SUPPLEMENTAL MATERIALS

"Functional Tuning of Photoactive Yellow Protein by Active Site Residue 46" by Philip et al.

Supplemental Methods

Protein purification. For purification of PYP in 96-well format, cells overproducing PYP were frozen, thawed, and each well was resuspended in 200 µL of Lysis Buffer with 200 µg/mL lysozyme and 0.5 µg/mL DNase I. The cells were lysed by incubation on a shaker at 4 °C for one hour, and 10 µL of pCA anhydride was added (Imamoto *et al.*, 1995) to each well in increments, shaking between additions. The insoluble cell debris was centrifuged down (2800 g, 30 min) and the soluble fractions were pipetted into GF/C filter plates (Whatman #7700-3301) stacked on top of polypropylene filter plates (Whatman #7700-3305) containing 40 µL Ni-NTA agarose (QIAGEN) per well that was washed with Lysis Buffer by filling the plates with buffer and draining them with a vacuum apparatus. The bottoms of the plates were sealed with Parafilm and folded paper towels were placed underneath the Parafilm to ensure that it sealed tightly against the plate. The soluble cell extract was centrifuged into the polypropylene plates and incubated on a shaker at 4 °C for one hour. After incubation, the resin was washed by filling the plates with buffer and draining into collection plates by centrifugation. A denaturing wash step was used first (100 mM sodium phosphate, 10 mM Tris-Cl, 6 M guanidine HCl, pH 8.0) followed by three washes with Lysis Buffer, and one wash with Elution Buffer lacking EDTA. The denaturing wash step was found to improve the purity index (for wtPYP: Abs_{280}/Abs_{446}) from ~0.7 to ~0.6. Protein was eluted from the resin by incubating for 30 min with 200 µL of Elution Buffer containing 50 mM EDTA and centrifuging into a collection plate. The purity of the resulting PYP samples was measured by visible absorbance spectroscopy and SDS gel electrophoreses (supplemental Fig. S1). Eighteen mutants were isolated from this random approach based on their absorbance spectrum. E46W was obtained using a specific mutagenic primer and found to have an absorbance spectrum very similar to that of E46Q, explaining why it was not readily identified. A complete list with the properties of the E46X mutants reported here is provided in Table 1 in the main text.

Fluorescence data collection. Initially, a 96-well plate accessory was used to measure fluorescence spectra of the E46X mutants. However, due to the low Φ_{fl} of the pCA in wtPYP (0.14 %), the signals had an unacceptably high degree of noise. Therefore, these spectra were measured for individual mutants using a microcuvet.

Supplemental Discussion

Considerations regarding the high pH range of the pH titrations–In the pH titrations of the E46X mutants we found that above pH 10.5, pCA thiol ester hydrolysis (Hoff *et al*., 1996) interfered with our ability to use the steady state visible absorbance of PYP for determining the pH titration curve. In many of the mutants this resulted in an increase in UV absorbance between 330 and 380 nm at pH values greater than 9 (data not shown). The result is a softening of the bend in the titration curve at the high pH range that may explain why most of the calculated *n* values for the remaining mutants are slightly less than one. For clarity the bleaching of the visible absorbance bands that occurs for all of the mutants above pH 10 due to thioester hydrolysis was omitted from Fig. 4B in the main text. For one mutant, E46K, high pH also resulted in a broadening and red-shift of the visible absorbance peak (data not shown). These results suggest that the titration behavior of E46K in the high pH range is more complex.

Analysis of pG recovery kinetics–We examined if any of the E46X mutants exhibits a major phase in the pB to pG that is beyond the 100 ms time resolution of the HP8453 diode-array spectrophotometer used in this study. This was prompted by the apparent moderately slow photocycle kinetics of this mutant, in contrast to the very fast kinetics reported in literature. In order to validate our data, we used the observed rate constants for pB decay to predict expected steady state photobleaching fractions, and compared these to the experimentally observed bleaching fractions. The nature of the

mathematical relationship between the recovery rates and the bleaching fractions is dependent on the model of the photocycle. We used the simplest model that could account for two recovery rates: that of two independent photocycles (supplemental Fig. S3A). In this model, the steady state bleaching fraction is given by:

$$f_B = \frac{k_1 k_4 + k_2 k_3}{k_2 k_4 + k_1 k_4 + k_2 k_3}$$

Note that the goal of this analysis is not to prove that this model is correct – we wish simply to prove that our data are self-consistent. This would allow us to determine if other mutants also have a significant fast (and thus undetected) component, as detected by an unexpectedly low bleaching percentage. This analysis assumes that the quantum yields (and therefore k_1 and k_3) of all mutants is the same. The predicted and actual bleaching fractions are plotted in supplemental Fig. S3B. The E46Q mutant stands out as having a bleaching fraction significantly lower than predicted by our simple model. We used a least squares fitting routine using the measured values of k_2 and k_4 for each mutant to find the one set of values of k_1 and k_3 that was best for all of the mutants. To check the degree of correspondence between predicted and observed bleaching percentage for a particular mutant, we ran the least squares fitting routine using all of the samples – except the mutant in question. The sum of the squared differences serves as an indicator of how well the fitting program achieved its goal. If a particular mutant shows a large discrepancy between its predicted and actual bleaching fraction, then the sum of the squared differences when we *exclude* that mutant should be significantly reduced. Supplemental Fig. S3C shows the sum of the squared differences when each of the mutants is individually excluded from the fit. When E46Q is excluded, the sum of the squared differences is by far the lowest, identifying this mutant as an outlier from the rest of our samples. We conclude that in our experimental setup we did not detect the very fast recovery rate of the majority of the E46Q sample, but instead only detected a small fraction of the sample with a much smaller recovery rate. The good match between predicted and observed bleaching fraction for the other mutants provides strong evidence that we accurately measured the photocycle rate of these samples. This is confirmed by the good match between the published values for the pG recovery kinetics of of wtPYP and its E46A and E46D mutants (supplemental Table S2).

Validation of microscale data–To validate the 96-well based high-throughput measurements reported here, we compared the results reported here with values published in literature (supplemental Table S2). For wtPYP we calculated a pK_a of 3.0, compared to the published value of 2.8 (Meyer 1985; Hoff *et al.*, 1997). This discrepancy is most likely due to the truncation of the titration curve at pH 2. Below that value, the impurities in the sample precipitated, making the solution cloudy and spectrophotometric measurements unreliable. Similarly, the apparent pK_a for E46D may be somewhat underestimated due to truncation of the titration curve above pH 10.

The stability that we measured for wtPYP is ~4 kcal/mol higher than the two published values (Genick *et al*., 1997; Lee *et al*., 2001). However, the two values reported in the literature themselves differ by 1.6 kcal/mol. These differences may be due to the composition of the buffers used or the methods used to analyze the titration data. A similar concern regarding differences in experimental conditions applies to detailed comparisons of the kinetics of pG recovery (see supplemental Table S2). However, the measurements reported here are consistent with the conclusion that wtPYP and the E46A mutant exhibit a pG recovery time constant near 1 second, while this process is much slower in the E46D mutant (~250 seconds).

Analysis of correlations in the data–To investigate possible correlations among the measured properties and with side chain structure, we grouped the mutants by hierarchical clustering using cluster software available from http://rana.lbl.gov/EisenSoftware.htm (Eisen *et al*., 1998) (supplemental Fig. S4). The mutants were clustered according to (1) pCA spectral parameters, (2) pCA Φ_{fl}, (3) pCA pH titration behavior, (4) guanidine titration behavior, and (5) photobleaching. Each of these general properties was given equal weight. Some of these properties were subdivided into parameters with fractional weights. The spectral parameters category was composed of four values, each with one-quarter weight: (a) λ_{max}, (b) absorbance peak bandwidth at ¾ height, (c) fluorescence emission maximum, and (d) emission peak

bandwidth at ¾ height (parameters b, c, and d will be presented and discussed elsewhere). These four values were grouped together with fractional weights because we expect them to be governed by similar factors. The pH titration behavior was subdivided into two values, each with one-half weight: (a) apparent pCA pK_a, and (b) *n* value, since these values correlate (see Table 1 in main text). The guanidine titration behavior was also subdivided into two values, each with one-half weight: (a) $\Delta G°_U$, and (b) *m* value. We treated wild type PYP as an equal member of the Glu46 library because here we are interested in how it would be classified relative to the entire group. Therefore, prior to clustering, each variable was re-centered at the group median. The resultant tree is shown in supplemental Fig. S4. The resulting clustering provides insights into the correlations between the properties of the side chain at position 46 and the resulting PYP chain active site properties.

The D, K, P, and R substitutions are similar with regards to several properties, and form a tightly knit group in the cluster analysis (supplemental Fig. S4). They all have blue-shifted λ_{max} values, high Φ_{fl}, high pK_a values, *n* values distinctly less than unity, low $\Delta G°_U$, and high bleaching fractions. This group of mutants deviates most from the rest of the Glu46 library. Surprisingly, it is also the most chemically heterogeneous group, with two being basic, one acidic, and one cyclic nonpolar. In contrast, all of the other groups (see supplemental Fig. S4) exhibit obvious chemical similarities among their members. These four substitutions appear to be the most disruptive to the PYP active site, and can be interpreted as having a less rigid pCA binding pocket, resulting in a shift in pCA properties towards its behavior in solution.

The mutant that most closely groups with wild type PYP is the conservative substitution Q. E and Q are of similar size and shape, and both form hydrogen bonds with the pCA. Biochemically, the properties that set them apart from most of the other samples are their low pK_a values, *n* values greater than unity, and high stabilities. The I, L, M, and V group of substitutions all involve small hydrophobic side chains that remove the wild type hydrogen bond, and likely do so without significantly perturbing the structure except for the disruption of the hydrogen bond to the pCA. This group has the most red-shifted absorbance spectra (and smallest bandwidths), pK_a values of ~7, *n* values very close to unity, moderate $\Delta G°_U$, and low degree of photobleaching. These two groups together confirm and extend the importance of the E46-pCA hydrogen bond in both spectral tuning and pK_a tuning.

Asn and Thr are of similar size and shape, and both are capable of donating a proton to a hydrogen bond. However, it is unlikely that E46T forms a hydrogen bond with the pCA given its large red-shifted λ_{max} (which is very close to that of the I/L/M/V group). In addition, using the DeepView program (http://us.expasy.org/spdbv/mainpage.html) to model the E46T substitution indicates that T can fit into the same space as the wild type E without any steric hindrance—as long as its side chain hydroxyl is directed away from the pCA. The similarity between the absorbance spectra of E46N and E46Q suggests that E46N does form a hydrogen bond with the pCA.

The last grouping that the cluster analysis produced contains E46F and E46Y. Again, the two members of this group have side chains of similar size and shape. It is likely that the properties of these two mutants are determined largely by their phenyl rings. They have high Φ_{fl}, and high pK_a values, similar to the D/K/P/R group. They differ from that group in that their *n* values are closer to unity, their stabilities are not as low, and they do not photobleach as much. One respect in which E46F and E46Y differ from each other significantly is in their λ_{max} values (470 and 445 nm, respectively). The similarity of the E46Y λ_{max} to that of wild type PYP suggests that Y hydrogen bonds with the pCA. The analysis presented here provides evidence for the preservation of a hydrogen bond between the pCA and 4 for the 19 substitutions: N, Q, W, and Y. This has been experimentally confirmed only for the E46Q mutant (Anderson *et al*., 2004; Sugishima *et al*., 2004); for the remaining three substitutions further structural studies are needed.

Spectral tuning effects–While spectral tuning in PYP depends critically on hydrogen bonding to the pCA, counter ion effects are central in the extensively studied spectral tuning mechanism in the rhodopsins. The wild type hydrogen bonds from the pCA to Glu46 and Tyr42 in the dark state of PYP are likely to function in a similar manner to the counter-ion in the opsins (Blatz et al., 1972; Nakanishi *et al.,* 1980): by stabilizing the charge-localized ground state. We should then expect to see a similar effect

whether the influence is from a hydrogen bond donor or a positively charged group, since this would also stabilize the ground state of the negatively charged pCA. When we substitute Lys or Arg into position 46, we do see a blue-shifted absorbance spectrum. We also see this however, when we substitute Asp or Pro. It should be noted that the protonation states of these substituted residues are not known at this stage. Detailed structural information will be needed to satisfactorily address this issue.

Supplemental References

Anderson, S., Crosson, S., and Moffat, K. (2004) Short hydrogen bonds in photoactive yellow protein. *Acta Cryst.* **D60**, 1008-1016

Blatz, P.E., Mohler, J.H., and Navangul, H.V. (1972) Anion-induced wavelength regulation of absorption maxima of Schiff bases of retinal. *Biochemistry* **11**, 848-855

Devanathan, S., Brudler, R., Hessling, B., Woo, T.T., Gerwert, K., Getzoff, E.D., Cusanovich, M.A., and Tollin, G. (1999) Dual photoactive species in Glu46Asp and Glu46Ala mutants of photoactive yellow protein: A pH-driven color transition. *Biochemistry* **38**, 13766-13772

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863-14868

Genick, U.K., Devanathan, S., Meyer, T.E., Canestrelli, I.L., Williams, E., Cusanovich, M.A., Tollin, G., and Getzoff, E.D. (1997) Active site mutants implicate key residues for control of color and light cycle kinetics of photoactive yellow protein. *Biochemistry* **36**, 8-14

Hoff, W.D., Devreese, B., Fokkens, R., Nugteren-Roodzant, I.M., Van Beeumen, J., Nibbering, N., and Hellingwerf, K.J. (1996) Chemical reactivity and spectroscopy of the thiol ester-linked *p*-coumaric acid chromophore in the photoactive yellow protein from *Ectothiorhodospira halophila*. *Biochemistry* **35**, 1274-1281

Hoff, W.D., van Stokkum, I.H.M., Gural, J., and Hellingwerf, K.J. (1997) Comparison of acid denaturation and light activation in the eubacterial blue-light receptor photoactive yellow protein. *Biochim. Biophys. Acta* **1322**, 151-162

Imamoto, Y., Ito, Y., Kataoka, M., and Tokunaga, F. (1995) Reconstitution of photoactive yellow protein from apoprotein and *p*-coumaric acid derivatives. *FEBS Lett*. **374**, 157-160

Imamoto, Y., Koshimizu, H., Mihara, K., Hisatomi, O., Mizukami, T., Tsujimoto, K., Kataoka, M., and Tokunaga, F. (2001) Roles of amino acid residues near the chromophore of photoactive yellow protein. *Biochemistry* **40**, 4679-4685

Lee, B.-C., Pandit, A., Croonquist, P.A., and Hoff, W.D. (2001) Folding and signaling share the same pathway in a photoreceptor. *Proc. Natl. Acad. Sci. USA* **98**, 9062-9067

Meyer, T.E. (1985) Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira halophila*. *Biochim. Biophys. Acta* **806**, 175-183

Meyer, T.E., Devanathan, S., Woo, T., Getzoff, E.D., Tollin, G., and Cusanovich, M.A. (2003) Site-specific mutations provide new insights into the origin of pH effects and alternative spectral forms in the photoactive yellow protein from *Halorhodospira halophila*. *Biochemistry* **42**, 3319-2533

Mihara, K., Hisatomi, O., Imamoto, Y., Kataoka, M., and Tokunaga, F. (1997) Functional expression and site-directed mutagenesis of photoactive yellow protein. *J. Biochem.* **121**, 876-880

Nakanishi, K., Baloghnair, V., Arnaboldi, M., Tsujimoto, K., and Honig, B. (1980) An external point charge model for bacteriorhodopsin to account for its purple color. *J. Am. Chem. Soc*. **102**, 7945-7947

Sugishima, M., Tanimoto, N., Soda, K., Hamada, N., Tokunaga, F., and Fukuyama, K. (2004) Structure of photoactive yellow protein (PYP) E46Q mutant at 1.2 Å resolution suggests how Glu46 controls the spectroscopic and kinetic characteristics of PYP. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2305-2309

van der Horst, M.A., van Stokkum, I.H.M., Dencher, N.A., and Hellingwerf, K.J. (2005) Controlled reduction of the humidity induces a shortcut recovery reaction in the photocycle of photoactive yellow protein. *Biochemistry* **44**, 9160-9167

Supplemental Table S1. Summary of the properties of the 20 proteins in the E46X library.

a.a.	λ_{max} (nm)	Φ_{fl}	pK_a	*n*	*m* (kcal/mol/M)	$\Delta G°_{unfolding}$ (kcal/mol)	Fraction Bleached
A	469 ± 1	.55 ± .02	7.88 ± .03	.95 ± .05	4.7 ± .3	12.3 ± .9	.31 ± .05
C	472 ± 1	.64 ± .05	7.0 ± .2	.71 ± .14	4.4 ± .7	10.3 ± 1.7	.09 ± .05
D	442 ± 1	.79 ± 0.03	9.0 ± .4	.72 ± .03	3.4 ± .3	6.5 ± .7	.72 ± .05
E	446 ± 1	.19 ± .04	3.05 ± .07	1.73 ± .05	4.50 ± .15	11.9 ± .6	.19 ± .05
F	470 ± 1	1.05 ± .09	8.64 ± .05	.92 ± .01	4.4 ± .3	9.9 ± .7	.07 ± .05
G	463 ± 1	.66 ± .12	7.84 ± .10	.94 ± .01	4.7 ± .2	9.4 ± .4	.40 ± .05
H	454 ± 1	.62 ± .02	7.10 ± .06	.85 ± .10	4.44 ± .06	9.07 ± .17	.29 ± .05
I	478 ± 1	.56 ± .02	7.01 ± .07	.92 ± .04	4.35 ± .15	11.0 ± .3	.08 ± .05
K	444 ± 1	1.27 ± .04	8.01 ± .06	.81 ± .03	3.7 ± .3	7.3 ± .7	.48 ± .05
L	475 ± 1	.44 ± .03	7.33 ± .02	.89 ± .03	4.89 ± .17	11.4 ± .7	.06 ± .05
M	476 ± 1	.39 ± .06	6.50 ± .07	.95 ± .03	4.4 ± .6	12.9 ± 1.4	.10 ± .05
N	463 ± 1	.60 ± .03	7.50 ± .01	.86 ± .03	4.8 ± 1.4	10 ± 3	.13 ± .05
P	441 ± 1	.85 ± .16	8.4 ± .4	.71 ± .06	3.6 ± .2	6.8 ± .6	.70 ± .05
Q	461 ± 1	.37 ± .02	4.64 ± .02	1.36 ± .10	4.5 ± .6	12.4 ± 1.6	.05 ± .05
R	442 ± 1	.98 ± .12	8.3 ± .4	.8 ± .3	3.92 ± .06	7.3 ± .9	.69 ± .05
S	464 ± 1	.54 ± .02	7.18 ± .06	.881 ± .015	4.57 ± .11	10.0 ± .5	.23 ± .05
T	475 ± 1	.61 ± .04	7.29 ± .06	.85 ± .05	4.80 ± .12	10.14 ± .17	.06 ± .05
V	478 ± 1	.53 ± .03	7.58 ± .03	.93 ± .03	4.71 ± .14	10.44 ± .10	.07 ± .05
W	461 ± 1	.65 ± .07	7.14 ± .02	.89 ± .06	4.29 ± .18	10.47 ± .06	.48 ± .05
Y	445 ± 1	1.44 ± .11	8.01 ± .01	.85 ± .06	4.33 ± .01	8.62 ± .22	.30 ± .05

Supplemental Table S2. Comparison of measured values reported here with published values.

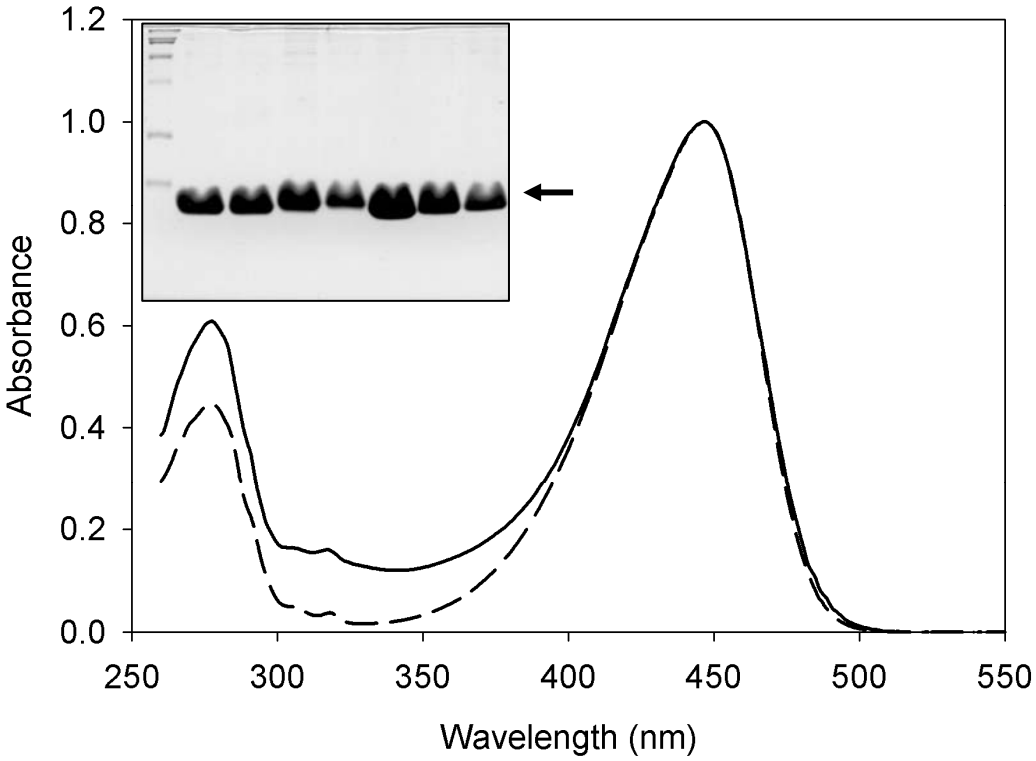
Property	Mutant	Determined here	Published Value
λ_{max} (nm)	wtPYP	446	446[f]
	E46Q	461	460[h]/462[b]
	E46A	469	465[a]/469[d]
	E46D	442	444[a]
pK_a	wtPYP	3.0	2.8[c]/3.0[d]
	E46Q	4.6	4.8[g]/5.3[d]
	E46A	7.9	7.8[d]/7.9[a]
	E46D	9.0	8.6[a]
$\Delta G°_U$ (kcal/mol)	wtPYP	11.9	7.1[b]/8.7[e]
k (s^{-1})	wtPYP	1.013 (61 %)	7 (46 %)
		0.03 (39 %)	1.6 (54 %)[i]
	E46Q[†]	0.609 (11 %)	21 (94 %)
		0.143 (89 %)	6 (3.4 %)[i]
	E46A	0.345 (98 %)	0.5[a]/0.77[d] [‡]
		0.087 (2 %)	
	E46D	0.012 (23 %)	0.0026[a]
		0.002 (77 %)	

[†]This discrepancy was explained by the presence of a small fraction of the E46Q sample that exhibits a slow photocycle, see text.

[‡]A fast phase of 8.1 x 10^3 s^{-1} has been observed in the literature (Devanathan *et al*., 1999), but is beyond the time resolution of our experimental setup.

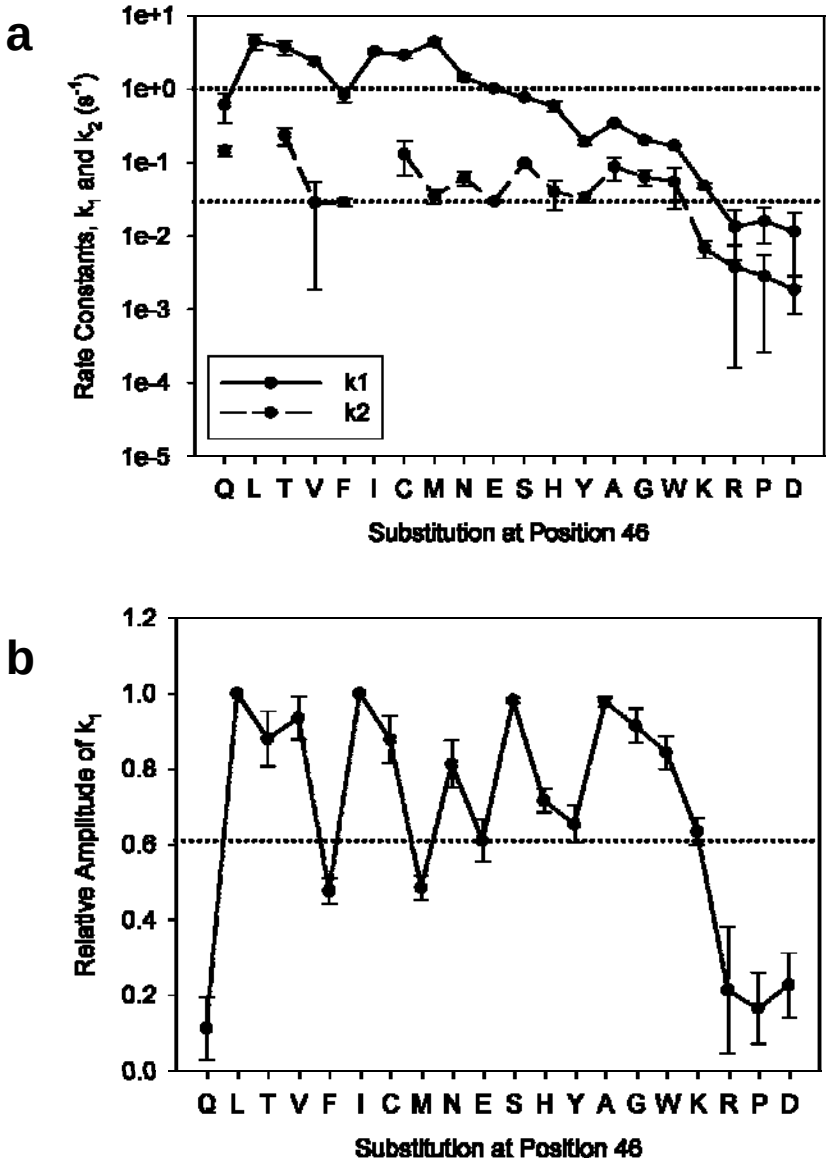
[a]Devanathan *et al*., 1999; [b]Genick *et al*., 1997; [c]Hoff *et al*., 1997; [d]Imamoto *et al*., 2001; [e]Lee *et al*., 2001; [f]Meyer, 1985; [g]Meyer *et al*., 2003; [h]Mihara *et al*., 1997; [i]van der Horst *et al*., 2005.

Supplemental Figure S1



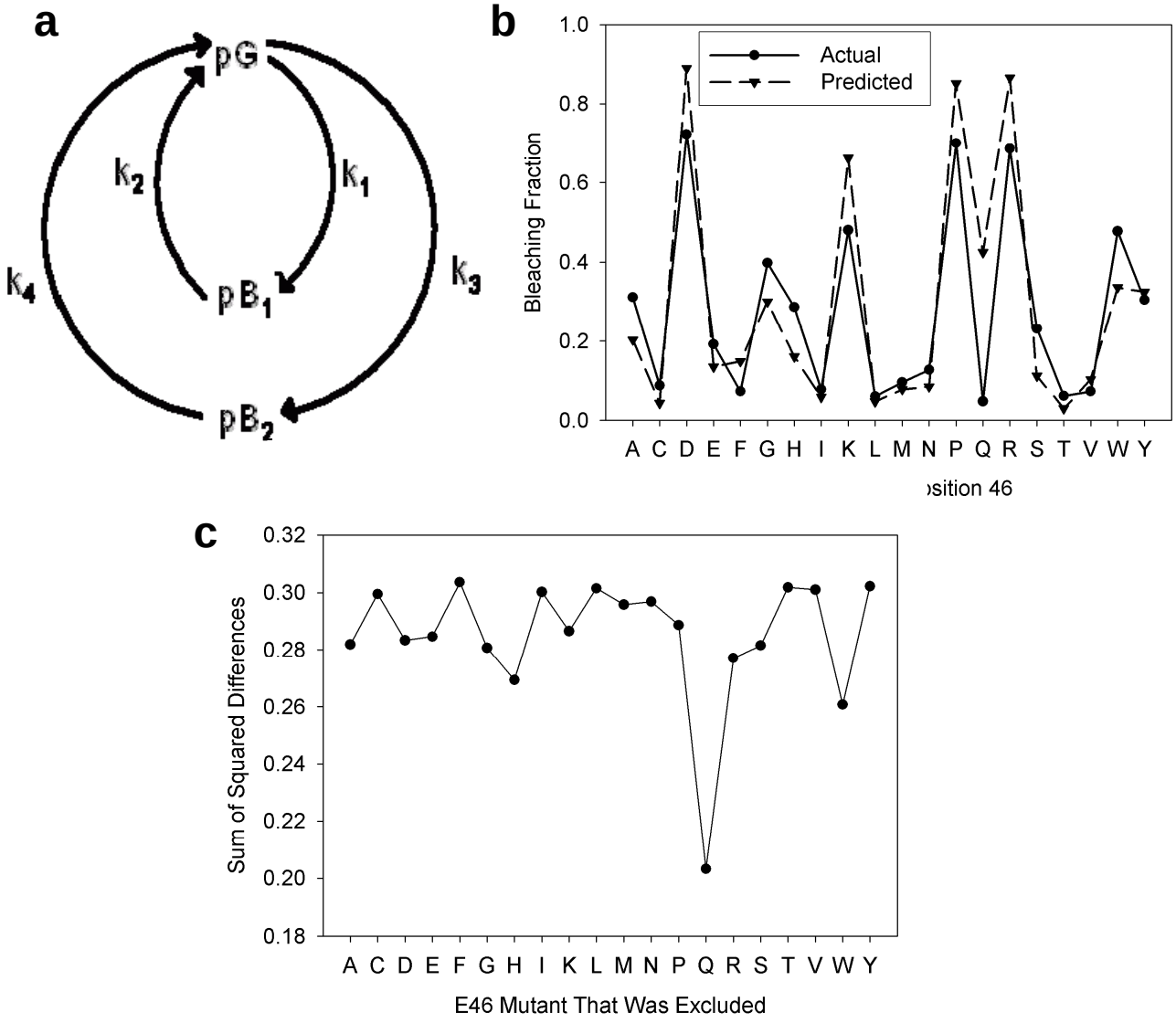
Supplemental Figure S1. Microscale purification of PYP. Normalized absorbance spectra of wtPYP prepared using the microscale method (solid) and a traditional large-scale three-step purification (dashed). Inset shows SDS-PAGE of seven different samples prepared using the microscale method. The first lane contains molecular weight markers. His-tagged PYP appears as a 16 kDa band (arrow). Even though the gel is overloaded, impurities are barely visible.

Supplemental Figure S2



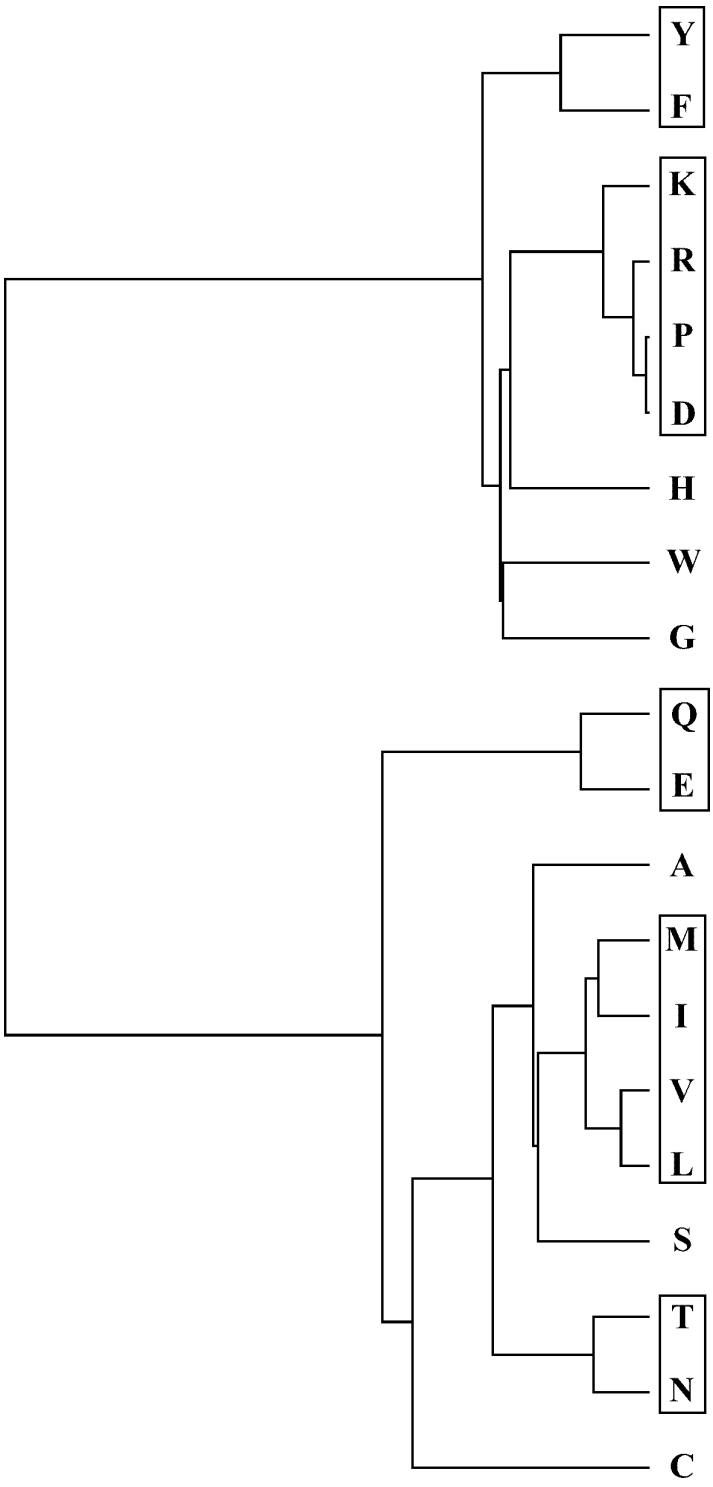
Supplemental Figure S2. Analysis of pG recovery kinetics in the E46X mutants. (**a**) The rate constants k_1 and k_2, sorted from the mutant with the lowest bleaching fraction to that with the greatest bleaching fraction. Error bars represent the standard deviations from three or more datasets. Some traces were fit with a monoexponential function, and therefore lack a value for k_2. Dotted lines mark the values of wild type PYP. (**b**) The relative contribution of k_1, sorted by bleaching fraction. For many mutants the kinetics of pG recovery from pB was biexponential, complicating direct comparison between the photocycle kinetics of different mutants. This prompted us to use the percentage of photobleaching to describe the kinetics of pB decay in a single value.

Supplemental Figure S3



Supplemental Figure S3. Comparison of experimentally determined and predicted bleaching fractions. (**a**) Photocycle scheme used to predict bleaching percentages based on observed pB decay kinetics. This simple model was used to fit the measured pG recovery rates and calculate steady state bleaching percentages. (**b**) Experimentally determined and predicted bleaching fractions. A clear discrepancy between the actual and predicted bleaching fractions was observed only for the E46Q mutant. (**c**) Testing for mutants in which the detected photocycle kinetics and photobleaching percentage do not match. We fit our model to the photocycle kinetics and photobleaching data for each of the twenty samples in the E46X library, specifically excluding one sample in each fit. The sum of the squared differences from the resulting twenty least squares fits reveal a clear improvement in the fit (i.e., the model fits better) when E46Q is excluded.

Supplemental Figure S4



Supplemental Figure S4. Hierarchical clustering of the E46X library of mutants based on their structural and functional characteristics.