pH Dependence of the Photoactive Yellow Protein Photocycle Recovery Reaction Reveals a New Late Photocycle Intermediate with a Deprotonated Chromophore*

Johnny Hendriks and Klaas J. Hellingwerf

From the Laboratory for Microbiology, Swammerdam Institute for Life Sciences and Netherlands Institute for Systems Biology, University of Amsterdam, 1018 WV Amsterdam, The Netherlands

The recovery reaction of the signaling state of photoactive yellow protein includes the following: (i) deprotonation of the p-coumaryl chromophore, (ii) refolding of the protein, and (iii) chromophore re-isomerization from the cis to the trans configuration. Through analysis of the pH dependence of this recovery reaction, we were able to provide proof for the existence of an additional photocycle intermediate. The spectral similarity between this new intermediate and the dark state indicates that the new intermediate has a deprotonated chromophore, which may facilitate chromophore re-isomerization. This spectral similarity also explains why this new intermediate has not been noticed in earlier studies. For our data analysis we introduce a photocycle model that takes into account the effect of the specific light regime selected, a model that was also used for simulations.

The photoactive yellow protein (PYP) from Halorhodospira halophila is a small, water-soluble photoreceptor protein. It makes use of a relatively simple chromophore, p-coumaric acid, which is bound to Cys-69 via a thiol ester bond (1, 2). Favorable in vitro characteristics have made this protein a popular model system for both experimental and computational studies of enzyme/protein structure and function (3). High resolution structures are available, not only of the PYP ground state (4–6) but also of several of its photocycle intermediates (7, 8) and mutants (9, 10). In fact, at least 46 PYP structures have already been deposited in the Protein Data Bank. The PYP photocycle has been intensively investigated in all relevant time domains (11–13). These studies have led to many different, and often conflicting, models for this photocycle, with a widely divergent nomenclature for the transient intermediates involved. Nevertheless, a consensus about the main intermediate states is emerging, the essence of which is summarized in Fig. 1. The ground state of PYP, pG (also named P and PYP), in which the chromophore is deprotonated and has a trans configuration, can absorb a blue photon. From the resulting excited state, a very short living ground state intermediate is formed, which relaxes directly back to pG (13). The excited state further results in the formation of the ultra-fast intermediate, Im, in which the chromophore has undergone trans to cis isomerization. Next, the pR intermediate (also named I1, and PYPt) is formed on a nanosecond time scale. Transient grating and time-resolved optical rotatory dispersion measurements revealed that further structural relaxation occurs in pR (14–16), which is accompanied by a subtle change in the absorption spectrum (12, 16). Accordingly, pR has been split up into pR1 and pR2 in Fig. 1. Next, the chromophore is protonated, thereby forming the intermediate pB’ (17) (also named I3), which is in equilibrium with pR (12). From pB’, pB (also named I1, I2’, and PYPm) is formed in which the protein structure accommodates the new protonation state of the chromophore. The exact dynamics of the transition between pB’ and pB is still under debate. Both an equilibrium and a unidirectional reaction between pB’ and pB have been proposed. We favor the interpretation of a unidirectional reaction. An equilibrium would result in the presence of pB’ throughout the recovery reaction, which is not supported by this study. Furthermore, the pB intermediate has been shown to have absorption characteristics that depend on conditions such as pH, to the extent that it can be highly similar to pB’. Based on available infrared difference spectra of both pB’ (17) and a pB intermediate with UV/visible absorption characteristics similar to pB’ (18), we conclude that pB’ and this pB are distinct intermediates. In this study, we will assume a unidirectional reaction between pB’ and pB. In the next step of the photocycle, pB recovers to the ground state pG. In this recovery reaction several alterations in the configuration of PYP and its coumaryl chromophore have to be accomplished as follows: (i) deprotonation of the chromophore; (ii) chromophore re-isomerization from the cis to trans configuration, and (iii) refolding of the polypeptide chain to the pG configuration. Here deprotonation of the chromophore may facilitate chromophore re-isomerization (12, 19).

In this study we focus on the recovery reaction of the PYP ground state (pG) from the putative signaling state pB. Measurement of PYP recovery kinetics has become a standard tool to determine the effect of site-directed mutagenesis (11, 20) and/or of changes in its mesoscopic context (17, 21) on protein function. However, the way in which this information has been obtained in various studies is far from standardized. Besides the obvious differences in used equipment, also the applied analysis
Here we report on the pH dependence of pG recovery, applying a new approach to analyze this reaction. In doing so we were able to positively identify the actual photocycle intermediate that catalyzes recovery to the ground state. Furthermore, we show the advantages and limitations of the use of photocycle models that incorporate the characteristics of the illumination conditions.

**EXPERIMENTAL PROCEDURES**

**Materials**—Recombinant wild type and H108F apoPYP were produced heterologously in *Escherichia coli*, as described previously (26). The apoPYP was reconstituted with a derivative of p-coumaric acid as described previously (27). The genetically introduced hexa-histidine containing the N-terminal tag was not removed for these experiments. The purity index ($A_{278}/A_{446}$) of the samples was <0.5.

**Sample/Buffer Preparation**—Buffer solutions with defined ionic strength were prepared by mixing 1 M acetic acid, 0.2 M Na$_2$HPO$_4$, 1 M Trizma® base, 0.5 M boric acid, 4 M NaCl, 1 M HCl, and 1 M NaOH. The required amount of these stock solutions was calculated (see supplemental material for details). Minor pH adjustments were made with either 1 M HCl or NaOH as necessary.

**Transient ms/s UV/Visible Spectroscopy**—UV/visible spectra were recorded using an HP8453 UV/visible spectrophotometer. Time-resolved spectra were recorded from 250 to 600 nm, using an integration time of 0.1 s. A “Kraayenhof” cuvette (28) was used to thermostat the sample at 25 °C and allowed measurement of the sample pH in the measurement setup. A Schott KL1500 LCD lamp was used to illuminate the sample. Actual sample illumination was computer-controlled via a shutter system, allowing synchronization of sample illumination with the spectroscopic measurement. A 2-s delay was used between the start of the measurement and the 5-s illumination period of the sample.

**Light Titration of H108F PYP**—Transient ms/s UV/visible spectroscopy was performed on a sample containing H108F PYP in 20 mM Tris/HCl with NaCl of pH 8 and an ionic strength of 0.25 M. The sample was illuminated with varying light intensities by changing settings on the Schott KL1500 LCD lamp. To minimize the effect of probe light, a 17.3% glass gray filter was placed between the probe lamp and the sample. The sample was illuminated until a light-induced steady state was reached. For the majority of the tested light intensities, the computer-controlled shutter system was used for sample illumination. However, to reach the higher light intensities, the shutter system was removed, and illumination was regulated by manually inserting/removing the light guide in/from the light source. See under “Lamp Characterization” for a description of how light intensities were determined.

**pH Titration of PYP Ground State**—The same spectrophotometer setup was used as described above. A buffer containing 5 mM each of acetic acid, phosphate, Trizma® base, and boric acid with an ionic strength of ~0.25 M was used. To a PYP solution of pH ~ 7.3 small aliquots of either NaOH or HCl was added repeatedly. After each addition, the pH and spectrum of the sample were recorded. Especially at extreme pH values, measurements were performed as quickly as possible to mini-
mize the effect of chromophore dissociation. Also, sample illumination was kept at a minimum to minimize light-induced effects at the pH extremes.

**Lamp Characterization**—Light intensities were measured with a LI-COR LI-250 light meter in combination with a LI-COR Quantum sensor, which measures light intensity from 400 to 700 nm. The lamp spectrum of the HP8453 UV/visible spectrophotometer was measured using the diagnostics mode of the spectrometer. The lamp spectrum of the Schott KL1500 LCD lamp was measured with an Ocean Optics USB4000 miniature fiber optic spectrometer. Light intensities for wavelength ranges other than 400–700 nm were extrapolated based on the measured lamp spectra.

**Data Analysis**—Analysis of two-dimensional datasets, such as the pH dependence of rate constants, was performed in OriginPro 8 (OriginLab Corp.). Three-dimensional datasets, such as time-resolved spectra and pH-dependent spectra, were analyzed in MatLab (The MathWorks™), using techniques described elsewhere (29).

Prior to analysis of time-resolved spectra, each spectrum in the series was corrected by subtracting the average absorbance of the steady state fractions of pB and pG (see under “Results”). In this way, the recovery rate is linked to the steady state fractions as described by Equation 1.

$$f_{pB} = \frac{f_{pB,ss} \cdot (k_1 \cdot e^{-k_1 \cdot (t - t_0)} + (1 - f_{pB,ss}) \cdot e^{-k_2 \cdot (t - t_0)})}{f_{pB,ss}} \quad \text{(Eq. 1)}$$

Here $f_{pB}$ and $f_{pG}$ are the calculated fractions of pB and pG, respectively; $f_{pB,ss}$ is the assumed steady state fraction of pB; $k_1$ and $k_2$ are rate constants; $f_{pB,ss}$ equals the fraction that the exponent with rate constant $k_1$ contributes; $t_0$ is the time at which the excitation light is turned off. Essentially, the model describes a bi-exponential recovery from pB to pG. As the value for $f_{pB,ss}$ has no influence on the quality of the fit nor on the obtained rate constants (see “Results”), this parameter was set to 0.5 for all analyses.

From the above analysis a pG spectrum and a pG/pB mixed spectrum was obtained. The mixed spectrum ($S_{mix}$) is related to the spectra for pG and pB ($S_{pG}$ and $S_{pB}$) according to Equation 2.

$$S_{mix} = \frac{(f_{pB,ss} - f_{pB,ss} \cdot S_{pG} + f_{pB,ss} \cdot S_{pB})}{f_{pB,ss}} \quad \text{(Eq. 2)}$$

Here, $f_{pB,ss}$ is the actual steady state fraction of pB, and $f_{pB,ss}$ is the assumed steady state fraction of pB used to obtain $S_{mix}$ (see Equation 1).

An initial estimation for $f_{pB,ss}$ was made, based on a simple photocycle model with a single-step recovery from pB to pG, using $k_1$ as recovery rate. Although this is not the optimal model to describe the data, it is able to provide a reasonable estimation of the steady state fractions of pB and pG (see under “Results”). In this way, the recovery rate is linked to the steady state fractions as described by Equation 3,

$$f_{pB,ss} = \frac{Q_g \cdot I_0 \cdot e \cdot \ln(10)}{Q_y \cdot I_0 \cdot e \cdot \ln(10) + k_{re}} \quad \text{(Eq. 3)}$$

Here $f_{pB,ss}$ is the steady state fraction of pB, and $k_{re}$ is the photocycle recovery rate. $Q_g$, $I_0$, and $e$ are constants describing the photocycle quantum yield (0.35) (30), light intensity for the wavelength range of 300 to 550 nm (2.154·10⁻⁴ mmol·cm⁻²·s⁻¹ = 2154 mmol·cm⁻¹·s⁻¹), and extinction coefficient for the wavelength range 300–550 nm (5179 M⁻¹·cm⁻¹), respectively, as described elsewhere (31). For a detailed derivation, see supplemental material. Note that the extinction coefficient is specific for the lamp and wavelength range used, as it was calculated from the lamp spectrum in combination with the absorption spectrum of the PYP ground state at pH 8, using an $e_{446}$ of 45,500 M⁻¹·cm⁻¹.

For the analysis of the pH dependence of a rate constant, with a single pH transition, use was made of the Henderson-Hasselbalch equation in the form of Equation 4. Here $k_{low,pH}$ and $k_{high,pH}$ are the values of $k$ at low and high pH, respectively.

$$k = \frac{k_{high,pH} - k_{low,pH}}{1 + 10^{(pH - pK)}} + k_{low,pH} \quad \text{(Eq. 4)}$$

For the analysis of multiple pH transitions within a single molecule Equation 5 was used,

$$f_i = \frac{1}{n_t + 1} \sum_{j=1}^{n_t} c_j \quad \text{Equation 5}$$

$i < j \Rightarrow c_j = 10 \sum_{k=1}^{i-1} n_k \cdot (pH - pK_k)$

Here $f_i$ is the fraction of species $i$, where the sum of all fractions is 1; $n_t$ is the number of pH transitions; $i$, $j$ are species number indicators; $k$, $l$ are pH transition number indicators; $pK$ is the $pK$ of a pH transition, where transition 1 has the lowest $pK$, and the $pK$ increases with increasing transition number; $n$ is a cooperativity constant for the pH transition. See supplemental material for buffer calculations for an example on how to use Equation 5.

**RESULTS**

In this study we have focused on the photocycle recovery reaction of PYP, i.e. the dark transition from pB to pG. For this we have recorded spectra from 250 to 600 nm with a time resolution of 100 ms. The broad wavelength range allows us to include the absorption changes of the absorption peak around 280 nm, a part of the spectrum that is generally not included in the analysis of PYP recovery. Because of the rather low time resolution, we ignore formation of pB and focus only on the PYP recovery reaction (see “Discussion”). In most of the recorded time trajectories the sample was...
illuminated for 5 s with actinic light, 2 s after the start of the measurement. This is enough time to reach a light-induced steady state in the sample.

The rate of photocycle recovery is strongly dependent on environmental conditions such as pH (11, 12), temperature (24), ionic strength (23), and buffer composition (21). Here we have opted to vary the pH over a wide range (pH 4.5 to 10.5), as we expect that this will provide us with the most relevant mechanistic information. The other environmental conditions were kept constant as much as possible. All our measurements were therefore performed at 25 °C (±0.1), with an ionic strength of 0.25 M (adjusted with NaCl), and a buffer mixture that contained 20 mM acetic acid, phosphate, Trizma® base, and boric acid.

The data obtained from our measurements was analyzed in three stages. Note that our analysis was performed on absorption spectra and not absorption difference spectra. First, the data were explored to determine the number of kinetically and spectrally distinguishable states that can be reasonably extracted from the data. Next, for each measured pH value, a kinetic analysis was performed. From this analysis also spectral information was extracted. The results of the analyses at the individual pH values were subsequently analyzed to extract pH dependences. To support the interpretation of these measurements, a light titration of the H108F mutant of PYP was performed to demonstrate that Equation 3 can be used for the analysis in stage 2. Also a pH titration of the dark state was performed, which is necessary for normalization of the data in stage 2 of the analysis.

**Analysis Stage 1, Data Exploration**—To determine the number of kinetically and spectrally distinguishable states, the data were explored using singular value decomposition (SVD) (29). With SVD a matrix containing the time-resolved spectra (A) is deconvoluted into three separate matrices such that $A = U SV^T$ (the superscript T indicates the matrix is transposed). Here the $U$ matrix contains columns representing spectral contributions to the A matrix. The $V$ matrix contains columns representing temporal contributions to the A matrix. The $S$ matrix is a diagonal matrix containing scaling factors that indicate the significance of each column in the $U$ and $V$ matrices (columns are sorted with respect to their significance). In Fig. 2 examples of two SVD explorations are shown. Above pH 7, the third component of both the $U$ and $V$ matrix consists mostly of noise (see Fig. 2, B and D, for an example at pH 9.0), which indicates there are only two kinetically and spectrally distinguishable states, i.e. the dark and a light-induced state. Below pH 7, the third component does contain significant structure (see Fig. 2, A and C, for an example at pH 5.5), and the fourth component of both the $U$ and $V$ matrix consists mostly of noise (data not shown). This indicates that there are three kinetically and spectrally distinguishable states below pH 7. To determine the significance of this third state, we took a closer look at the $S$ matrix. For pH 5.5 and 9.0, the logarithm of the scaling values of the first nine ranks from the $S$ matrix are shown in Fig. 2, E and G. It is clear that the scaling value for increasing ranks drops quickly and then stabilizes. To be able to compare the scaling values from the different data sets, the scaling values in each $S$ matrix were normalized by the sum of all scaling values in that $S$ matrix, i.e. effectively expressing the scaling values as fractions. It then becomes clear that the first component of the SVD contributes around 94% to the dataset for pH values above 6.0, and it gradually drops to 86% below this pH value (see Fig. 2F). The third component of the SVD only contributes around 0.04% for pH values above 7.0 and gradually increases to 0.4% below this pH. As such, it is reasonable to assume that this third component can be neglected, which then allows the assumption that only two kinetically/spectrally distinguishable states (i.e. the dark and a light-induced state) are relevant for the entire measured pH range. This is further corroborated by comparing the original data with data reconstructed from the $U$, $V$, and $S$ matrices, using either the first 2 or 3 ranks (see supplemental Fig. S1 through S3). When 3 ranks are used, the difference with the original data results in unstructured noise with a maximum level of 1 mOD units at all measured pH values. Therefore, for subsequent kinetic analyses, datasets reconstructed using the first 3 ranks were used.

Note, when only 2 ranks are used, similar unstructured noise of the same level is obtained above pH 7. Below pH 7 some structure is observed in the noise, the maximum deviation of which gradually increases going to lower pH values. However, deviations are still lower than 3 mOD units at pH 5, whereas noise levels remain at most 1 mOD units (data not shown).

**Analysis Stage 2, Kinetic Analysis**—For the kinetic analysis we make use of the fact that the $A$ matrix, which represents the data, can be split into two matrices such that $A = DF^T$. Here the matrix $D$ contains the spectrum of each relevant species, and the matrix $F$ contains the temporal profile of the contribu-
contribution of each of the species in matrix D. As such, by modeling the F matrix, we can obtain the spectrum of each species via D = A·F⁻¹, where F⁻¹ is the pseudo-inverse of F, and therefore we can also determine a simulated A matrix (29). In short, we fit a temporal profile to our data, from which we obtain both kinetic and spectral information.

From our SVD exploration, we concluded a two-state model should sufficiently describe our kinetic data. In our experiment the recovery can be seen as the transition from a light-induced steady state mixture of pG and pB back to a dark state with only pG present. The simplest way to model the F matrix in this two-state situation is by having the two-states be interconverted via first-order kinetics (i.e., using exponentials). Here the decay of the fraction of pB is mirrored by the growth in the fraction of pG. To obtain a good fit of the data (i.e., minimal structure in the residuals) two exponentials are required (see Equation 1), as was also shown previously (12). The results from this analysis are shown in Fig. 3. For the majority of the measured pH range, the faster of the two exponentials contributes more than 95% to the data. Only at the extremes of pH (pH < 5 and >10), where the two rate constants become more similar, does the contribution of the faster rate constant drop below 95%. Furthermore, the pH dependence of the fast and slow component appears to be slightly different (see Fig. 3). Most notably, the pH at which the highest rate is achieved is slightly shifted toward the lower pH for the slow component. Note that varying the steady state fraction of pB in Equation 1 does not shift the curve e.g. tuned visually based on spectral shape (e.g., pG) or by the high pH pB species. One of the assumptions that is generally made to obtain the pB spectrum is that pB does not absorb above ~420 nm (32, 33). Fig. 4B indicates that this assumption does not hold. Furthermore, this assumption is not applicable at alkaline pH values. Therefore, we adopted another approach. To obtain the true pB spectra at the various pH values from the kinetic analysis described above, we have used Equation 2 to subtract the contribution of pG from the mixed spectrum obtained in the analysis. A simple photomultiplier model was used to estimate f_{pB,ss} (see Equation 3), where the rate constant that contributes most to pG recovery was used for k_{re}. The value for f_{pB,ss} was then fine-tuned visually based on spectral shape (e.g., no negative absorption allowed), shape of the first four ranks of an SVD analysis on the pH dependence of the pB spectra (e.g., curve smoothness), pH dependence of the f_{pB,ss} parameter, and analysis of the pH dependence of the resulting pB spectra (described below). The set of pB spectra obtained after this optimization is shown in Fig. 5A.

**Light Titration of H108F PYP**—To make sure Equation 3 is able to estimate the values for f_{pB,ss} accurately enough, a light titration was performed on a PYP sample, i.e. with the H108F mutant of PYP, as the recovery rate of this mutant protein at pH 8 is slow enough to allow for a nearly complete titration, from 0 to 100% bleach of the sample, with the light source available to us. The recovery rate at each measured light intensity was determined as described above. Again two exponentials were required to fit the data. The major recovery component, which contributes 94% and was the same in all measurements (as...
**PH Dependence of PYP Recovery**

**FIGURE 5. PH dependence of the spectrum of pB. A**. pB spectrum obtained for each pH. B, pH dependence of the fractional contribution of each pB species to the spectra in A. C, deconvoluted spectra for the low pH (dotted), medium pH (solid), and high pH (dashed) pB species. Also shown is the spectrum of pG at pH 8 (thin solid line). Note that the pB spectrum at medium pH contains a contribution of the pG intermediate. The inset in C shows the model that accounts for the observed pH dependence of the pB spectrum, from which only pK_{a1} (5.9 with an n of 0.9) and pK_{a2} (9.9 with an n of 1.2) could be extracted from the spectra depicted in A.

**FIGURE 6. LIGHT TITRATION OF PYP H108F.** The fraction of 446 nm absorption remaining in the light-induced steady state (relative to the dark state) is plotted against the light intensity used (in the 300–550 nm range). Measurements performed with illumination via the shutter system (see "Experimental Procedures") are indicated by circles. Measurements performed by manually inserting/removing a light guide in/from the light source are indicated by triangles. The fitted line, Equation 3 was used with a k_e = 0.35 s^{-1} (average of experiments with I_0 between 38 and 14,000 μmol·m^{-2}·s^{-1}); Q_p = 0.35; ε_{446} = dependent on lamp setting as indicated by number above symbol in plot (setting f = 4021; 2 = 4252; 3 = 4537; 4 = 4742; 5 = 5089; 6 = 5236 μmol·cm^{-1}·mol^{-1}). To fit the f_{ap} profile obtained with Equation 3, which varies between 0 and 1, it needed to be scaled such that it varies between 1.00 and 0.03.

The expected, was found to be 0.35 s\(^{-1}\). In Fig. 6, the fraction of the absorption at 446 nm that is left after reaching the light-induced steady state is plotted as a function of light intensity. The fitted line was made by scaling Equation 3, with all parameters fixed to measured values (Q_p = 0.35; k_{pe} = 0.35 s\(^{-1}\); ε dependent on lamp setting). Except for the four highest light intensities in the plot, all data points were obtained using the same shutter setup to illuminate the sample, as the one that was also used for the pH titration. To obtain the four highest light intensities, the shutter had to be removed. At these four light intensities the sample temperature increased significantly, which may explain their sub-optimal fit to the curve. Nonetheless, it is clear that Equation 3 is able to provide a reasonable estimate of f_{apB,5236} especially under the conditions used for the pH titration.

**Analysis Stage 3, PH Dependence of the pB Spectrum**—From the pH dependence of the pB spectra we were able to confidently extract two pH transitions, resulting in three distinct spectral species for pB. These are shown in Fig. 5, B and C. Note that the species dominant at neutral pH appear to be a combination of two separate species, as indicated by its bi-lobed absorption spectrum. These two species apparently cannot be separated based on pH, at least not in our experimental approach. As noted above, the species with an absorption maximum around 450 nm cannot be explained either by pG or by high pH pB.

**pH Titrations of pG**—Note that small variations in the amount of protein in the samples (caused by pipetting errors) made it necessary to normalize the spectra obtained from the kinetic analysis above. We selected the absorption of pG at 446 nm to normalize the spectra that were used to determine the true pB spectra. As the spectrum of pG also varies with pH, the pH dependence of the UV/visible absorption spectrum of the dark state of PYP (i.e. pG) was determined for the pH range from pH 0.7 to 13.2. A more thorough analysis of this dataset, using the complete spectra, showed that seven pH transitions were required to describe all pH-induced spectral transitions for this pH range (see Fig. 7). At low pH, transitions with a pK_{a1} of 2.0, 3.1, and 4.1 were found. The corresponding spectra in Fig. 7A reveal that the transition with a pK_{a1} of 4.1 is accompanied by a minor decrease in absorption around 446 nm and a slight increase around 350 nm. The transition with a pK_{a2} of 3.1 results in a pB_{dark} species that significantly absorbs around 450 nm. This latter contribution is absent in the most acidic species, which is formed with a pK_{a1} of 2.0. At high pH, transitions with a pK_{a1} of 10.0, 11.5, and 12.2 were found, as well as an apparent pK_{a2} of 13.0 for chromophore dissociation (because of the hydrolysis of the thiol ester linkage (34)). Note that although we could fit the transition associated with chromophore dissociation as a pH transition, it actually is the result of hydrolysis of the thiol ester bond, which occurs with a rate that is pH-dependent. The corresponding spectra in Fig. 7C reveal that the transition with a pK_{a1} of 10.0 is accompanied by a minor change in the shape of the peak at 446 nm and a broadening of the peak around 280 nm that is typical for deprotonated tyrosine residues (34, 35). The transition with a pK_{a1} of 11.5 results in a blue shift of the main absorption peak to 427 nm. The transition with a pK_{a1} of 12.2 results in an additional blue shift of the main absorption peak to 409 nm. Note that for the pH range that we used for the kinetic analysis (pH 4.5–10.5), only three of these seven PY ground state species contribute significantly (see Fig. 7B). Furthermore, the transition with a pK_{a1} of 4.1 appears to coincide with the appearance of the third component in the SVD exploration below pH 7 (data not shown), indicating that
these two characteristics are possibly related. Also note that the pH dependences in Fig. 3 and Fig. 5B do not coincide with the pH dependence observed in the SVD exploration.

DISCUSSION

The photocycle of PYP has been studied extensively, especially with respect to signaling state formation. This has led to detailed models of the events that lead to the formation of the pB intermediate, which involves chromophore isomerization, proton transfer from the chromophore to Glu-46, and conformational transitions of the protein. However, with respect to the dark recovery reaction of the ground state, much less is known. It has been suggested that this reaction occurs via an intermediate with a specific fold, and an absorption spectrum that is similar to that of the ground state (12), but no solid proof could be provided at that point. Here, we provide this proof by applying an improved data analysis method on an extensive kinetic dataset that covers both a large pH (4.5–10.5) and wavelength (250 – 600 nm) range.

Photocycle Models—Photocycle models are often used to analyze light-induced events in PYP, but rarely do these models explicitly take into account the illumination conditions of the experiment. For this study we have derived a photocycle model that does take this aspect into account (see supplemental material). Initial analyses of the data using a basic model with a one-step recovery step from pB to pG resulted in fits comparable in quality to mono-exponential analyses of the recovery reaction. This is not surprising as mono-exponential recovery is inherent to the basic model. However, the major advantage of using the model is that not only recovery from the light-induced steady state can be fitted but also the formation of the steady state pB/pG mixture. Furthermore, the use of this type of model has the added advantage that it is straightforward to incorporate the effect of sample illumination by the probe lamp, which becomes significant at the pH extremes where photocycle recovery is very slow. Use of this basic model is, however, not sufficient as it has already been shown that mono-exponential decay is not sufficient to accurately describe the data (12).

Even though the use of more complicated photocycle models was initially promising, a few complications emerged. Our analysis of the recovery reaction thus far indicates that two states are involved with a bi-exponential transition. To explain such behavior of the recovery reaction with any photocycle model, a third state needs to be introduced (apart from the electronically excited state of pG). However, spectrally this species could not be distinguished from either pB or pG. Without a clear spectral constraint for this additional state, a dependence between several of the rate constants for the models is introduced that does not allow for a confident estimation of these rate constants. The assumption that the additional species has the same spectrum as either pB or pG also did not improve this situation.

Another complication is that during the experiment a small amount of sample is lost. Most likely this is because of aggregation, as spectra of the sample between the start and end of the measurement show a small increase in scattering (data not shown). Our model does not take this process into account, which negatively influences analysis of the data using all extended photocycle models we tested. We therefore decided to analyze our data with a bi-exponential analysis of the photocycle recovery reaction from a steady state situation, as described under “Results.” This latter method is less sensitive to measurement artifacts, such as aggregation, although it still allows us to extract information relevant to our study.

Two-state Model Analysis—Basically our data confirm previous measurements on PYP; however, our data analysis approach allowed us to get more detailed information from the dataset. As explained above, initially it was our intention to use photocycle models that incorporate illumination conditions of both the probe and actinic light to analyze our data. The idea was that such a model would act as a constraint/assumption that would allow accurate determination of the absorption spectra of photocycle intermediates. Furthermore it would describe artifacts introduced by the probe light. However, we soon found that the complexity of the models needed made such an approach impractical. In part this is caused by the presence of several fast equilibria between spectrally distinct species in the photocycle. This makes it impossible to distinguish between these species, based on kinetic information alone. Therefore, we limited our analysis to dark-state recovery from a light-induced steady state. Furthermore, based on SVD exploration of the data, we assumed that only two kinetically and spectrally distinguishable states are present. The following question then arises. To what extent are photocycle intermediates other than the dark and the signaling state present in a light-induced steady state? To get an answer to this question, we simulated the experiment using a photocycle model that, besides pG and pB, includes the pR and pB’ intermediate. pH-dependent rate constants were taken from our previous work (12). We compared these results with a simulation of a two-state model that only has the pG and pB species. The simulations revealed that neither pR nor pB’ accumulates significantly under our measurement conditions. Table 1 lists the simulated steady state amounts of pR and pB’ for the illumination condi-
tions used in our experiment. Only at the extremes of pH enough of either pR or pB′ may accumulate for them to show up in our data. Temporal comparison of the two models shows them to be as good as identical within the time resolution of our experiment (data not shown). Only at the pH extremes, where pR or pB′ start to accumulate in higher amounts, do significant differences become visible. Furthermore, when illumination of the sample is ended, the small amounts of pR and pB′ that do accumulate decay instantly with respect to the time resolution of our measurement (data not shown). These simulations further validate our assumption of a two-state situation for our measurements. The final validation for the two-state assumption is the analysis itself. Recovery from the light-induced steady state could be described by just two spectral species. However, at the pH extremes the quality of the fits was somewhat diminished (see supplemental Fig. S4 through S6). With all the signs pointing to a two-state model for the recovery reaction, it is rather surprising that two exponents were required to satisfactorily describe the transition from the light-induced steady state to the dark state. One would expect mono-exponential behavior from a two-state model. This suggests that most likely a third state is present that cannot be distinguished spectrally from the other two states, presumably because it differs from the other states only in protein conformation. The equilibrium between these structural forms should then be relatively slow to observe the bi-exponential behavior of the recovery. In fact, in a recent publication it was suggested that the bi-exponential behavior of the recovery is caused by trans/cis isomerization of Pro-54 (36), i.e. presumably a spectrally silent transition between different structural states. Based on the slight difference in the pH dependence of the two obtained rate constants (see Fig. 3), one can conclude that the equilibrium between such structural forms is pH-dependent.

Obtaining the pB Spectrum—The spectral information obtained from our two-state analysis consisted of a spectrum for pG and a mixed spectrum of pB and pG. To obtain the pB spectrum from the mixed spectrum requires a good estimation of the amount of pB present in the light-induced steady state. Typically, one would have used spectral constraints to achieve a good estimation. As we were not able to use any spectral constraints other than that the spectrum cannot have a negative component, we had to use other constraints. Here we have made use of a photocycle model again. As the two-state model is able to describe the data reasonably well, we derived an equation from it for the light-induced steady state fraction of pB (Equation 3). This fraction is dependent on the illumination conditions, which we measured, and on the recovery rate. Of the two rates we obtained from our analysis, the dominant one contributes more than 95% to the signal for most pH values. Therefore, we used this dominant rate to estimate the amount of pB in the light-induced steady state. The obtained values were subsequently fine-tuned visually, based on spectral and pH dependence characteristics. We are aware that this is a weak point in the analysis, as it introduces a human bias into the analysis. However, typical changes in the calculated fraction of pB were less than 0.01; they never exceeded 0.04 (see also supplemental Fig. S7). As expected, the largest corrections were necessary at the extremes of pH. Furthermore, a test with the PYP mutant H108F, in which the amount of light was varied, showed that using Equation 3 allows one to make a reasonable estimate of the amount of pB in the light-induced steady state (see Fig. 6).

Identification of a New Intermediate (pG′) —From previous work it was already clear that the pB spectrum is pH-dependent. From our data we were able to separate three spectral species based on pH (see Fig. 5). Basically, these results confirm previous work (18, 25). At low pH pB has an absorption maximum of 368 nm, at medium pH an absorption maximum of 355 nm, and at high pH an absorption maximum of 430 nm, with corresponding $K_a$ values at 5.9 and 9.9, respectively. However, the spectral species at medium pH appears to be a mixture of at least two species, with distinct absorption maxima at 355 and 450 nm. The peak at 450 nm has not been observed before and can be explained by neither the pG nor the high pH pB species. It must therefore be a distinct species related to the 355 nm pB species. The goal of our experiment was to find the intermediate through which recovery to pG occurs. We find it likely that the 450 nm species represents this intermediate. The suggested characteristics of this intermediate, a specific fold and an absorption spectrum that is likely very similar to that of the ground state (12), seem to fit. As this intermediate is spectrally similar to pG, and also is a precursor to pG, we have decided to name it pG′. Note that in a previous publication (12) we suggested the rather confusing name pBdep, which we will no
longer use. As the function of pG’ is to catalyze the recovery reaction, it would be logical for the pH dependence of PYP recovery to be governed by the relative presence of this intermediate. We were therefore surprised that we could not separate pG’ from the related pB intermediate on the basis of pH dependence. The only explanation can be that pG’ and the 355 nm pB species are involved in a pH-independent equilibrium, as is illustrated in the inset of Fig. 5C. Furthermore, if pG’ indeed catalyzes the recovery reaction, it is likely that the rate of recovery via pG’ is much faster than the rate of direct recovery via any one of the pB species. The pH dependence of the rate of recovery via pG’ should therefore become visible by division of the most prominent rate constant obtained with the bi-exponential fit (see Fig. 3, circles) by the fraction that the pG’-containing pB spectrum contributes (see Fig. 5B, solid line). Indeed such a treatment results in a curve with a clear, single pH transition. Analysis with Equation 4 revealed a pK_a of 7.07 (see Fig. 8). Deviations at the pH extremes are likely because of the low abundance of pG’ at these pH extremes, allowing direct recovery from pB to become significant. The result further shows that at low pH, recovery via pG’ is more efficient with an apparent rate constant of 2.70 s⁻¹ compared with high pH for which an apparent rate constant of 1.59 s⁻¹ is found.

Based on the analysis so far, it would seem that a multitude of pB species are in a fast equilibrium of which there is one, pG’, through which accelerated recovery of pG is achieved. If this hypothesis would completely describe the photocycle events, our analysis should only have needed a single exponent to describe the recovery from pB to pG. However, it turns out that a minor second exponent is required for a proper fit of the data. One might explain this via the equilibrium reaction of pG’ with pB. However, if so, it should be possible to distinguish between pG’ and the other pB species based on kinetics, which is not the case. Still the spectral similarity of pG and pG’ may be a complicating factor here. However, we feel that the explanation given above is more likely, i.e. that the bi-exponential behavior is caused by a spectrally silent difference in structure that is involved in a relatively slow equilibrium.

If the amount of pG’ that is present indeed governs the observed recovery rate, then logic would dictate that this should be reflected in the pB spectrum of the H108F mutant used in the light titration test. At pH 8 the recovery reaction of H108F is slower than that of wild type PYP. When we compare the ratio between the absorption at 358 and 446 nm for wild type (2.9) and H108F (9.4), we indeed see that the amount of pG’ is significantly less in H108F in comparison with wild type PYP (see supplemental Fig. S14).

pH Dependence of pG—The minor third component, detected below pH 7 by the SVD exploration (see Fig. 2), might be explained by heterogeneity in pG. The pH titration on the dark state of PYP revealed this heterogeneous nature of pG. It appears that one of the pH transitions (with a pK_a of 4.1) between different ground state species has a pH dependence, comparable with the contribution of the third component in the SVD exploration. The pH dependence of the dark state of PYP further revealed that formation of pBdark appears to have two pH transitions. One, with a pK_a of 3.1, results in a pBdark species that also has a small contribution of a species that absorbs around 452 nm. A pH-independent equilibrium between these two species could explain this. The second transition with a pK_a of 2.0 then results in a pure pBdark spectrum. Furthermore, the pH titration of the dark state of PYP revealed two transitions at alkaline pH, with pK_a values of 11.5 and 12.2, each of which results in a species with a distinct absorption maximum. The species with an absorption maximum of 409 nm is most likely an (partially) unfolded PYP (34). Also, an additional transition at alkaline pH with a pK_a of 10.0 was found. Unlike the former transitions that were accompanied by a major change of the absorption spectrum (from 446 to 427 to 409 nm), the transition with a pK_a of 10.0 is more subtle; it causes only a small change in the main absorption peak. The pK_a of 10.0 suggests that the transition is because of deprotonation of tyrosine residues. This is corroborated by a broadening of the absorption peak around 280 nm, typical for the presence of tyrosinates (34, 35). It is therefore likely that a similar transition also exists for pB. However, for pB we observed a pK_a of 9.9, which is accompanied by a major spectral change. A separate transition due to deprotonation of tyrosine residues, like for the dark state of PYP, may therefore not be visible for pB. The quality of the pH-dependent fit of the pB spectra was relatively poor at very alkaline pH. It is therefore likely that an additional high pH pB species exists. However, as we do not have spectral data for pB above pH 10.5, it was not possible to confidently determine its characteristics.

PYP Photocycle Tree—Based on this study we have built a model for the photocycle of PYP (see Fig. 9) that includes all the different pH transitions described above. At first sight, this photocycle model looks rather complex. With a total of 18 different species, not including electronically excited states, naming all these species starts to become a problem. Therefore, we have grouped these species into photocycle functional groups. Essentially, we stay with the naming convention that we have used before, only now names such as pG, pR, and pB represent a group of species instead of a single species. This is meaningful as proteins are very dynamic, both structurally and with regard to the protonation state of their amino acid residues. Three basic states of PYP can be discriminated, i.e. pG, pR, and pB (also known in other nomenclatures as P, I₁, and I₂ (or I₂’)) or
The pR state is the dark state of PYP. Absorption of blue light by this state can lead to the formation of the pR state. Upon formation of pR, the chromophore is isomerized from trans to cis, while maintaining a structural state highly similar to that of pG. Adaptation of the protein to the new isomerization state of the chromophore results in the formation of pB. Subsequently, pG is recovered from pB, completing/closing the photocycle. Besides these basic states of PYP, several others need to be defined as well. The pB-dark state can be considered a chemically/pH-induced nonfunctional state of PYP. Although this state may be spectrally similar to pB, the chromophore is still in the trans state. As the protein structure is perturbed in this state, the protein has become nonfunctional. Although this state may be irrelevant biologically, it has to be taken into account in certain in vitro experiments. Though generally pB-dark is considered the low pH form of pG, the high pH form of pG is a pB-dark form as well. At very low and very high pH values one also has to take into account chromophore hydrolysis (34, 37).

In its simplest form, the photocycle of PYP has three states as follows: pG, pR, and pB. In between these basic states of PYP, additional intermediates can be detected. In Fig. 9 only those intermediates have been depicted that occur in the pR to pB and in the pB to pG transition. Interestingly, these additional intermediates, i.e. pB' and pG', only seem to occur in the biologically relevant path of the photocycle (indicated by boxes with thick lines in Fig. 9). The pB' state is characterized by the protonation of the chromophore by the carboxylic acid group of Glu-46 (17). This protonation is reversible, but pB' can also be converted to pB (12). At low pH the chromophore may well be protonated directly via the solvent, thereby bypassing pB'.

Note that in the model the reaction between pB' and pB is unidirectional. It has been suggested that this reaction actually is an equilibrium. In that case we should have observed pB'-like absorption characteristics mixed with our pB spectra. This is not the case, except for the low pH pB form, which is spectrally similar to pB'. Therefore, if
based on either kinetics, pH, or both, at least for the conditions used in this study. Combined with the apparent spectral similarity of pG' to pG, it is easy to see how the pG' intermediate remained undetected for so long. Recovery via pG' appears to have a pH dependence with only a single pH transition (see Fig. 8). It is unclear what causes the pK_a of 7.07 for recovery via pG'. The more complex pH dependence of the PYP recovery rate (see Fig. 3) can be explained by the fact that the amount of pG' is pH-dependent as well. Direct recovery of pB to pG seems to be possible as well, although it is much slower than recovery via pG'. Direct recovery from pB to pG is only significant at the extremes of pH, where the amount of pG' is minimal.

The absorption characteristics of pG' not only indicate that the chromophore is deprotonated but also that this deprotonated chromophore has specific interactions with the protein. Interestingly, a recent molecular dynamics study on PYP indicates that helix formation is a dynamic bottleneck in the recovery reaction of PYP (38). It might be that this helix formation is one of the key elements involved in the equilibrium between pB and pG'.

**Concluding Remarks**—The model depicted in Fig. 9 can easily be extended to take into account other aspects of PYP. Here we have limited ourselves to the following: the isomerization state of the chromophore, the protonation state of the chromophore, Glu-46 and tyrosine residues, the folding state of the protein, and its visible absorption maximum. Other aspects that could be considered are as follows: the protonation state of other amino acid residues, the presence/absence of salt bridges (23), and the secondary structure of a specific part of the protein (38). Furthermore, we have limited ourselves to pathways induced by blue light excitation of pG. Other paths such as light-induced branching reactions from pR or pB, for example, or illumination of pBdark species can of course also be added. One thing is very clear from this model, whenever PYP is used for experiments chances are that mixtures of intermediates are formed that cannot always be distinguished based on kinetic information alone. On the practical side, the complexity of the model makes it very difficult to apply it directly to the analysis of datasets. However, it is a powerful tool that can help with the choice of experimental conditions. Furthermore, it can serve as a guide for the analysis of data.

The use of a light-induced steady state pB/pG mixture has enabled us to identify/visualize the photocycle intermediate pG'. Although this approach can be used in many other cases as well, it does have its limitations. In order for a significant amount of signaling state to accumulate, the recovery rate has to be relatively slow. For example, for the E46Q mutant of PYP, significant amounts of pB cannot be accumulated in part of the relevant pH range (25). The advantage of using a light-induced steady state is that it is fairly easy to estimate the amount of signaling state that is accumulated in that steady state, using measurable parameters such as light intensity and recovery rate. This is valuable information that can be used in the data analysis process. In our case it was a key part of the identification of pG'. This intermediate may play a crucial role in further studies of thermal chromophore isomerization in PYP.

**Acknowledgments**—We thank Dr. I. H. M. van Stokkum and Dr. T. Gensch for providing valuable feedback regarding the manuscript.

**REFERENCES**

1. Baca, M., Borgstahl, G. E., Boissinot, M., Burke, P. M., Williams, D. R., Slater, K. A., and Getzoff, E. D. (1994) *Biochemistry* **33**, 14369–14377.
2. Hoff, W. D., Dux, P., Hard, K., Devreeze, B., Nugteren-Roodzant, I. M., Crielaard, W., Boelens, R., Kaptein, R., van Beeumen, J., and Hellingwerf, K. J. (1994) *Biochemistry* **33**, 13959–13962.
3. Hellingwerf, K. J., Hendriks, J., and Gensch, T. (2003) *J. Phys. Chem. A* **107**, 1082–1094.
4. Borgstahl, G. E., Williams, D. R., and Getzoff, E. D. (1995) *Biochemistry* **34**, 6278–6287.
5. Dux, P., Ruibinestenn, G., Vuister, G. W., Boelens, R., Mulder, F. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J., and Kaptein, R. (1998) *Biochemistry* **37**, 12689–12699.
6. Getzoff, E. D., Gutwin, K. N., and Genick, U. K. (2003) *Nat. Struct. Biol.* **10**, 663–668.
7. Ihee, H., Rajagopal, S., Srajer, V., Pahl, R., Anderson, S., Schotte, F., Anfinrud, P. A., Wulf, M., and Moffat, K. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7145–7150.
8. Schmidt, V., Pahl, R., Srajer, V., Anderson, S., Ren, Z., Ihee, H., Rajagopal, S., and Moffat, K. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4799–4804.
9. Anderson, S., Srajer, V., Pahl, R., Rajagopal, S., Schotte, F., Anfinrud, P., Wulf, M., and Moffat, K. (2004) *Structure (Loud.)* **12**, 1039–1045.
10. Shimizu, N., Kambik, H., Yamazaki, Y., Imamoto, Y., and Kataoka, M. (2006) *Biochemistry* **45**, 3542–3547.
11. Genick, U. K., Devanathan, S., Meyer, T. E., Canestrelli, I. L., Williams, E., Cusanovich, M. A., Tollin, G., and Getzoff, E. D. (1997) *Biochemistry* **36**, 8–14.
12. Hendriks, J., van Stokkum, I. H. M., and Hellingwerf, K. J. (2003) *Biophys. J.* **84**, 1180–1191.
13. van Wilderen, L., Van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J., van Grondelle, R., and Groot, M. L. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15050–15055.
14. Takeshita, K., Imamoto, Y., Kataoka, M., Mihara, K., Tokunaga, F., and Terazima, M. (2002) *Biophys. J.* **83**, 1567–1577.
15. Takeshita, K., Imamoto, Y., Kataoka, M., Tokunaga, F., and Terazima, M. (2002) *Biochemistry* **41**, 3037–3048.
16. Chen, E. F., Gensch, T., Gross, A. B., Hendriks, J., Hellingwerf, K. J., and Kliger, D. S. (2003) *Biochemistry* **42**, 2062–2071.
17. Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J., and Hoff, W. D. (2001) *Biochemistry* **40**, 1510–1517.
18. Shimizu, N., Imamoto, Y., Harigai, M., Kambik, H., Yamazaki, Y., and Kataoka, M. (2006) *J. Biol. Chem.* **281**, 4318–4325.
19. Sergi, A., Gruning, M., Ferrario, M., and Buda, F. (2001) *J. Phys. Chem. B* **105**, 4386–4391.
20. Kumauchi, M., Hamada, N., Sasaki, J., and Tokunaga, F. (2002) *J. Biochem. (Tokyo)* **132**, 205–210.
21. Shimizu, N., Kambik, H., Mihara, K., Imamoto, Y., and Kataoka, M. (2002) *J. Biochem. (Tokyo)* **132**, 257–263.
22. Borucki, B., Joshi, C. P., Otto, H., Cusanovich, M. A., and Heyn, M. P. (2006) *Biophys. J.* **91**, 2991–3001.
23. Hoersch, D., Otto, H., Joshi, C. P., Borucki, B., Cusanovich, M. A., and Heyn, M. P. (2007) *Biophys. J.* **93**, 1687–1699.
24. van Brederode, M. E., Hoff, W. D., Van Stokkum, I. H., Groot, M. L., and Hellingwerf, K. J. (1996) *Biophys. J.* **71**, 365–380.
25. Hendriks, J., Hoff, W. D., Crielaard, W., and Hellingwerf, K. J. (1999) *J. Biol. Chem.* **274**, 17655–17660.
26. Kort, R., Hoff, W. D., Van West, M., Kroon, A. R., Hoffer, S. M., Vlieg, K. H., Crielaard, W., Van Beeumen, J. I., and Hellingwerf, K. J. (1996) *EMBO J.* **15**, 3209–3218.
27. Hendriks, J., Gensch, T., Hvid, L., van Der Horst, M. A., Hellingwerf, K. J., and van Thor, J. J. (2002) *Biophys. J.* **82**, 1632–1643.
pH Dependence of PYP Recovery

28. Kraayenhof, R., Schuurmans, J. J., Valkier, L. J., Veen, J. P. C., Van Marum, D., and Jasper, C. G. G. (1982) *Anal. Biochem.* **127**, 93–99
29. Hendler, R. W., and Shrager, R. I. (1994) *Biochem. Biophys. Methods* **28**, 1–33
30. van Brederode, M. E., Gensch, T., Hoff, W. D., Hellingwerf, K. J., and Braslavsky, S. E. (1995) *Biophys. J.* **68**, 1101–1109
31. Hendriks, J., Avila-Perez, M., Bruggeman, F., and Hellingwerf, K. J. (2007) in *From Computational Biophysics to Systems Biology (CBSB07)* (Hansmann, U. H. E., Meinke, J., Mohanty, S., and Zimmermann, O., eds) pp. 149–153 John von Neumann Institute for Computing, Juelich, Germany
32. Joshi, C. P., Borucki, B., Otto, H., Meyer, T. E., Cusanovich, M. A., and Heyn, M. P. (2005) *Biochemistry* **44**, 656–665
33. Hendriks, J., van Stokkum, I. H., Crielaard, W., and Hellingwerf, K. J. (1999) *FEBS Lett.* **458**, 252–256
34. Hoff, W. D., Devreese, B., Fokkens, R., Nugteren-Roodzant, I. M., Van Beeumen, J., Nibbering, N., and Hellingwerf, K. J. (1996) *Biochemistry* **35**, 1274–1281
35. Holiday, E. R. (1936) *Biochem. J.* **30**, 1795–1803
36. Lee, B. C., and Hoff, W. D. (2008) *Protein Sci.* **17**, 2101–2110
37. Meyer, T. E., Devanathan, S., Woo, T., Getzoff, E. D., Tollin, G., and Cusanovich, M. A. (2003) *Biochemistry* **42**, 3319–3325
38. Vreede, J., Hellingwerf, K. J., and Bolhuis, P. G. (2008) *Proteins* **72**, 136–149