation of five 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 three-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 five 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**:

$$\mathbf{D} = \mathbf{U}\cdot\mathbf{S}\cdot\mathbf{V}^{\mathbf{T}} \quad (1)$$

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

$$\mathbf{A} = (\mathbf{S}\cdot\mathbf{V}^{\mathbf{T}})^{\mathbf{T}} \quad (2)$$

is a matrix whose elements are the combination coefficients in the equivalent principal component analysis. Any jth column of the **D** matrix is the linear combination of the spectral eigenvectors in **U** with the combination coefficients in the jth 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

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

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 five-component vector **R**, which minimizes the deviation of eq 3 from unity, with the 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 five-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 three-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 coefficient, by which the M − BR difference spectrum was calculated, was obtained by fitting the >540 nm portion of the $-p\times$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. The LM side of the KLM triangle could be located by moving along lines parallel to the direction defined by the first 4 data points, and monitoring the onset of the corresponding difference spectra 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, assuming that neither L nor M absorbs here, only K does. The vertex of L is obtained by moving along the LM line away from M until the average of the difference spectra in the <410 nm range reaches zero, supposing that the average absorption of L in this region equals

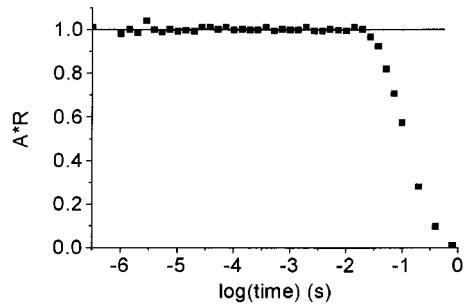

Figure 2. The stoichiometric behavior of the data from Figure 1. The left-hand side of eq 3 is plotted versus time. Data points represent the difference spectra in Figure 1.

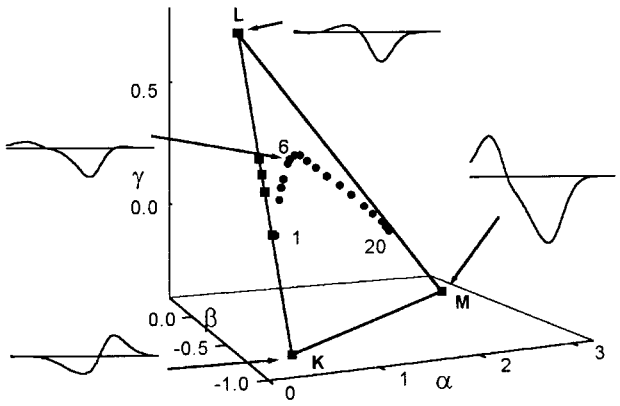

Figure 3. Demonstration of self-modeling for the first half of the photocycle. Three-dimensional plot of the combination coefficients in matrix **A** (●) representing the original difference spectra in Figure 1A ($\alpha_j = \mathbf{A}_{j,1}$ $\beta_j = \mathbf{A}_{j,2}$ $\gamma_j = \mathbf{A}_{j,3}$, where **A** is defined by eq 2, and *j* refers to the *j*th difference spectrum, so that spectrum$_j = \alpha_j\mathbf{U}_1 + \beta_j\mathbf{U}_2 + \gamma_j\mathbf{U}_3$, $\mathbf{U}_k$ being the *k*th column of **U**). The points which determine the K−L line and the vertexes 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).

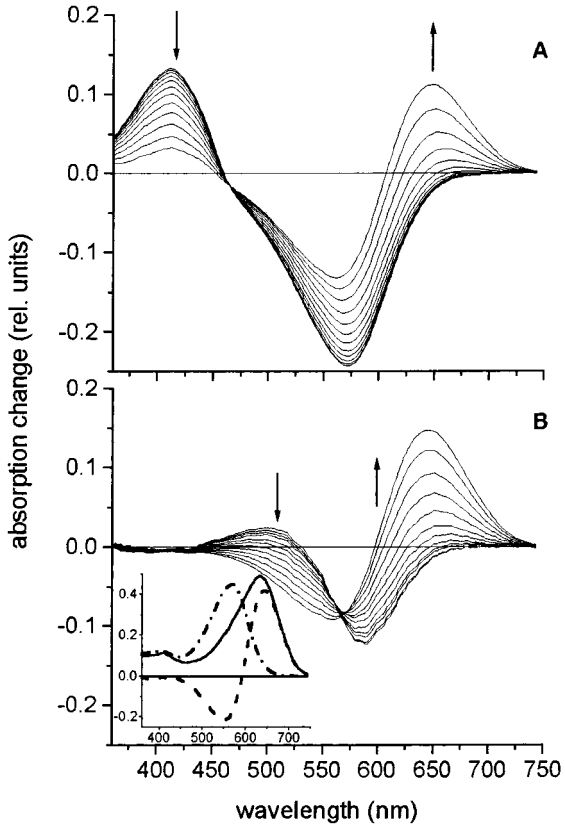

Figure 4. (A) Difference spectra Nos. 21−32 from Figure 1B, which serve as the basis to determine the spectra of the pure N and O intermediates. (B) Spectra from part A after subtraction of the M intermediate. The resulting spectra are assumed to be contributed by the L, N, and O intermediates. For details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

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 M is faster than in either the WT or the D96N mutant, even the earliest 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 four data points, until the <410 nm absorbance diminished (■). The slightly lower absorbance of K relative to BR in this region was ignored by this assumption. The K vertex was located along the KL line while monitoring the absolute spectra, 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 vertexes 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 spectra Nos. 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 spectrum (No. 37, dashed line) is negative in this region. By adding an appropriately scaled BR spectrum (dash dotted line) to it, we obtained the solid line, which demonstrates that the 37th 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. SVD of the matrix plotted in Figure 4B confirmed that it is a three-component system, and the stoichiometric plane was also determined. The adherence of the data points to the stoichiometric plane is demonstrated in Figure

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

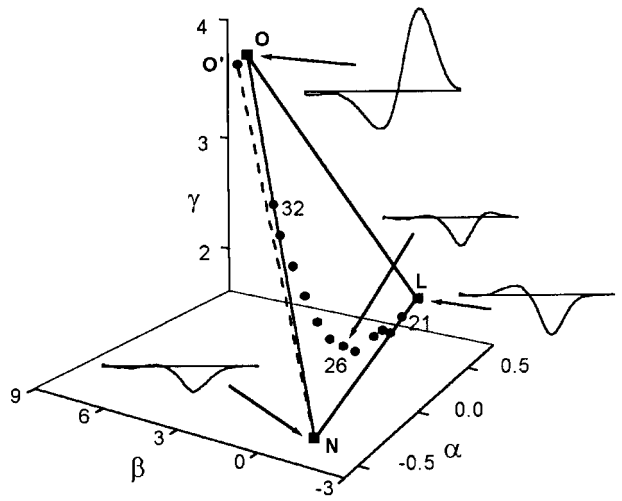

Figure 5. Demonstration of self-modeling on the difference spectra in Figure 4B. Three-dimensional plot of the combination coefficients in matrix **A** (●) (for details see the caption to Figure 3) and the vertexes corresponding to the pure difference spectra of the L, N, and O intermediates (■) outlining the stoichiometric plane. Insets: The difference spectra corresponding to the pure intermediates and a mixture difference spectrum (No. 26).

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. On the basis of 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, 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 1). 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, while the latter provided unrealistically tall and narrow spectra. There is 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 26th data point.

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.

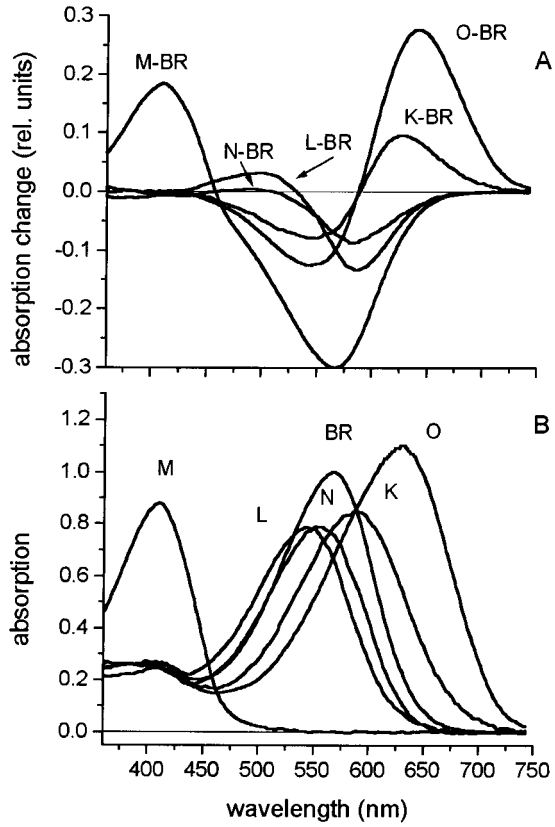

Figure 6. Difference spectra (A) and absolute spectra (B) of the photocycle intermediates K, L, M, N, and O, as determined by SVD-SM, and the spectrum of the initial state BR. The spectra are the average of spectra determined from nine independent experiments on E204Q BR at pH 5, 7, and 9 and $T = 15$,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

The intermediate spectra were calculated at all three pH values and all three temperatures with SVD-SM in a similar manner. The N spectra were blue shifted by ca. 2 nm at pH 9, otherwise all spectra proved to be pH and temperature independent. The pH 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 1. 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

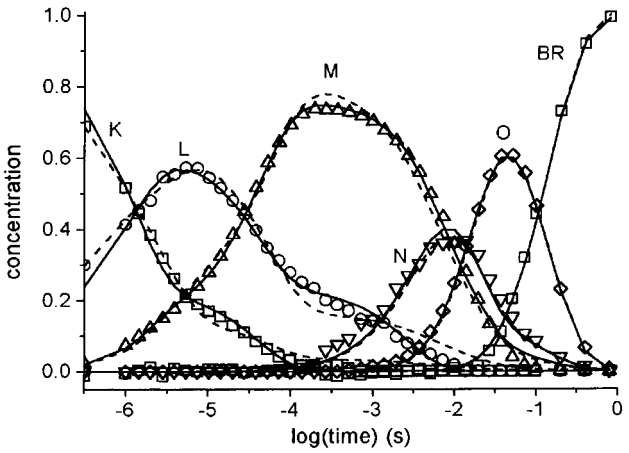

Figure 7. Kinetics of the photocycle intermediates of E204Q BR at pH 5, $T = 15$ °C (symbols). Dashed lines represent the fit by the photocycle in reaction scheme I, solid lines the fit by reaction scheme II (see text).

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 Nos. 22−40 with L, M, N, and O, in accordance with the findings that the first 20 spectra constitute a three-component system, and that K decays before the second 20 time points. Spectra Nos. 19−21 were fitted with all five 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 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 pH 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 the selection of the NO line 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 given in reaction I:

$$\mathrm{BR} \rightarrow \mathrm{K} \rightleftharpoons \mathrm{L}_1 \rightleftharpoons \mathrm{L}_2 \rightleftharpoons \mathrm{M}_1 \rightleftharpoons \mathrm{M}_2 \rightleftharpoons \mathrm{N} \rightleftharpoons \mathrm{O} \rightarrow \mathrm{BR} \quad \text{(I)}$$

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 pH values. A better fit with $\chi^2 = 0.0424$ was obtained with reaction II (solid lines), which contains L_2 in a cul-de-sac as proposed

$$\mathrm{BR} \rightarrow \mathrm{K} \rightleftharpoons \mathrm{L}_1 \rightleftharpoons \mathrm{M}_1 \rightleftharpoons \mathrm{M}_2 \rightleftharpoons \mathrm{N} \rightleftharpoons \mathrm{O} \rightarrow \mathrm{BR} \quad \text{(II)}$$
$$\Updownarrow$$
$$\mathrm{L}_2$$

earlier for the D96N mutant,[17] and which gave a similarly good fit to D96N data as reaction I in another study.[27] 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 reaction scheme did not affect the other rate constants. In the same manner, the $M_1 \leftarrow M_2$ back reaction was negligible at all pH and temperature values, as found for the WT photocycle at alkaline pH.

The rate constants obtained with reaction II increase monotonically with increasing temperature. There is no pH dependence of the rates up to intermediate N, except for a nonsystematic dependence of the $L_2 \rightarrow M_1$ transition (not shown). Further complicating the photocycle scheme with additional intermediate substates and pathways could probably remove this uncertainty; however, the possibilities are numerous and may not yield conclusive evidence. Instead, we concentrate on the systematic pH dependence at all three temperatures of the N → O and the N ← 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. It has been reported[28,29] that the titration of three 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 pH dependence of phenomenological rate constants, thereby avoiding model-dependent conclusions. Other investigators,[30−32] however, fitted the measured O kinetics both with sums of exponentials to obtain phenomenological rates, and with the model $\mathrm{M} \rightleftharpoons \mathrm{N} \rightleftharpoons \mathrm{O} \rightarrow \mathrm{BR}$, which provides the pH 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 extracellular side in the O form. According to their model, above this pK proton release from this cluster is followed by its reprotonation by D85, a fast process, whereas below this pK proton release from the cluster is inhibited, and release must proceed directly from D85, a slower process. This pH 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 the kinetics not only 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 reaction III:

$$\mathrm{N} \rightleftharpoons \mathrm{O} \rightarrow \mathrm{BR} \quad \text{(III)}$$

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 rates at pH 5 and 15 °C were 34, 0.3, 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

(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.

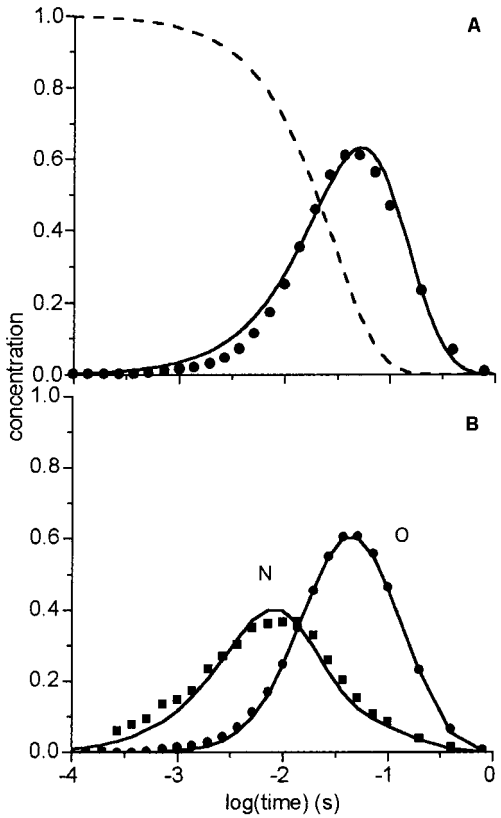

Figure 8. (A) Fit of the kinetics of the O intermediate (symbols) separately, by the simplified reaction scheme III. The assumed kinetics of the precursor to O (N, or the equilibrated M + N) is shown by the dashed line. (B) Fit of the kinetics of the O intermediate by reaction scheme II, which correctly 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).

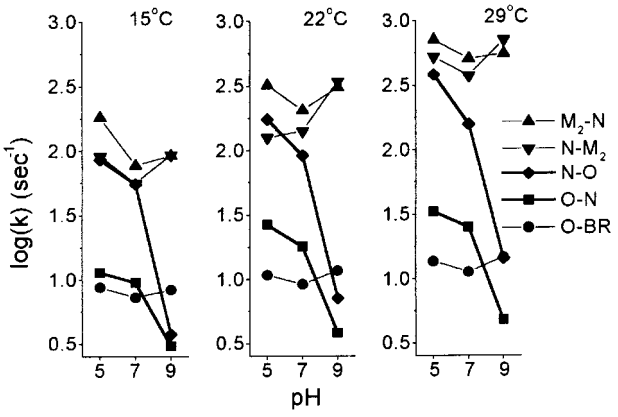

Figure 9. Dependence of the rate constants on pH and temperature, obtained from the fit of the photocycle reaction scheme II to the intermediates kinetics.

if the precursor to O, the N intermediate, was taken into account correctly (Figure 8B). The corresponding rates were 86, 11, 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 not only the knowledge of the kinetics of O (measured usually at 630 nm), but also the time evolution of its precursors. The latter can only be determined reliably from multichannel spectroscopy, with global analysis by methods such as SVD-SM.

Figure 9 shows that the rates of both the N → O and N ← O transitions decrease with increasing pH. This behavior can be explained by assuming that the transition state barrier between these two intermediates is increasing with increasing pH. Since the O → BR rate is pH -independent, the free energy of the O state is probably also pH 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 pH (Table 2). Although there is a little shift of the free energy content of N meaning its stabilization with increasing pH, the main factor controlling the rise, decay, and yield of O as a function of pH and temperature is the pH 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 pK should be close to 8, based on the similar pH dependence of the N → O and N ← O rates (Figure 9).

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 pH 5, 7, and 9

	$\Delta\Delta G$ (kJ/mol)			$\Delta G^{\ddagger}$ (kJ/mol)		
	pH 5	pH 7	pH 9	pH 5	pH 7	pH 9
N vs O	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

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 pH-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 nontrivial task to determine the correct intermediate kinetics from absorption spectroscopy. The D96N mutant provided a simpler system with three 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 power of SVD-SM 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 pH-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 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 four- rather than a five-component system. This is not the case in WT, where at pH below neutral some K intermediate remains during the onset of N.[33]

The spectrum of M is the most accurate one 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 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, and 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 vertexes, to avoid negative intermediate concentrations. It is reassuring that the intermediate spectra obtained from the nine separate experiments independently were almost identical, and their average 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 could 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 (reaction schemes I and II, respectively) gave similarly good fits for WT and D96N, our data clearly distinguish between them in favor of reaction scheme II at least in the case of the E204Q mutant. 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 toward 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 pH and temperature. 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 pH. 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 requires 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 pH, while others are pH-independent. The estimated pK for the control of both rates is around 8. The simultaneous pH-dependencies of the forward and reverse rates suggest that the barrier height between the two intermediates is pH-dependent, rather than the free energy levels of the intermediates themselves. It appears that the reisomerization of the retinal, the rate-limiting step in the N → O transition in E204Q,[37] may be catalyzed by a group with this pK, 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.[38,39]

It has been suggested[31] that at pH lower than the transient pK 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 pK plays no role in the lifetime and yield of O in E204Q,

(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. U.S.A.* **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) Brown, L. S.; Needleman, R.; Lanyi, J. K. *Biochemistry* **2000**, *39*, 938−945.

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

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

rendering their pH-dependence less complicated. Our results with E204Q 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 behavior of O as a function of pH and temperature. Complications are expected in WT due to the alternative deprotonation routes of D85 during O decay and BR recovery, 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.

JA0030414