Developing a Model System to Understand and Predict Photoactive Yellow Protein Behavior

A. Talattof, Dr. M. Cusanovich, Dr. R. Indik
Department of Mathematics, University of Arizona

May 17, 2005

1 Introduction

A photosensing protein directs light energy captured by its chromophore into a photocycle. The protein's structure must accommodate the photocycle and promote the resulting chemical or conformational changes that lead to signal transduction. The 1.4 Å crystallographic structure of photoactive yellow protein provides (PYP) the first view at atomic resolution of a protein with a photocycle [1]. PYP is a water soluble signaling protein involved in sensing blue light isolated from a halophilic phototrophic bacterium [2]. It serves as a prototype for a wide array of signaling proteins that contain the PAS domain structural motif [2]. PYP undergoes a self-contained light cycle. Light-induced trans-to-cis isomerization[1] and coupled protein rearrangements produce a new set of active-site hydrogen bonds. Resulting changes in shape, hydrogen bonding and electrostatic potential at the protein surface form a likely basis for signal transduction. The structural results suggest a general framework for the interpretation of protein photocycles [4].

The photocycle of PYP can be divided into several stages. First, the chromophore undergoes isomerization, initiated by the absorption of a photon in the ground state (P) and completed within a few nanoseconds to form the intermediate I_1. A deprotenated chromophore changes its isomerization state to cis by flipping its carbonyl group [5]. By convention, I_2 is designated to be structurally the furthest intermediate from the ground state before recovery, though the formation of an I_2' intermediate has been observed, during which proton transfer and conformational changes occur [2].

Functional activity of almost any protein, particularly of all enzymes, requires rapid dynamical fluctuations in its structure. In photoactive proteins light absorption by the chromophore leads to excited state formation, followed by a series of transient and reversible colour changes, usually referred to as its photocycle. This reversibility facilitates kinetic studies through the application of signal-averaging techniques in e.g. laser-induced transient absorption spectroscopy (TAS). The measured light-induced absorption changes as a function of time can be subjected to mathematical analysis to determine properties of relevant intermediates [6].

Several key mathematical terms discussed later will be defined. Let V and W be finite-dimensional inner product spaces, and let $T : V \to W$ be a linear transformation of rank r. Then there exist orthonormal bases $\{v_1, v_2, \ldots, v_n\}$ for V and $\{u_1, u_2, \ldots, u_n\}$ for W and positive scalars $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ such that $T(v_i) = \sigma_i u_i$ if $1 \leq i \leq r$ or 0 if $i \geq r$. Let A be an $m \,\mathrm{x}\, n$ matrix of rank r. Then there exists a unique $n \,\mathrm{x}\, m$ matrix B such that $(L_A)^\dagger : F^m \to F^n$ is equal

[1] Coordination compounds that exist in two or more isomeric forms may undergo reactions that convert one isomer (in this case 4-hydroxycinnamic acid chromophore) to another.

to the left-multiplication transformation L_B. We call B the pseudoinverse of A and denote it by $B = A^\dagger$ [3]. The pseudoinverse can computed using a singular value decomposition $A = U\Sigma V^*$. Let β and γ be the ordered bases whose vectors are the columns of V and U, respectively, and let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ be the nonzero singular values of A. Then β and γ are orthonormal bases for F^n and F^m, respectively. Defining $\Sigma^\dagger_{ij}$ as $\frac{1}{\sigma_i}$ if $i = j \leq r$ and 0 otherwise, $A^\dagger = V\Sigma^\dagger U^*$, and this is a singular value decomposition of $A^\dagger$.

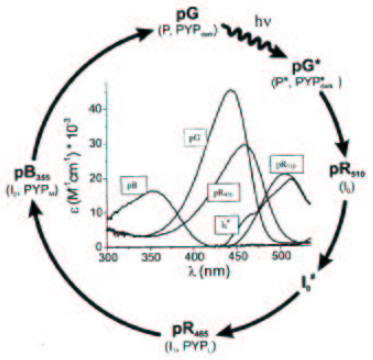

Figure 1: Overview of the photocycle of PYP. The key intermediates in the photocycle of PYP as detected with transient absorption spectroscopy at ambient temperature are indicated in the (redundant) nomenclature that is currently in use in the scientific literature. The inset shows the UV/vis spectra of these intermediates [5].

It is of interest to be able to model the kinetic phenomena that PYP undergoes leading to signal transduction under physiological conditions. Assuming there exists a direct relationship between protein folding and signaling [6], factors such as amplitude, wavelength, time, temperature and pH could thus be investigated in relation to one another. For example, both photocycle kintetics and spectra of photocycle intermediates are strongly dependent on pH [7]. Moreover, the ability to predict kinetic properties, relating structure to function, would prove useful in analyzing mutants of this protein. In this paper I present the findings thus far, including a simple model relating absorption to time, as well as an outline of what is yet to be completed.

2 Findings

The initial model is created by solving a system of linear equations with given [experimental] rate constants. Suppose $y'(t) = (y'_0(t), y'_1(t), y'_2(t), y'_3(t))$, the absorption rates, and that

$$A = \begin{pmatrix} -x1 & z1 & 0 & 0 \\ x1 & -(x2+z1) & z2 & 0 \\ 0 & x2 & -(x3+z2) & z3 \\ 0 & 0 & x3 & -z3 \end{pmatrix}$$

where x_i's and z_i's are the forward and reverse rate constants, respectively, as given in Fig. 2. Solving $\frac{d}{dx}y = log(2) * A * y$ yields a 232x4 matrix y consisting of absorbance values changing over time, each column representing an intermediate. Hence we can plot this, as seen in Fig. 4. The method of solving this exponential matrix equation lies is the diagonalizability of A. Let $A = PDP^{-1}$, where the columns of P are eigenvectors and the diagonal entries of D are eigenvalues

of A. In this case,

$$P = \begin{pmatrix} 0.6674 & 0.0400 & 0.0133 & 0.0000 \\ -0.7411 & 0.6837 & 0.2657 & 0.0001 \\ 0.0738 & -0.7287 & 0.5277 & 0.0002 \\ -0.0001 & 0.0050 & -0.8067 & 1.0000 \end{pmatrix}$$

and

$$D = 10^3 * \begin{pmatrix} -5.2776 & 0 & 0 & 0 \\ 0 & -0.7241 & 0 & 0 \\ 0 & 0 & -0.0033 & 0 \\ 0 & 0 & 0 & -9.7518 * 10^{-18} \end{pmatrix}$$

Hence $e^{tA} * y_0 = e^{tPDP^{-1}} * y_0 = P * e^{tD} * P^{-1} * y_0 = \sum_{j=1}^{m} e^{t\lambda_j} * w_j$ where w_j depends on y_0.

$$I1 \underset{250}{\overset{5000}{\Leftrightarrow}} I2 \underset{250}{\overset{500}{\Leftrightarrow}} I2' \underset{0.001}{\overset{5}{\Leftrightarrow}} P$$

Figure 2: A simplified reaction scheme showing the reaction rates (forward above, reverse below) in inverse seconds between the intermediate states.

To model the behavior of the intermediates I used Matlab's stiff ordinary differential equation solver, ode15s. An initial value of 1 was assumed for the absorbance ($y_0 = [1, 0, 0, 0]$). To make sense of the model, since it is on a time scale of ten powers of ten, a semi-logarithmic plot was made, taking the absorbance versus the log of time (see Fig. 3, 4).

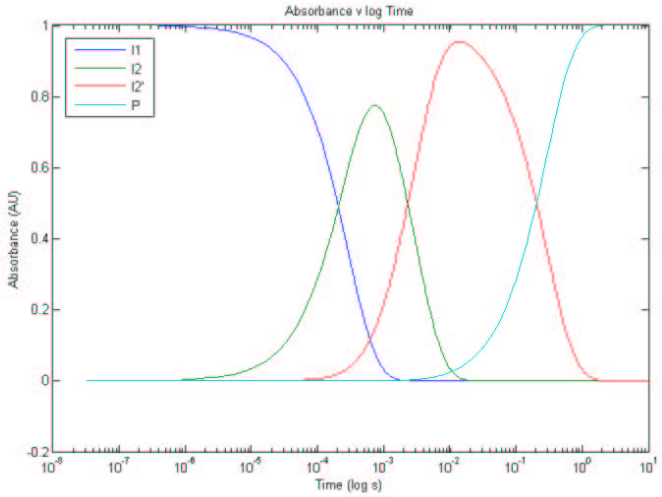

Figure 3: Semi-logarithmic scale of absorbance simulating a forward rate constants.

When considering only forward-rate reactions (Fig. 3), as each intermediate begins decaying, the following intermediate begins formation in a relatively smooth manner. However, when both forward and reverse rates are used, the decay and formation of intermediates becomes a more complex function (Fig. 4).

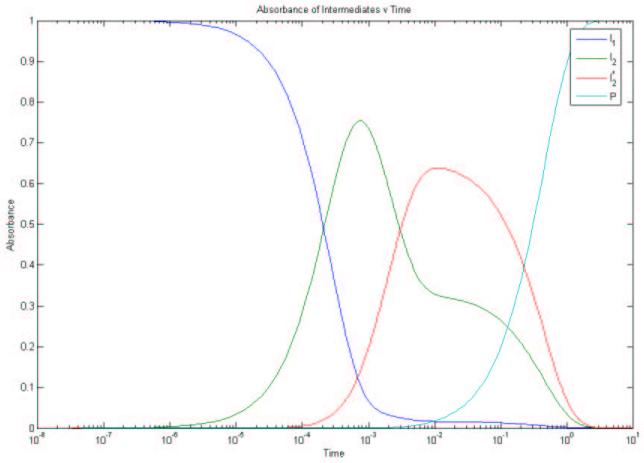

Figure 4: Semi-logarithmic scale of absorbance simulating forward and backward rate constants.

To determine the model that best fits the absorbance values at different wavelengths (see fig. 5), I constructed a matrix using four different wavelengths

$$W_\lambda = \begin{pmatrix} 460^4 & 460^3 & 460^2 & 460 & 1 \\ 420^4 & 420^3 & 420^2 & 420 & 1 \\ 380^4 & 380^3 & 380^2 & 380 & 1 \\ 350^4 & 350^3 & 350^2 & 350 & 1 \end{pmatrix}$$

and determined its pseudoinverse (using $pinv(W_\lambda, 1e-6)$ to allow for a tolerance of 10^{-6}), and multiplied that by each intermediate's absorbances given by column matrices $(A_{460}, A_{420}, A_{380}, A_{350})$, yielding the coefficient matrices $(a, b, c, d, e)_{intermediate}$ corresponding to $ax^4 + bx^3 + cx^2 + dx + e$. These experimental numbers, however, can only provide an insight as to how the intermediates actually behave at different wavelenghts.

Given a 17x4 matrix representing wavelengths (ranging from 500nm to 340nm in intervals of

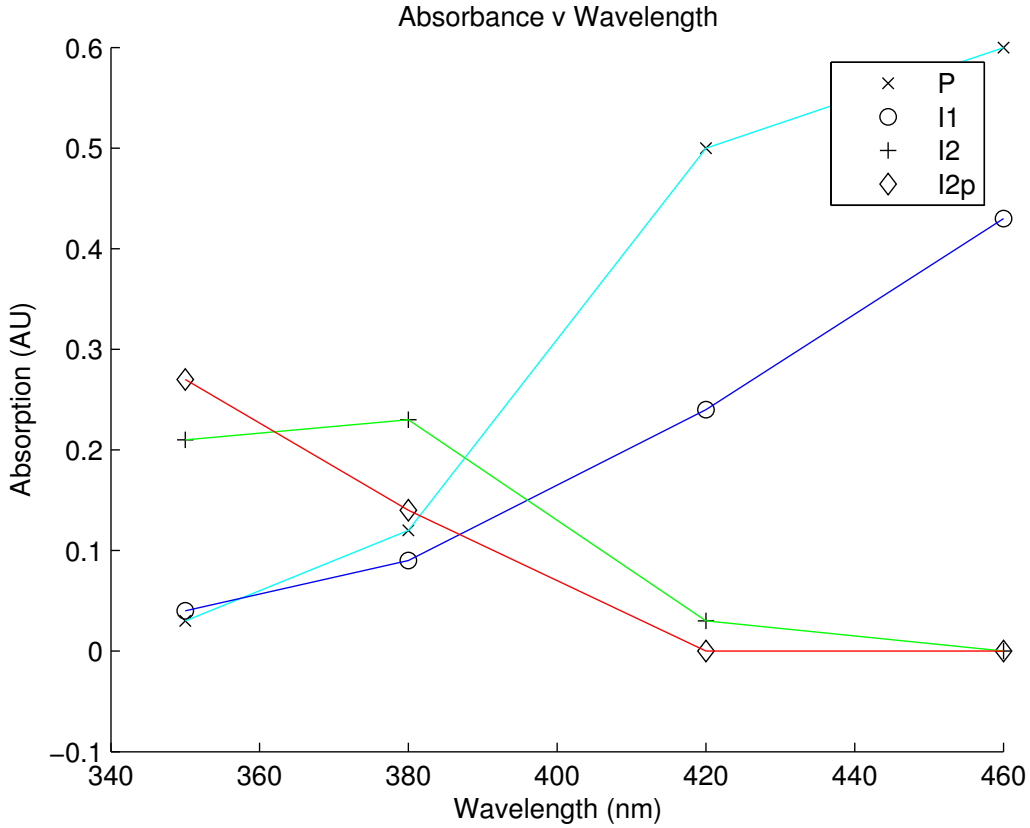

Figure 5: Plot showing absorbances of intermediates P, I_1, I_2, and I_2' at 350nm, 380nm, 420nm, and 460nm. The model that fits this data has been reduced to a 4th-degree polynomial, as anything above that behaves in the same manner.

10) and absorbance values of the intermediates $(I_1, I_2, I_2^{'}, P)$,

$$A_\lambda = \begin{pmatrix} .07 & .00 & .00 & .00 \\ .14 & .00 & .00 & .02 \\ .26 & .00 & .00 & .10 \\ .37 & .00 & .00 & .30 \\ .43 & .00 & .00 & .60 \\ .43 & .01 & .00 & .74 \\ .38 & .02 & .00 & .74 \\ .31 & .02 & .00 & .65 \\ .24 & .03 & .00 & .50 \\ .19 & .07 & .00 & .40 \\ .14 & .12 & .01 & .28 \\ .09 & .18 & .05 & .20 \\ .09 & .23 & .14 & .12 \\ .07 & .25 & .21 & .08 \\ .05 & .23 & .25 & .05 \\ .05 & .21 & .27 & .03 \\ .04 & .18 & .25 & .01 \end{pmatrix}$$

we can create a function relating these wavelength-specific absorbances showing, at each wavelength, the transition between intermediates over time.

For mapping the wavelength-specific absorbance values, we can create another matrix M_λ of size 232x17 where each column represents a particular wavelength. To create this matrix we use the following:

$$M_\lambda = (A_\lambda * y^{'})^{'} \tag{1}$$

Therefore if we plot M_λ against log-scaled time we can see the transition of absorbance from each wavelength (Fig. 6), going from I_1 on the left (faster times) to P on the right (slower times). Given the complexity of the graph, a useful tool in Matlab (version 7+) is employed, allowing the user to select a given point on any curve and receive the coordinates (in this case, time and absorbance). This should allow us to predict absorbance, given a wavelength, at a specific time, and thus see what intermediate the reaction is in.

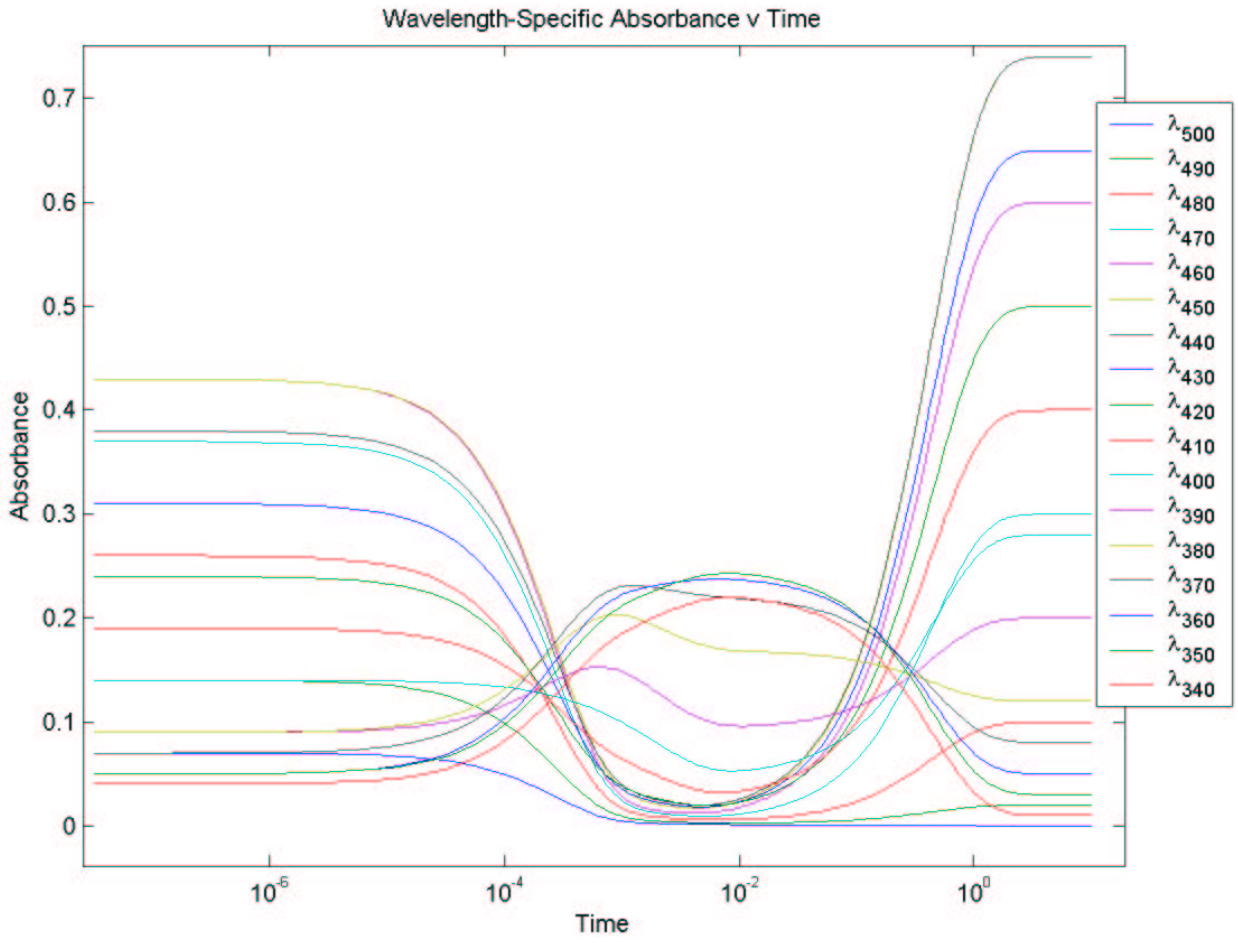

Figure 6: Semi-logarithmic scale of wavelength-specific absorbance showing transition between between intermediates.

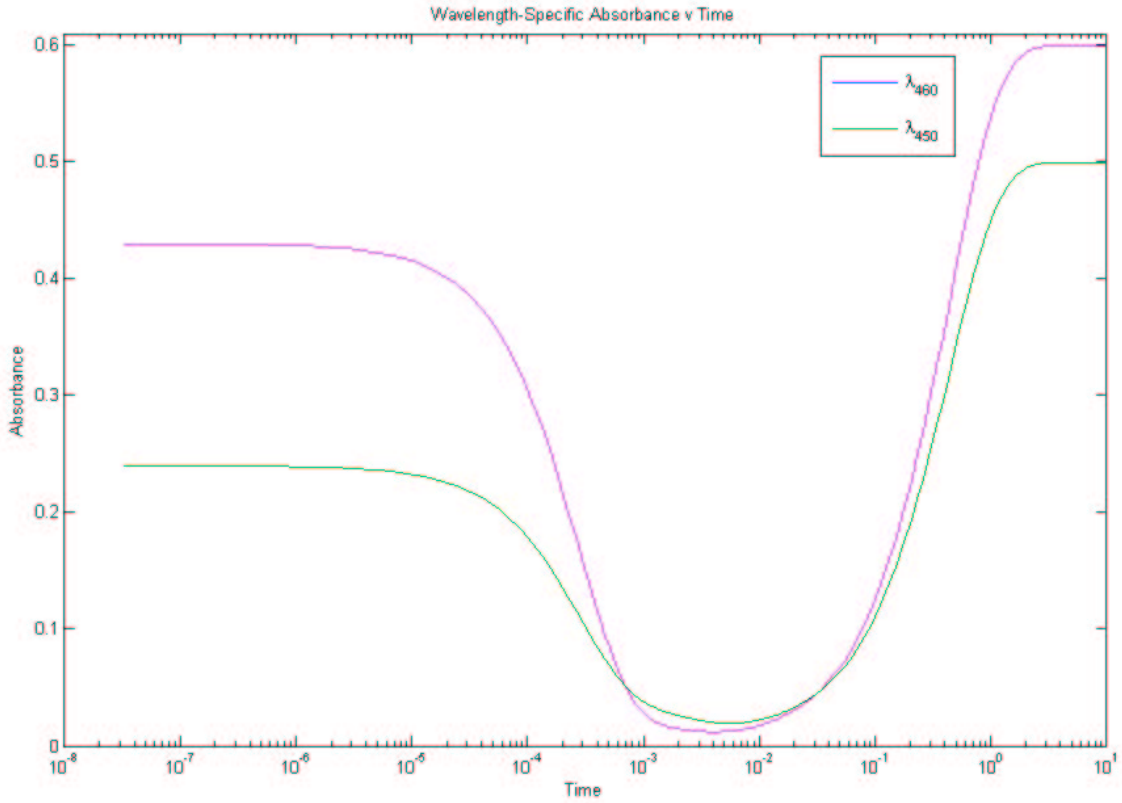

Figure 7: Semi-logarithmic scale of wavelength-specific absorbance showing transition between between intermediates having selected two plots in Matlab.

Furthermore, Matlab (again version 7+) permits a selection of plots to be shown. For example, one may want to compare absorbance as a function of time for only two wavelengths, say 460nm and 450 nm. This is shown in Figure 7.

3 Future Work

So far this model system is functional enough to provide rough predictions of results obtained in a laboratory setting. Certain parameters (initial absorbance, rate constants, wavelength-specific absorbance) can be modified to some extent, though more work is required to allow for inputs such as pH and temperature. In addition, inherent noise has not been accounted for, thus experimental results can not be judged for accuracy using this model as a comparison, as the noise may be indicative of strange behavior (especially at very fast times, where the noise is most abundant). Alas, given the few numbers of parameters that can be varied, a graphical user interface has not been developed, though this could easily be employed.

References

[1] Borgstahl, G.E.O., Williams, D.R., and Getzoff, E.D. The 1.4 Structure of Photoactive Yellow Protein, a Cytosolic Photoreceptor: Unusual Fold, Active Site and Chromophore. *Biochemistry*, 34:6278–6287, 1995.

[2] Cusanovich, M. A., and Meyer, T. E. Photoactive yellow protein: a prototypic PAS domain sensory protein and development of a common signaling mechanism. *Biochemistry*, 42:4759–4770, 2003.

[3] Friedberg, Stephen H., Arnold. J. Insel, Lawrence E. Spence. *Linear Algebra, 4th ed.* Pearson Education, Inc., 2003.

[4] Genick, U.K., Borgstahl, G.E.O., Ng, K., Ren, Z., Pradervand, C., Burke, P.M., Srajer, V., Teng, T.Y., Schildkamp, W., McRee, D.E., Moffat, K. and Getzoff, E.D. Millisecond Time-Resolved Laue Crystallography: Structure of a Protein Photocycle Intermediate. *Science*, 275:1471–1475, 1997.

[5] Hellingwef, K. J., Hendriks, J., and Gensch, T. On the Configurational and Conformational Changes in Photoactive Yellow Protein that Leads to Signal Generation in Ectothiorhodospira halophila. *J. Biol. Phys.*, 28:395–412, 2002.

[6] Hendriks, J., Stokkum, I.H.M., and Hellingwerf, K.J. Deuterium isotope effects in the photocycle transitions of the photoactive yellow protein. *Biophys. J.*, 84:1180–1191, 2003.

[7] Lee, B., Pandit, A., Croonquist, P. A., and Hoff, W. D. Folding and signaling share the same pathway in a photoreceptor. *Proc. Natl. Acad.*, 98:9062–9067, 2001.