Kinetic Evidence for an Obligatory Intermediate in the Folding of the Membrane Protein Bacteriorhodopsin†

Amjad Farooq‡

Department of Biochemistry, Imperial College of Science, Technology & Medicine, London SW7 2AY, U.K.

Received June 23, 1998; Revised Manuscript Received August 5, 1998

ABSTRACT: A photodiode array in conjunction with a rapid stopped-flow mixing method, with a millisecond time resolution, is used here to study the refolding of the membrane protein bacteriorhodopsin from an apoprotein state with a native-like secondary structure in mixed phospholipid/detergent micelles. Refolding to the native state is initiated by the rapid mixing of *all-trans*-retinal and the apoprotein bacterioopsin in mixed micelles. A lag phase of several seconds is observed in the appearance of the native state, as monitored by the increase in absorbance of the native chromophore. This observation demonstrates unequivocally that an intermediate is obligatory in the formation of bacteriorhodopsin. It is further shown that this intermediate is spectroscopically distinct from free retinal (absorbance maximum ∼ 380 nm) and bacteriorhodopsin (absorbance maximum ∼ 560 nm) and absorbs maximally at 430 nm. Evidence for the decay of the 430 nm intermediate into bacteriorhodopsin via three distinct parallel pathways is also provided. Taken together, these findings are used to describe a model in which distinct populations of the apoprotein in mixed micelles appear to fold along separate pathways via their corresponding intermediates into the native state. How the results of this study provide new insights into the mechanisms of protein folding is discussed.

The elucidation of the mechanisms of protein folding remains one of the mammoth challenges in modern biology. The widespread interest in this area does not merely arise from its relevance to processes such as protein synthesis, degradation, and transport in the body but more so from the recognized impact that this pursuit might have on medicine and biotechnology. Thus, folding studies hold promise for the treatment of disease (*1*) as well as in the commercial exploitation of recombinant proteins (*2*).

In the late 1960s, in a pioneering paper, Levinthal argued that a protein folds to its native state via specific intermediates and a specific pathway (*3*). Ever since, the requirement of intermediates in the protein folding process has been considered *de rigueur* (the classical view of protein folding). Intermediates have indeed been detected in many proteins, including barnase (*4*), ribonuclease A (*5*), ubiquitin (*6*), cytochrome *c* (*7*), and bovine pancreatic trypsin inhibitor (*8*). Recent experimental and theoretical advances in this area however suggest that intermediates may not be necessary, at least in the case of small proteins. Experimental observations show that some proteins fold via a simple two-state process with no accumulation of intermediates. Examples include chymotrypsin inhibitor-2 (*9*), the cold-shock protein CspB (*10*), monomeric λ repressor (*11*), and acyl-coenzyme A binding protein (*12*). Theoretical advances, largely based on statistical mechanical ideas and Monte Carlo simulations (*13*), have led to the new view of protein folding. In this new view, folding is envisaged as a distribution or ensemble of structures from the unfolded to the native state, and thereby, the requirement of specific intermediates and pathways is eliminated (*14*, *15*).

There is now widespread speculation that intermediates merely represent kinetic traps and off-pathway reactions that slow rather than accelerate productive folding (*10*, *16*, *17*). This view is corroborated by the direct observation of multiple folding pathways in a number of studies (*18*, *19*). Further, studies on cytochrome *c* (*17*) and lysozyme (*20*) indicate that folding can proceed in the absence of kinetically detectable intermediate species. Thus, the controversy surrounding the role of intermediates in protein folding remains unresolved. A kinetic criterion for the involvement of an obligatory intermediate in a reaction is the observation of a lag phase in the appearance of the native protein from a denatured or a partially refolded state (*16*, *19*). In the case of most proteins, the rates of refolding reactions are too fast for the detection of a lag phase using conventional stopped-flow mixing techniques. The folding of the membrane protein bacteriorhodopsin however is a relatively slow process with lifetimes in the second to minute range (*21*, *22*). Bacteriorhodopsin thus serves as a candidate *par excellence* for investigating the presence of a lag phase in the formation of its native state.

Bacteriorhodopsin is the sole protein of the purple membrane of *Halobacterium salinarium*, where it acts as a light-driven proton pump (*23*, *24*). This small integral

† This work was supported by the BBSRC.

‡Present address: Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine, One Gustave L Levy Place, Box 1677, New York, NY 10029-6574. Telephone: (212) 824-8236. Fax: (212) 849-2456. E-mail: amjad@inka.mssm.edu.

10.1021/bi981485v CCC: $15.00
Published on Web 10/02/1998

membrane protein is comprised of an apoprotein bacterioopsin to which a retinal moiety is covalently attached via a Schiff base at Lys216 (*25*). The biological activity of bacteriorhodopsin is dependent upon the photoisomerization of its retinal chromophore. The native protein bacteriorhodopsin traverses the membrane seven times in the form of α-helices, and its three-dimensional structure is known to near-atomic resolution (*26*).

In this report, a stopped-flow mixing approach in conjunction with photodiode array spectroscopy is used to study the refolding reaction of bacteriorhodopsin from an apoprotein state with a native-like secondary structure in mixed DMPC/CHAPS[1] micelles. It is shown that the formation of the final native state is preceded by a lag phase of several seconds. This leads to the conclusion that an intermediate species must be involved in the folding of bacteriorhodopsin. The subsequent decay of this intermediate into bacteriorhodopsin via three distinct parallel pathways is also demonstrated. These observations form the basis of a model in which three distinct populations of the apoprotein are suggested to fold along separate pathways via their corresponding intermediates into the native state. The results of this study are discussed in light of both the classical and new views of protein folding.

MATERIALS AND METHODS

Materials. DMPC was purchased from Avanti (Alabaster, AL) and CHAPS from Calbiochem (Nottingham, U.K.). *all-trans*-Retinal (vitamin A aldehyde) and SDS (electrophoresis grade) were supplied by Sigma (Dorset, U.K.). All other reagents and chemicals were of analytical grade. Mixed DMPC/CHAPS micelles were prepared in sodium phosphate buffer (pH 6.0) and *all-trans*-retinal in ethanol as described previously (*22*, *27*).

Protein Purification and Delipidation. The apoprotein bacterioopsin was isolated from purple membrane using the standard procedure (*28*) and delipidated as described previously (*22*, *29*).

Refolding Experiments. Bacteriorhodopsin was refolded by the addition of retinal to the apoprotein bacterioopsin in mixed micelles (bO/DMPC/CHAPS). Briefly, *all-trans*-retinal in 1% (w/v) DMPC/1% (w/v) CHAPS/0.1% (w/v) SDS micelles and 50 mM sodium phosphate buffer (pH 6.0) was mixed, in the stopped-flow cuvette, with an equal volume of bO/DMPC/CHAPS. bO/DMPC/CHAPS was prepared by mixing manually 8 μM bO in 0.2% (w/v) SDS and 50 mM sodium phosphate buffer (pH 6.0) with an equal volume of 2% (w/v) DMPC/2% (w/v) CHAPS micelles in 50 mM sodium phosphate buffer (pH 6.0) and the mixture allowed to equilibrate for 30 min. Final bO and retinal concentrations were 2 μM. The final concentrations of all other reagents were 0.1% (w/v) SDS, 1% (w/v) DMPC, 1% (w/v) CHAPS, 50 mM sodium phosphate buffer, and <0.5% (v/v) ethanol. All measurements were taken at 22.0 ± 0.1 °C in the dark or dim red light. The final pH of the reaction mixture at this temperature was 6.0 ± 0.1. All experiments were carried out on bO preparations with refolding yields of 95% or more at pH 6.0. Several bO preparations differ in that they are delipidated separately but the source of purple membrane may be the same. Refolding yields were determined as described previously (*22*, *27*).

Stopped-Flow Spectroscopy. Time-resolved absorption spectra were recorded using an Applied Photophysics Sx.17MV stopped-flow spectrometer (Leatherhead, U.K.), equipped with a photodiode array detector. This detector comprises a linear 256-element diode array covering the range of 305−1100 nm, with a diode wavelength separation of about 3.3 nm and a dead time of ∼2−3 ms. Water was used as a reference for all experiments. An excitation bandwidth of 1 nm and a path length of 1 cm were used for all experiments. A 150 W xenon arc lamp was used as a light source, and its alignment was checked at the beginning of each experiment. A thermostatic water bath was used to maintain the temperature at 22.0 ± 0.1 °C. Time-resolved absorption spectra were collected over different time scales, ranging from 10 to 1000 s, over the wavelength range of 305−750 nm. For each time scale, 400 spectra were recorded. To improve the signal-to-noise ratio below 400 nm, where the intensity of the xenon lamp is lower, data over the wavelength range of 305−750 nm were collected in two parts. Data over the wavelength range of 305−400 nm were taken using an integration time of 100 ms, while data over the wavelength range of 400−750 were taken using an integration time of 10 ms. The two sets of data in the range of 305−400 and 400−750 nm were normalized to any changes in light intensity incurred during the course of the experiment prior to data analysis.

Exponential Kinetics. Time-resolved absorption spectra $S[t,\lambda]$ were analyzed to a sum of exponentials in a global analysis using the function

$$S[t,\lambda] = \sum\{\alpha_i[\lambda]\exp(-\nu_i t)\} + \delta[\lambda] \qquad (1)$$

where $\alpha_i[\lambda]$ is the observed wavelength-dependent amplitude and ν_i is the observed rate constant of a kinetic phase i and $\delta[\lambda]$ is the wavelength-dependent baseline offset. The parameters $\alpha_i[\lambda]$, $\delta[\lambda]$, and ν_i were directly calculated from the raw data by iterative reconvolution based on the Marquardt algorithm (*30*). The lifetimes are the reciprocals of the experimentally observed rate constants. The kinetic software LOOK was used for data analysis (Imperial College, London, U.K.). Transient spectra collected over different time scales were analyzed simultaneously. The quality of fits was assessed with plots of residuals. The quality of residuals for simultaneous analysis of data over different time scales is usually poor in comparison to the quality of the residuals resulting from the fit of the same data over separate time scales. The former method however is the method of choice in that it generates more reliable values for the kinetic parameters.

Gaussian Decomposition. The refolding reaction of bacteriorhodopsin proceeds through an intermediate. The absorption spectrum of the intermediate, calculated from the

[1] Abbreviations: bO, partially denatured apoprotein bacterioopsin in SDS; SDS, sodium dodecyl sulfate; DMPC, L-α-1,2-dimyristoylphosphatidylcholine; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; R, retinal; SVD, singular value decomposition; S, state of the apoprotein bacterioopsin in mixed micelles; N, native state of bacteriorhodopsin; I_{430}, 430 nm intermediate; ν, experimentally observed rate constant; k, intrinsic rate constant.

wavelength-dependent amplitudes resulting from a global analysis of time-resolved spectra to a sum of exponentials (vide supra), includes a contribution from free retinal. To determine the absorbance maximum of the intermediate precisely, the calculated spectrum of the intermediate $\epsilon[\lambda]$ was decomposed into Gaussian components corresponding to free retinal and the intermediate. This was accomplished using the equation based on a combination of Gaussian and Rayleigh functions

$$\epsilon[\lambda] = [a_r/\sigma_r\sqrt{(2\pi)}] \exp\{-^1/_2[(\lambda - \lambda_r)/\sigma_r]^2\} + [a_i/\sigma_i\sqrt{(2\pi)}] \exp\{-^1/_2[(\lambda - \lambda_i)/\sigma_i]^2\} + (\beta/\lambda^4) + \delta \quad (2)$$

where a_r and a_i are the areas, σ_r and σ_i equal 0.425 times the full width half-maxima, and λ_r and λ_i are the absorbance maxima of the absorption bands corresponding to free retinal and the intermediate, respectively. λ is the wavelength; β is a constant related to the Rayleigh factor, and δ is the baseline offset. Fitting was achieved by iterative reconvolution based on the Marquardt algorithm using ORIGIN (Microcal). The inclusion of the Rayleigh function (β/λ^4) improves fitting by separating the contribution of light scattering by the micelles from retinal absorption. Full details of this procedure can be found elsewhere (*27*).

SVD Analysis. Time-resolved absorption spectra were subjected to singular value decomposition (SVD) analysis so the number of independently absorbing species in the reaction of the apoprotein bacterioopsin with retinal could be determined. The SVD algorithm (*31*) decomposes a data matrix **Y** (time × wavelength) containing a set of time-dependent spectra into a matrix product

$$\mathbf{Y} = \mathbf{USV}^T \quad (3)$$

where columns of **V** contain time-dependent amplitudes that are related to the time-dependent concentration profiles of physical intermediates, columns of **U** contain wavelength-dependent amplitudes (basis spectra) that are related to the absolute spectra of physical intermediates ($\mathbf{V}^T$ is the transpose of **V**), and diagonal elements of **S** constitute singular values, i.e., weighting terms needed to normalize elements in **U** and **V** to the corresponding elements in **Y**. The result of a typical SVD analysis consists of a series of wavelength-dependent amplitudes, weighted by their singular values, and time-dependent amplitudes. Typically, a plot of singular values falls from the first and largest value and plateaus out at the value characteristic of the number of spectrally distinct species. SVD analysis thus gives a model-free indication of the number of independently absorbing species in a reaction.

RESULTS

The Formation of Bacteriorhodopsin Occurs via an Intermediate. It is well established that the apoprotein bacterioopsin in mixed phospholipid/detergent micelles contains fully formed transmembrane helices (*32*, *33*). Thus, conversion of this apoprotein state in mixed micelles into the native state of bacteriorhodopsin can be studied in the presence of retinal without the complications of the kinetics of helix formation and/or rearrangement/association (*22*, *33*). Refolding of bacteriorhodopsin from the apoprotein in mixed micelles, as opposed to the partially denatured state bO in SDS, is not only simple, but in so doing, one also avoids the problems associated with micelle mixing (*27*, *34*). Although protein fluorescence is a highly specific probe of folding, retinal absorbance is employed here to monitor both the disappearance of free retinal and the appearance of the native protein. Free retinal absorbs maximally at 380 nm, whereas the native protein bacteriorhodopsin, in which retinal is covalently attached via a Schiff base at Lys216 (*25*), absorbs maximally around 560 nm. This shift in the retinal absorption band is believed to result from the interaction of retinal Schiff base with amino acid residues lining the retinal binding pocket (*35*). Thus, time-resolving changes in the retinal absorption band should provide a convenient measure of the extent of the refolding reaction of bacteriorhodopsin.

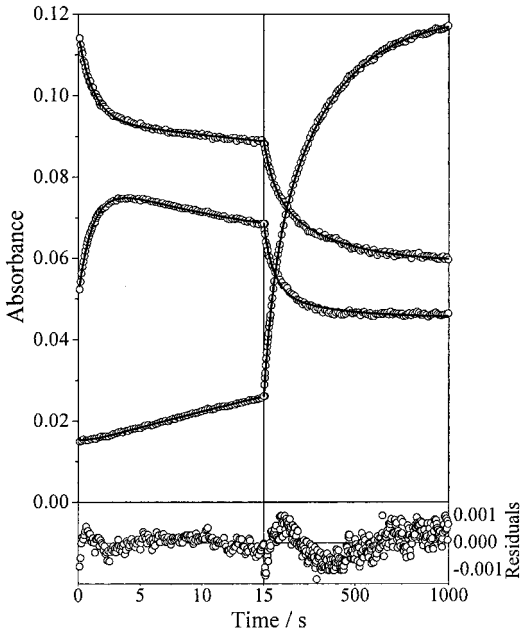


FIGURE 1: Bacteriorhodopsin refolding. *all-trans*-Retinal in DMPC/CHAPS micelles was mixed, in the stopped-flow cuvette, with an equal volume of the apoprotein bacterioopsin in mixed micelles (bO/DMPC/CHAPS). bO/DMPC/CHAPS was prepared by mixing manually DMPC/CHAPS micelles with an equal volume of bO partially denatured in SDS and allowed to equilibrate for 30 min. Time-resolved absorption spectra were collected over the 305−750 nm region using a photodiode array (not shown). Kinetic traces derived from these spectra at 380, 430, and 560 nm are shown (○). Data are shown over two time periods. Data shown over the 0−15 s region were collected over a full scale of 16 s. Data shown over the 15−1000 s region were collected over a full scale of 1000 s. For each full scale, 400 spectra were recorded. Not all data points are shown. The curves from top to bottom shown over the 0−15 s region correspond to wavelengths of 380, 430, and 560 nm, respectively. Time-resolved spectra collected over the 0−1000 s region and wavelength range of 305−750 nm were fit to a sum of four exponentials (phases ν_1−ν_4) in a global analysis. Transient spectra collected over different time scales were analyzed simultaneously using the kinetic analysis software LOOK. The fits to raw data are indicated (solid lines). The residuals at 430 nm are also shown underneath the main panel.

Time-resolved absorption spectra were collected using a photodiode array (not shown). The lack of a single isobestic point in these transient spectra has been reported previously (*34*), implying that an intermediate is involved in the refolding reaction of bacteriorhodopsin. Figure 1 shows the corresponding kinetic traces at 380, 430, and 560 nm. The

absorbance at 430 nm is that of an intermediate species (vide infra). As can be seen, the initial rate of decay of absorbance at 380 nm is matched by the initial rate of increase of absorbance at 430 nm. There is however little change in absorbance at 560 nm over the first few seconds of the refolding reaction of bacteriorhodopsin. The presence of this so-called lag phase demonstrates unequivocally that the intermediate is obligatory in the refolding reaction of bacteriorhodopsin. The time-resolved absorption spectra collected over various time scales in the region of 0−1000 s were best described by a sum of four exponentials in a global analysis, with rate constants of 2.2 (ν_1), 0.13 (ν_2), 0.013 (ν_3), and 0.0032 s^{-1} (ν_4). These rate constants and their corresponding wavelength-dependent amplitudes were determined using eq 1. The information that can be obtained from the wavelength-dependent amplitudes of these four kinetic phases is outlined below.

A preliminary account of the data presented in Figure 1 has been briefly reported previously (*34*). In this previous study, however, data were only collected over the first 10 s and could be approximated by a biexponential. Collection of data over the 0−1000 s time scale leads to a much more accurate representation of the process of bacteriorhodopsin folding that occurs over a minute range. Thus, analysis of data over a longer time scale has resulted in resolution of two additional kinetic phases here that could not possibly have been detected from the analysis of the previous data over a much shorter time scale. It is also worth mentioning that there is an 8-fold increase in the density of the data presented here compared with that of the data mentioned in the previous report (*34*). In the previous study, only 100 transient absorption spectra were collected, whereas the data analyzed here represent a total of 800 spectra. Also worth pointing out is the fact that the rate constant of 0.13 s^{-1} reported here is not to be confused with the rate constant of 0.16 s^{-1} reported in the previous study (*34*). In fact, these two rate constants represent two different molecular processes. The rate constant of 0.16 s^{-1} results from the addition of retinal to a denatured state of the apoprotein in SDS, whereas the process under investigation here is the addition of retinal to the apoprotein in mixed micelles. Moreover, the rate constant of 0.16 s^{-1} was assigned to the decay of retinal into the intermediate, whereas the rate constant of 0.13 s^{-1} here is assigned to the decay of the intermediate into bacteriorhodopsin (vide infra).

The Intermediate Decays into Bacteriorhodopsin via Multiple Pathways. The wavelength-dependent amplitudes of four kinetic phases $\nu_1-\nu_4$ can be used to assign them to particular stages during the folding of bacteriorhodopsin (Figure 2). These amplitudes are essentially difference spectra between the decaying and nondecaying species (assuming a priori that each exponential denotes an independent first-order reaction), and thus consist of positive and negative bands. A positive amplitude indicates a decrease in absorbance and a negative amplitude an increase. The amplitude of the fastest phase ν_1 consists of a positive band centered at 380 nm and a negative band in the 420−440 nm region, an indication of the decay of retinal into the intermediate absorbing in the 420−440 nm region. The amplitudes of slower phases $\nu_2-\nu_4$ are comprised of positive bands in the 400−440 nm region and negative bands around

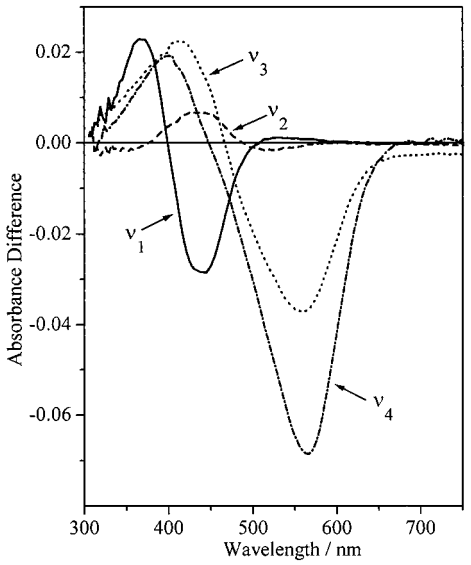


FIGURE 2: Wavelength-dependent amplitudes. The amplitudes of four kinetic phases $\nu_1-\nu_4$ resulting from the fit of a four-exponential function to the time-resolved spectra (see the legend of Figure 1) in a global analysis are shown.

560 nm, and thus indicate the decay of the intermediate into bacteriorhodopsin. The fact that an intermediate appears to be obligatory (vide supra) implies that phases $\nu_2-\nu_4$ cannot possibly represent the decay of retinal directly into bacteriorhodopsin. The amplitude of phase ν_2 appears not to be clearly resolved from phases ν_3 and ν_4 (Figure 2). This is probably the result of multiple routes involving the decay of the intermediate into bacteriorhodopsin.

The area under a wavelength-dependent amplitude provides a qualitative measure of the change in the oscillator strength of the retinal absorption band (*27*). As previously reported (*34*), the oscillator strength of the retinal absorption band in free retinal and in the intermediate is the same (the area under the amplitude of the kinetic phase ν_1 is close to zero) (Figure 2). This observation is taken to imply that retinal is noncovalently bound to the apoprotein within the intermediate. In contrast, the oscillator strength of the retinal absorption band undergoes significant change upon the decay of the intermediate into bacteriorhodopsin (the areas under the amplitudes of kinetic phases $\nu_2-\nu_4$ do not equal zero) (Figure 2). This change in the oscillator strength of the retinal absorption band is correlated to the formation of Schiff base between retinal and the apoprotein (see Discussion).

The Intermediate Absorbs Maximally at 430 nm. The wavelength-dependent amplitudes are related to the absolute spectra of the various species involved in the refolding reaction of bacteriorhodopsin. Thus, the absorption spectrum of free retinal can be determined from the summation of wavelength-dependent amplitudes of all four kinetic phases ($\nu_1-\nu_4$) and the wavelength-dependent baseline offset (parameter $\delta[\lambda]$ in eq 1) (*27*). Likewise, the absorption spectrum of the intermediate was obtained by subtraction of the wavelength-dependent amplitude of phase ν_1 (see Figure 2) from the calculated spectrum of free retinal. The absorption spectrum of the intermediate however includes a contribution from free retinal (Figure 3). This is a consequence of the transient nature of the intermediate. To obtain the absorbance maximum of the intermediate accurately, its calculated

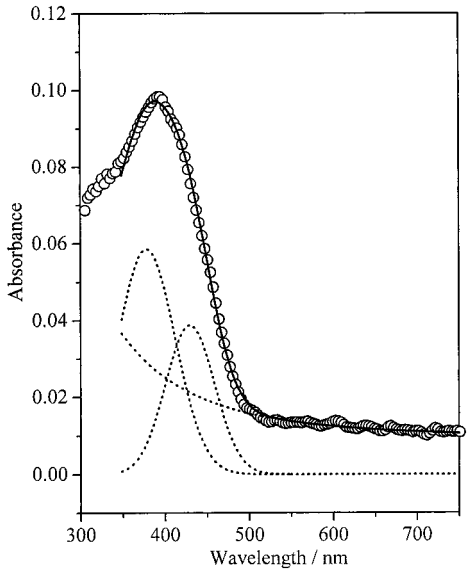


FIGURE 3: Decomposition of the absorption spectrum of the intermediate. The absorption spectrum of the intermediate (○) was decomposed into Gaussian components (···) corresponding to free retinal (absorbance maximum ~ 380 nm) and the intermediate (absorbance maximum ~ 430 nm) using eq 2. This equation is based on a combination of Gaussian and Rayleigh functions (see Materials and Methods). The fit of eq 2 to raw data is indicated (solid line). The absorption spectrum of the intermediate was calculated by subtracting the wavelength-dependent amplitude of phase ν_1 (see Figure 2) from the absorption spectrum of free retinal. The absorption spectrum of free retinal (not shown) was determined from the summation of the wavelength-dependent amplitudes of all four kinetic phases ($\nu_1-\nu_4$) and the wavelength-dependent baseline offset (parameter $\delta[\lambda]$ in eq 1). The dotted curve corresponds to the Rayleigh component as a result of the scattering of light by the micelles.

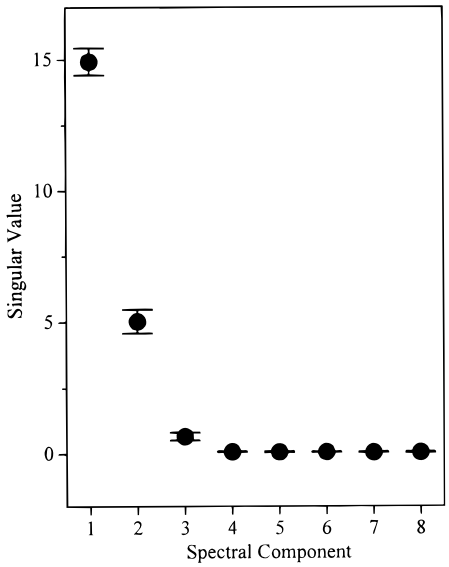


FIGURE 4: SVD analysis. Time-resolved absorption spectra collected over the 0−1000 s region and the wavelength range of 305−750 nm (see the legend of Figure 1) were analyzed by SVD algorithm. The SVD algorithm decomposes the data using eq 3. A plot of the singular value vs the number of spectrally distinct components is shown. Errors were obtained from measurements on three different bO preparations. Error bars are one standard deviation.

absorption spectrum was decomposed using eq 2 on the basis of a combination of Gaussian and Rayleigh functions. From the decomposed Gaussian bands shown in Figure 3, it is evident that the intermediate absorbs maximally at 430 nm. The shift in the retinal absorption band from 380 nm in free retinal to 430 nm in the intermediate is probably due to a change in the polarity of the solvent environment rather than due to Schiff base formation (see Discussion).

There Is Only One Spectrally Distinct Intermediate. SVD analysis provides a model-free indication of the number of independently absorbing species in a reaction (*31*). Figure 4 shows a plot of a singular value as a function of the number of possible spectral components for the time-resolved absorption spectra collected in the 0−1000 s region and over the wavelength range of 305−750 nm for the refolding reaction of bacteriorhodopsin. As can be seen, the singular value plateaus out at the third spectrally distinct component. This implies that there are only three spectrally distinct species in the refolding reaction of bacteriorhodopsin. They are free retinal, the intermediate, and bacteriorhodopsin. Thus, on the basis of absorbance measurements presented here, the formation of bacteriorhodopsin from the apoprotein in mixed micelles occurs via only one kinetically detectable intermediate.

DISCUSSION

This report describes the use of a simple stopped-flow mixing method in conjunction with photodiode array spectroscopy in demonstrating the obligatory role of an intermediate in the refolding reaction of bacteriorhodopsin. This demonstration is based upon the observation of a lag phase of several seconds in the appearance of the native protein (Figure 1). Thus, an intermediate species populates over the first few seconds in the refolding reaction (phase ν_1), and during this time, the concentration of the native protein remains practically zero. Only upon the decay of this intermediate does the concentration of the native protein begin to increase.

On the basis of observation of a lag phase in the appearance of the native protein, an obligatory intermediate has also been shown to be involved in the refolding of interleukin-1β (*36*) and staphylococcal nuclease (*37*). In these systems, in particular, the presence of a lag in the appearance of the native protein is of paramount importance to the mechanisms of protein folding. The lag seen in the formation of the native state of bacteriorhodopsin results from a second-order reaction between retinal and the apoprotein followed by the Schiff base formation (*25*), which makes the appearance of an intermediate perhaps less surprising. Although intermediates have been detected in a number of other systems (*4*−*7*), whether their role in folding is essential is not clear. The studies on interleukin-1β, staphylococcal nuclease, and bacteriorhodopsin however indicate that the role of intermediates in protein folding may be more than just to act as kinetic traps or to induce off-pathway reactions (*10*, *16*, *17*).

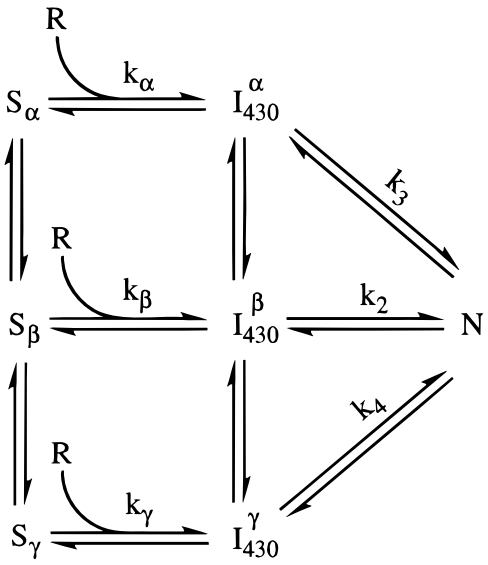
It is also shown here that the obligatory intermediate in the refolding reaction of bacteriorhodopsin absorbs maximally at 430 nm (Figure 3). On the basis of the oscillator strength of the retinal absorption band, this 430 nm intermediate was designated a noncovalent retinal−protein complex (*34*). The fact that the oscillator strength of the retinal absorption band in free retinal and in the 430 nm intermediate

is the same lends little support to the assignment of this intermediate to a noncovalent retinal−protein complex, since the formation of a covalent linkage does not necessarily alter the oscillator strength. In fact, the oscillator strength of free retinal and retinal Schiff base is the same in methylpentane (*38*). It is however unlikely that this would be the case within the apoprotein that contains both charged and polar residues.

The oscillator strength of the retinal absorption band is highly sensitive to electrostatic interactions between the retinal Schiff base and amino acid residues surrounding the retinal binding pocket within the apoprotein (*39*). Thus, if retinal were to be covalently attached to the apoprotein within the 430 nm intermediate, it is unlikely that the oscillator strength of the retinal absorption band in free retinal and in the 430 nm intermediate would be the same. Further, an intermediate with a double-peaked absorbance maximum has also been observed in the regeneration of bacteriorhodopsin from apomembrane (purple membrane minus retinal) and retinal at lower temperatures (*40*). Retinal was shown to be noncovalently bound to the protein within this 430 nm−460 nm intermediate. In fact, the spectral properties of bacteriorhodopsin mutant Lys216Ala resemble this 430 nm−460 nm species (*41*). On the basis of these considerations, it is therefore concluded that the apoprotein is noncovalently bound to retinal within the 430 nm intermediate observed here.

Another interesting feature of the results presented above is that the 430 nm intermediate appears to decay into the native protein via three distinct parallel pathways with lifetimes ranging from several seconds (phase ν_2) to several minutes (phases ν_3 and ν_4) (Figure 2). The fact that an intermediate is obligatory in the refolding of bacteriorhodopsin rules out the possibility that retinal may decay directly into the native chromophore. Thus, the assignment of phases ν_2−ν_4 to the formation of bacteriorhodopsin via parallel routes is justified. Although the decay of the 430 nm intermediate into bacteriorhodopsin via three distinct pathways can in principle account for the kinetic behavior observed here, it is unlikely that this is the case in practice. This is because for the stage involving the decay of the 430 nm intermediate into bacteriorhodopsin most of the population will follow the fastest pathway (phase ν_2), resulting in at most two distinct kinetic phases instead of three. The most likely scenario in which the three distinct kinetic phases may be observed for the decay of the intermediate into bacteriorhodopsin is one in which the intermediate transiently exists in three kinetically distinct populations, and if these are in slow exchange with each other (relative to phase ν_2), then their decay into bacteriorhodopsin via kinetically distinct pathways can be experimentally observed.

The foregoing argument implies that three distinct populations of the apoprotein in mixed micelles fold along separate pathways into three distinct populations of the 430 nm intermediate, overriding the fact that this process can only be experimentally observed as a single kinetic event (phase ν_1). The three distinct populations of the apoprotein in mixed micelles can also be considered to be in slow exchange with each other (relative to phase ν_1). A plausible kinetic scheme that accommodates the above observations then follows as



in which S_α, S_β, and S_γ are the three substates within the apoprotein in mixed micelles (the starting state), I^α_{430}, I^β_{430}, and I^γ_{430} are the three corresponding substates within the 430 nm intermediate (noncovalent apoprotein−retinal complex), R is the retinal chromophore, and N is the final native state of bacteriorhodopsin. The starting state of the apoprotein in mixed micelles is not a denatured or unfolded protein but rather contains substantial secondary structure and, indeed, has been suggested to be an intermediate per se in the refolding of bacteriorhodopsin from a partially denatured state in SDS (*22*, *27*, *34*). It is in fact quite reasonable to suggest that this state of the apoprotein could be analogous to the molten globule observed in the folding studies of water-soluble proteins (*42*).

In the scheme presented above, the steps characterized by the second-order intrinsic rate constants k_α, k_β, and k_γ correspond to the experimentally observed phase ν_1. That is, these steps cannot be experimentally distinguished. The steps characterized by the first-order intrinsic rate constants k_2−k_4 correspond to the experimentally observed phases ν_2−ν_4, respectively. The back intrinsic rate constants have not been included in the above scheme to keep the matters simple (back reactions are considered to be negligible relative to the forward reactions). The vertical arrows in the above scheme indicate the slow exchange of the distinct populations with each other within the apoprotein in mixed micelles and within the 430 nm intermediate and are not under experimental investigation here.

Parallel pathways have also been detected in a number of other systems, including lysozyme (*20*, *43*), ribonuclease A (*5*), cytochrome *c* (*44*), and dihydrofolate reductase (*45*). One possible cause of multiple folding routes in these systems is assigned to the slow cis−trans proline isomerization. The parallel nature of the pathways in bacteriorhodopsin folding observed here may also be a consequence of proline isomerization. There is a total of 11 proline residues in bacteriorhodopsin, three of which are found within the transmembrane helices (*26*). It is thus conceivable that the apoprotein in mixed micelles exists in a mixture of native and non-native proline conformers, which are converted to distinct populations of the 430 nm intermediate via distinct pathways (rate constants k_α, k_β, and k_γ). The population of

the 430 nm intermediate with native proline isomers could thus decay into bacteriorhodopsin via the fastest route (rate constant k_2), while the populations of the 430 nm intermediate that contain proline isomers in a non-native conformation may take longer (rate constants k_3 and k_4).

The classical view of protein folding invokes the role of specific intermediates and a specific pathway (*3*, *46*). In contrast, the new view of protein folding envisages folding as a distribution or ensemble of structures from the unfolded to the native state (*14*, *15*). These two views however do not necessarily have to be mutually exclusive. This report indeed shows that although the 430 nm intermediate is obligatory (consistent with the classical view) it subsequently decays into bacteriorhodopsin via alternative parallel pathways (consistent with the new view). To further bridge the gap between the classical and new views of protein folding, the kinetic mechanism presented above inherently implies that both the apoprotein in mixed micelles and the 430 nm intermediate do not conform to specific structures but rather represent an ensemble of subspecies. In the language of protein folding funnels (*15*), the apoprotein state in mixed micelles and the native state could represent the top (rim) and the bottom of the funnel, respectively, while the 430 nm intermediate could be the many bumps and valleys that the molecules may have to overcome in their quest to reach the native fold. That is, the surface of the folding funnel of bacteriorhodopsin could be quite rough.

To what extent the decay of the 430 nm intermediate into bacteriorhodopsin involves structural changes in the extramembraneous loops and/or the transmembrane regions cannot be assessed at this stage. Protein engineering techniques (*47*) however hold a great deal of promise for investigating the role of the 430 nm intermediate in bacteriorhodopsin folding at a structural level to further our understanding of the protein folding mechanisms.

ACKNOWLEDGMENT

I am sincerely grateful to James Barber and Stephen Matthews for their much needed advice and support. I acknowledge Martin Bell for providing the kinetic analysis software LOOK. I express my deep thanks to the Wellcome Trust for the award of a postdoctoral fellowship. I dedicate this paper to Peter Fryer.

REFERENCES

1. Kelly, J. W. (1996) *Curr. Opin. Struct. Biol. 6*, 11−17.
2. Pain, R. H. (1987) *Trends Biochem. Sci. 12*, 309−312.
3. Levinthal, C. (1968) *J. Chim. Phys. 65*, 44−45.
4. Bycroft, M., Matouschek, A., Kellis, J. T., Serrano, L., and Fersht, A. R. (1990) *Nature 346*, 488−490.
5. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature 335*, 694−699.
6. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) *Nat. Struct. Biol. 3*, 193−205.
7. Roder, H., Elove, G. A., and Englander, S. W. (1988) *Nature 335*, 700−704.
8. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497−526.
9. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10428−10435.
10. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol. 2*, 663−673.
11. Huang, G. S., and Oas, T. G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6878−6882.
12. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M., and Poulsen, F. M. (1995) *Biochemistry 34*, 7217−7224.
13. Sali, A., Shakhnovich, E., and Karplus, M. (1994) *Nature 369*, 248−251.
14. Baldwin, R. L. (1995) *J. Biomol. NMR 5*, 103−109.
15. Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.
16. Creighton, T. E. (1994) *Nat. Struct. Biol. 1*, 135−138.
17. Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994) *Nat. Struct. Biol. 1*, 149−156.
18. Zhang, J.-X., and Goldenberg, D. P. (1993) *Biochemistry 32*, 14075−14081.
19. Creighton, T. E. (1997) *Trends Biochem. Sci. 22*, 6−10.
20. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9029−9033.
21. Booth, P. J., Flitsch, S. L., Stern, L. J., Greenhalgh, D. A., Kim, P. S., and Khorana, H. G. (1995) *Nat. Struct. Biol. 2*, 139−143.
22. Booth, P. J., Farooq, A., and Flitsch, S. L. (1996) *Biochemistry 35*, 5902−5909.
23. Oesterhelt, D., and Schuhmann, L. (1974) *FEBS Lett. 44*, 262−265.
24. Khorana, H. G. (1988) *J. Biol. Chem. 263*, 7439−7442.
25. Bayley, H., Huang, K.-S., Radhakrishnan, R., Ross, A. H., Takagaki, Y., and Khorana, H. G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2225−2229.
26. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M., and Henderson, R. (1996) *J. Mol. Biol. 259*, 393−421.
27. Farooq, A. (1997) Ph.D. Thesis, University of London, London.
28. Oesterhelt, D., and Stoeckenius, W. (1974) *Methods Enzymol. 31*, 667−678.
29. Braiman, M. S., Stern, L. J., Chao, B. H., and Khorana, H. G. (1987) *J. Biol. Chem. 262*, 9271−9276.
30. Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431−441.
31. Henry, E. R., and Hofrichter, J. (1992) *Methods Enzymol. 210*, 129−192.
32. London, E., and Khorana, H. G. (1982) *J. Biol. Chem. 257*, 7003−7011.
33. Riley, M. L., Wallace, B. A., Flitsch, S. L., and Booth, P. J. (1997) *Biochemistry 36*, 192−196.
34. Booth, P. J., and Farooq, A. (1997) *Eur. J. Biochem. 246*, 674−680.
35. Nakanishi, K., Balogh-Nair, V., Arnaboldi, M., Tsujimoto, K., and Honig, B. (1980) *J. Am. Chem. Soc. 102*, 7945−7947.
36. Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol. 4*, 725−731.
37. Walkenhorst, W. F., Green, S. M., and Roder, H. (1997) *Biochemistry 36*, 5795−5805.
38. Schaffer, A. M., Waddell, W. H., and Becker, R. S. (1974) *J. Am. Chem. Soc. 96*, 2063−2068.
39. Birge, R. R., Murray, L. P., Zidovetzki, R., and Knapp, H. M. (1987) *J. Am. Chem. Soc. 109*, 2090−2101.
40. Schreckenbach, T., Walckhoff, B., and Oesterhelt, D. (1978) *Biochemistry 17*, 5353−5359.
41. Schweiger, U., Tittor, J., and Oesterhelt, D. (1994) *Biochemistry 33*, 535−541.
42. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20−24.
43. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature 358*, 302−307.
44. Elove, G. A., Bhuyan, A. K., and Roder, H. (1994) *Biochemistry 33*, 6925−6935.
45. Jennings, P. A., Finn, B. E., Jones, B. E., and Matthews, C. R. (1993) *Biochemistry 32*, 3783−3789.
46. Dill, K. A. (1985) *Biochemistry 24*, 1501−1509.
47. Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) *Nature 346*, 440−445.

BI981485V