Protonation States and pH Titration in the Photocycle of Photoactive Yellow Protein†

Eugene Demchuk,‡ Ulrich K. Genick, Tammy T. Woo, Elizabeth D. Getzoff, and Donald Bashford*

Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received June 30, 1999; Revised Manuscript Received October 25, 1999

ABSTRACT: Photoactive yellow protein (PYP) undergoes a light-driven cycle of color and protonation states that is part of a mechanism of bacterial phototaxis. This article concerns functionally important protonation states of PYP and the interactions that stabilize them, and changes in the protonation state during the photocycle. In particular, the chromophore pK_a is known to be shifted down so that the chromophore is negatively charged in the ground state (dark state) even though it is buried in the protein, while nearby Glu46 has an unusually high pK_a. The photocycle involves changes of one or both of these protonation states. Calculations of pK_a values and protonation states using a semi-macroscopic electrostatic model are presented for the wild-type and three mutants, in both the ground state and the bleached (I_2) intermediate state. Calculations allowing multiple H-bonding arrangements around the chromophore also have been carried out. In addition, ground-state pK_a values of the chromophore have been measured by UV−visible spectroscopy for the wild-type and the same three mutants. Because of the unusual protonation states and strong electrostatic interactions, PYP represents a severe test of the ability of theoretical models to yield correct calculations of electrostatic interactions in proteins. Good agreement between experiment and theory can be obtained for the ground state provided the protein interior is assumed to have a relatively low dielectric constant, but only partial agreement between theory and experiment is obtained for the bleached state. We also present a reinterpretation of previously published data on the pH-dependence of the recovery of the ground state from the bleached state. The new analysis implies a pK_a value of 6.37 for Glu46 in the bleached state, which is consistent with other available experimental data, including data that only became available after this analysis. The new analysis suggests that signal transduction is modulated by the titration properties of the bleached state, which are in turn determined by electrostatic interactions. Overall, the results of this study provide a quantitative picture of the interactions responsible for the unusual protonation states of the chromophore and Glu46, and of protonation changes upon bleaching.

This paper presents theoretical calculations and experimental measurements designed to elucidate the factors responsible for functionally significant protonation states in the ground state and bleached state (I_2) of the light cycle of PYP.[1] We have focused on the chromophore, which is deprotonated and negatively charged in the ground state but protonated in the bleached state, and on the nearby side chains of Tyr42, Glu46, and Arg52, which may influence the chromophore. The methods used include calculations based on a macroscopic electrostatic model incorporating atomic detail, and spectroscopic measurements on site-directed mutants under varying pH conditions. We also present a reinterpretation of data on the pH-dependence of the photocycle kinetics (*1*) to provide an explicit model of protonation states of the chromophore and Glu46 in an intermediate thought to be involved in signal transduction.

PYP is a small cytosolic, globular, light-sensing protein implicated in the negative phototaxis of the bacterium *Ectothiorhodospira halophila*. It consists of 125 amino acid residues and a 4-hydroxycinnamic acid chromophore attached through a thioester linkage to Cys69 (*2*−*4*). In the ground state, the protein has a bright yellow color resulting from an absorbance maximum at 446 nm (*5*), and the chromophore is buried within the protein structure and has a negative charge due to the deprotonation of its phenolic group (*3*, *6*, *7*). The photocycle (*8*−*10*) includes an early red-shifted (λ_{max} = 495 nm) intermediate I_1 in which the chromophore has undergone a *cis*−*trans* isomerization but structural changes are otherwise minor (*11*); and a blue-shifted (bleached) intermediate state I_2 (λ_{max} = 340 nm) in which the chromophore swings out into solution and is protonated (*12*). Although the phototropic signal transduction pathway is not fully known, the bleached state is thought to be a "signaling" state that interacts with another (unknown) protein. Thus, the bleached state lifetime is important for signal modulation,

† Supported by the National Institutes of Health grants GM45607 (DB) and GM37684 (EDG).

‡ Permanent address: National Institute for Occupational Safety and Health, 1095 Willowdale Road, Morgantown, WV 26505-2888

* Corresponding author. Telephone: 858-784-9612. Fax: 858-784-8896. E-mail: bashford@scripps.edu.

[1] Abbreviations used: PYP, photoactive yellow protein; MEAD, macroscopic electrostatics with atomic detail; OPLS, optimized potentials for liquid simulations, united-atom model; RHF, restricted Hartree−Fock; ESP, electrostatic potential (for charge fitting); HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-[2-ethanesulfonic acid]; MES, 2-[*N*-morpholino]ethanosulfonic acid; rms, root-mean-square.

10.1021/bi991513p CCC: $19.00
Published on Web 01/08/2000

and its pH-dependence implies electrostatic control of protonation states as a signal modulation mechanism.

The relatively simple photocycle of PYP involves some of the same elements, such as isomerization, conformational, and protonation state changes, as more complex photoactive proteins, such as bacteriorhodopsin (*4*, *10*, *13*). The ground-state structure of PYP has been determined by X-ray crystallography at 1.4 Å resolution (*6*), and the structure of the bleached state of the photocycle has been determined at 1.9 Å resolution by Laue techniques (*12*). An expression system for mutagenesis of PYP has been developed (*1*). These factors have made PYP an attractive case study for photoactivity in proteins.

The anionic character of the chromophore in the ground state of PYP is unusual in that burial would be expected to favor neutral states over charged states. Moreover, this deprotonated state of the chromophore persists under decreasing pH down to 2.7 (*5*), while the pK_a of the free chromophore is 9.0 (*3*). Thus, the negatively charged state is even more heavily favored in the particular buried protein environment provided by PYP than it would be in solution. Examination of the protein structure suggests that hydrogen bonding interactions with the carboxylic acid group of Glu46 and the hydroxyl group of Tyr42, and proximity to the positive charge of Arg52 could contribute to the stabilization of this negative charge (*6*). Resonance delocalization of the charge could also have a stabilizing effect. The protonated form of the solvent-exposed chromophore in the bleached state is what might be expected from the pK_a of the free chromophore or the chromophore in the unfolded state (9.0 in both cases (*3*)). Because the energetic terms involved in desolvating and stabilizing a buried charge are potentially large, PYP also presents a severe challenge for theoretical calculations of protein electrostatics.

The computational studies reported here include ab initio quantum chemical calculations of the charge distribution on the chromophore in the protonated and deprotonated states, and classical electrostatic calculations of the pH titration behavior of the protein's ionizable groups, and the resulting pH-dependence of the protein's stability. The electrostatic calculations use the model, macroscopic electrostatics with atomic detail (MEAD), in which the protein is modeled as a low dielectric material with embedded partial charges surrounded by a high dielectric medium, the atomic details of the protein structure are used to define the boundary and charge placement, and finite numerical methods are used to solve the resulting partial differential equations for the electrostatic potential (*14*). Models of this kind have been useful for the calculation of pH titration behavior in proteins (*15*, *16*), including the photoactive proteins, bacteriorhodopsin (*17*−*19*) and photosynthetic reaction center (*20*). The experimental studies reported here include determination of the chromophore protonation state vs pH by UV−visible spectroscopy for the active-site mutants, Y42F, E46Q, and R52A. We find that the computational results are in good agreement with the available experimental studies, including those first reported here, for the ground state of the wild-type protein and several mutants. The reinterpretation of the previous experimental data on the pH-dependence of the photocycle kinetics presented here suggests that the lifetime of the bleached state is modulated by specific pK_a shifts in the bleached state.

COMPUTATIONAL METHODS AND DATA

The method of calculating the pH titration behavior of ionizable groups in proteins using the MEAD (macroscopic electrostatics with atomic detail) model has been described previously (*17*) and is only briefly outlined here. It is based on the idea that the difference between the titration behavior of a chemical group in a protein and a group of the same kind in a small model compound arises from differences in the electrostatic environment. It is further supposed that these electrostatic effects can be treated by a model in which the protein is regarded as an object of dielectric constant ϵ_{in} containing embedded partial charges, the solvent is regarded as a continuum with the bulk dielectric properties of water, and that salt effects can be treated by Debye−Hückel theory. The dielectric boundary between protein and solvent is the molecular surface (*21*), which depends on the atomic radii and the radius of a solvent-sized spherical probe.

In this theory, three kinds of electrostatic terms influence the titrating groups: (1) The interaction of the titrating group's charges with the polarization of dielectric media and redistribution of mobile ions that they themselves create in the surroundings. This is called the Born self-energy term $\Delta\Delta G_{Born}$, and it tends to destabilize the charged state because the protein reduces the accessibility of solvent to the group. (2) The interaction of the titrating group's charges with nontitrating charges in the protein or model compound. This is the background interaction term, $\Delta\Delta G_{back}$. It includes the effect of permanent protein dipoles and the dipoles of the neutral forms of the ionizable groups on the titration of other ionizable groups. (3) Interactions between the charged forms of the titrating groups, which can be represented as a matrix, W_{ij} where i and j label the interacting sites. All three kinds of terms can be calculated by solving the Poisson−Boltzmann equation for the electrostatic potential generated by the protonated and deprotonated forms of each site in the protein and model compound, and forming the appropriate sums and differences of the products of these potentials with the appropriate charges.

Since the Born and background terms are independent of the protonation states of other sites, it is convenient to combine them with the model compound pK_a (pK_{mod}) to form the "intrinsic pK",

$$pK_{int} = pK_{mod} - (\Delta\Delta G_{Born} + \Delta\Delta G_{back})/2.303RT \quad (1)$$

which is the pK_a that the site would have if all other sites in the protein were held in their neutral form (*22*).

Given the set of pK_{int} values and site−site interactions, it is possible, in principle, to calculate the protonation fraction θ_i of each site i at any given pH by a Boltzmann-weighted average over all of the possible protonation microstates of the protein. Since a molecule with N titrating sites will have 2^N possible protonation states, and N is typically several tens or hundreds for proteins, this direct approach is rarely practical. However, several approximation methods have been developed to overcome this problem (*20*, *23*−*25*). The pH at which a plot of θ_i vs pH crosses the $\theta = 0.5$ line is designated as pK_{half} of site i (i.e., the pH at which i is half protonated). In most cases, the individual site titration curves resemble the usual Langmuir form, and the pK_{half} can be interpreted as the pK_a of the site. However, when strongly

interacting sites compete for a proton, it is possible to obtain θ_i vs pH curves that show multistep titration, or even reverse direction (*17, 23*). We shall see that such cases occur in the active site of PYP.

The pH-dependence of the free energy of a conformational change from A to B can be expressed in terms of integrals over the proton-binding isotherms of the conformers (*26*−*28*),

$$\Delta G_{AB}(\mathrm{pH}) - \Delta G_{AB}(\mathrm{pH_o}) = 2.302RT \int_{\mathrm{pH_o}}^{\mathrm{pH}} [Q_B(\mathrm{pH}') - Q_A(\mathrm{pH}')]\, d\mathrm{pH}' \quad (2)$$

where the Q are either the total charge or total number of protons bound in state A or B. Here, this formula will be used to obtain populations of alternative active-site hydrogen-bonding schemes, and to estimate the pH-dependence of the stability of PYP. Equilibration of hydrogen-bonding schemes can also be handled by methods that sample different proton positions simultaneously with the sampling of ionization states (*29*), but the above formalism is equivalent (*25*) and suitable when only a small number of different conformational states are being considered.

Parameters and Coordinates. The model compounds were taken as the *N*-formyl-*N*-methylamide derivatives of the amino acid residues, and their coordinates were taken directly from the corresponding atoms of the relevant protein structure. The pK_{mod} values were as follows: chromophore, 9.0; Asp, 4.0; Glu, 4.4; His, 6.8, for the $N^{\epsilon 1}$ tautomer or 6.4 for the $N^{\epsilon 2}$ tautomer; Tyr, 9.6; Lys, 10.4; Arg, 12.0; C-terminus, 3.8; N-terminus, 7.5. Calculations of the W_{ij}, $\Delta\Delta G_{Born}$, and $\Delta\Delta G_{back}$ terms was done using the multiflex program from the MEAD computer program suite (*17, 30*), which is available from our laboratory.[2] Calculation of individual-site titration curves and pK_a values was carried out using Gilson's clustering algorithm (*24*). The OPLS-UA (optimized potentials for liquid simulations, united-atom model) charge and radius set (*31*) was used for the MEAD calculations, since these parameters have been found to give the closest agreement to experimental small molecule solvation energies, among the empirical force fields in common use for protein calculations (*32*). The OPLS radii used were the $2^{5/6}\sigma$, which corresponds to the minimum of the Lennard−Jones potential. With this choice, all hydrogen atoms (which have zero radius in OPLS) are enclosed within the heavy atoms to which they are attached. Details of the preparation of OPLS-like parameters for the chromophore and other titrating groups are described below. In all MEAD calculations, the ionic strength was set to 0.02 M, which corresponds to the experimental conditions (*1*). Several different interior dielectric constant values were used, as described in the results section.

Calculations were based on the 1.4 Å resolution structure of the ground state of the photoactive yellow protein from *Ectothiorhodospira halophila* (*6*), and the 1.9 Å resolution structure of the bleached form (*12*).[3] Structural models of the site-directed mutants, Y42F and R52A, were prepared by deleting the appropriate side-chain atoms from the wild-type structure, and a model of E46Q was prepared by substituting an NH2 group for the Glu46 carboxylic-group oxygen atom nearest the chromophore. The SYBYL program (Tripos Associates Inc., St. Louis) was used to add initial hydrogen-atom coordinates to these structures assuming their usual protonation state at pH 7, unless otherwise noted. Protonation states of histidine residues were assigned iteratively (*33*). Crystallographically observed water molecules were depicted by the TIP3P model (*34*). Energy minimization of selected hydrogen atom positions in the active site was carried out using the all-atom Amber 3.0 force field (*35*), as implemented in SYBYL, with the distance-dependent dielectric constant, 2r. In contrast to more recent force fields (*31, 36, 37*), Amber 3.0 uses an explicit hydrogen-bonding term rather then relying on electrostatics and Lennard−Jones terms to reproduce hydrogen bonding, and so is more robust with respect to distance dependent reductions in the dielectric constant introduced in order to mimic screening without explicit bulk solvent. Energy minimization proceeded in two steps, first by simplex energy minimization of atoms with forces above 500 kcal/mol/Å, then by Powell conjugate gradient energy minimization until the rms gradient was less than 0.05 kcal/mol/Å. Only polar and aromatic hydrogen atoms were kept for continuum electrostatic calculations. Explicit water molecules were also removed prior to this stage. Minimum-energy positions for the titrating hydrogen atoms of the chromophore and Glu46 were determined in a systematic search of the hydroxyl-group dihedral angle. This minimization procedure was applied to both the wild-type and the mutant structures.

Chromophore Charge Parameters. Although both Amber and OPLS provide atom types that may be appropriate for the functional groups found in the chromophore of PYP, neither provides a directly corresponding set of charges. It was, therefore, necessary to develop atomic partial charges following two different strategies, one corresponding to Amber for molecular mechanics minimization calculations, and the other corresponding to OPLS for MEAD calculations. The quantum mechanical calculations for these purposes, described below, were carried out using the Gaussian94 program (*38*). The model molecule used was *N*-acetyl-*N*-methylamide cysteine with 4-hydroxycinnamic acid attached through a thioester linkage to the cysteine S^{γ} atom.

The Amber-style charges were based on electrostatic potential fitting to the results of a single-point STO-3G calculation on an extended molecular structure that was minimized using a molecular mechanics force field. Starting coordinates for the chromophore model molecule were taken from the corresponding atoms of the ground-state structure (in which the chromophore is in a somewhat extended trans state), and the SYBYL program was used to add hydrogen atoms and rotate dihedral angles so as to produce an extended side-chain conformation in which nonbonded side chain/backbone interactions were minimal. The bond types and topology were initially deduced by the SYBYL program and corrected by hand. The Amber 3.0 force field was then used within SYBYL to energy minimize the structure. During this minimization a preliminary charge set for the chromophore based on the parameters for half-cysteine and tyrosine was used. We then performed a Hartree−Fock calculation of the electronic structure using the STO-3G basis set, and charge

[2] The MEAD computer programs can be obtained over the Internet by anonymous FTP from the host, ftp.scripps.edu, in the directory, pub/electrostatics.

[3] The PDB file names for the ground and bleached forms are 2phy and 2pyp, respectively.

parameters were obtained by fitting to the resulting electrostatic potentials using the "ESP" procedure (*39, 40*), as implemented within Gaussian94. The resulting charges and geometries of the neutral form are similar to those of analogous groups in the Amber 3.0 parameter set (*35*), while for the ionized species, the formal charge is confined within the *p*-vinyl-phenol part of the chromophore without propagation onto the cysteine side chain and backbone parts of the residue (results not shown).

The OPLS approach to charge parameters relies more on the fitting of calculated vaporization and solvation thermodynamics to experiment than on fitting to quantum chemical calculations (*31*). We, therefore, took charge parameters for the neutral form of the chromophore directly from the OPLS atom-type tables for cysteine in its disulfide form and the tyrosine side chain. Carboxylic acid carbonyl charges were used for the carbonyl group of the thioester linkage. This left only the vinyl group's charges undetermined. Isolated aliphatic groups have zero partial charges in OPLS. In joining the vinyl group to phenol and carbonyl groups, we placed the residual charge of the phenol, from which a hydrogen atom is removed in linking to the vinyl, and the residual charge of the carbonyl, on the vinyl carbon atoms to which these groups are respectively linked. This procedure is similar to methods used in other OPLS-based studies in which parameters were estimated by assembly from existing group parameters (*31, 41*). A groupwise approach of this kind assumes that there is not a significant redistribution of charges among the functional groups of the neutral form, and that the charge of the sulfur atom in the thioester linkage is similar to that of a sulfur atom in a half-cysteine side chain (see Results and ref *42*).

For the charged form, such assumptions are probably not valid, so we modeled the difference between the neutral and charged forms on the difference between quantum chemical calculations for the neutral and charged species. Electronic structure calculations and geometry optimization at the RHF level using the 6-31G** basis set were carried out for both the protonated (neutral) and deprotonated (−1 charge) forms of the chromophore, and ESP atomic charges were determined for both using Gaussian94. Initial coordinates were taken from the STO-3G calculations described above. Geometry optimization was carried out in redundant internal coordinates. Dihedral angles along the $C^{\beta}-S^{\gamma}$ and two C−C bonds adjacent to the double bond were assumed redundant to prevent collapse of the side chain driven by nonbonded interactions. The ESP charges were adjusted slightly to maintain equal charges on equivalent atoms, such as the ortho carbons of the phenyl group. Backbone ESP charges for the ionized form were constrained to match those of the neutral form, and any excess charge was distributed over the side chain in proportion to Mulliken charges, as in the recent Amber charge model (*37*). An OPLS-like partial charge set for the charged state was produced by adding to the OPLS neutral-state charges the normalized difference between the adjusted ESP charges for the charged and neural forms for each atom.

Charge Parameters for Other Groups. For the backbone, the nontitrating side chains, and the neutral forms of His and Tyr, the OPLS charges were used directly in the MEAD calculations. For the neutral form of carboxylic acid groups, primary amines, and the arginine side chains, the hydrogen atoms were not explicitly represented because their correct placement is not generally known; instead, a smeared-charge modification to the OPLS parameter set was used (*33*). Except as noted below, the ionization of titratable groups was represented simply as the addition of a formal point charge to the neutral group. Accordingly, a charge of −1 was added to the carboxyl carbon of carboxylic acid groups, or the phenolic oxygen atom of tyrosine side chains, and a charge of +1 was added to the nitrogen atom of lysine side chains, the guanidinium-group carbon atom of arginine, or the C^{β} atom of histidine. The SYBYL program's hydrogen-building/energy-minimization facility produces all the required hydrogen-atom positions.

In the active site, explicit placements of crucial hydrogen atoms provided a more detailed charge model. In these cases, the OPLS charges for both the neutral and charged forms of Glu46, and the neutral form of Tyr42 were used directly. No OPLS charges are available for the charged form of Tyr. We developed them from RHF/6-31G** calculations by the same procedure described above for the development of OPLS-like charges for the deprotonated form of the chromophore.

EXPERIMENTAL METHODS AND MATERIALS

The charge state of the 4-hydroxy-cinnamic acid chromophore in the ground state of wild-type and mutant PYPs as a function of pH was monitored by UV−visible spectroscopy. Wild-type and mutant forms of PYP were generated, isolated, and purified as described previously (*1*). Concentrated stock solutions of wild-type PYP and mutants E46Q, Y42F, and R52A in 20 mM HEPES (*N*-(2-hydroxyethyl)piperazine-*N*′-[2-ethanesulfonic acid]) at pH 7.0 were first diluted with water to obtain concentrations corresponding to OD 5 at the wavelength of maximal absorbance λ_{max} (pH 7) (wild-type, 446 nm; Y42F, 458 nm; R52A, 450 nm; E46Q, 462 nm). These samples were then diluted 10-fold with solutions containing buffering compounds at 100 mM concentrations. The buffers used for the tested pH range were as follows: sodium acetate, pH-adjusted with glacial acetic acid for pH 3.0, 3.5, 4.0, 4.5, and 5.0; MES (2-[*N*-morpholino]ethanosulfonic acid), pH-adjusted with sodium hydroxide for pH 6.0 and 6.5; and HEPES, pH-adjusted with sodium hydroxide for pH 7.0, 7.5, and 8.0 UV−visible spectra of the resulting solutions were recorded on a Hewlett-Packard 8453 diode array spectrophotometer. The average fractional charge (afc) on the chromophore was calculated as:

$$\mathrm{afc}_{\mathrm{pH}x} = -\mathrm{OD}_{\mathrm{pH}x}/\mathrm{OD}_{\mathrm{pH8.0}}$$

assuming that the sample at pH 3.0 (OD = 0) represents a fully protonated chromophore and the sample at pH 8.0 represents a fully deprotonated chromophore. The OD values used in this calculation were those measured at the respective λ_{max} (pH 7) of the wild-type and mutant PYPs.

RESULTS

The present work consists primarily of calculations and experimental determinations of the hydrogen ion equilibrium of key groups in PYP, and its mutants, in both the ground state and the bleached state. However, direct experimental

determinations of the equilibrium properties of the bleached-state are made difficult by its limited lifetime. We, therefore, begin with a reinterpretation of previously published measurements of the pH-dependence of the recovery kinetics of the bleached state, in terms of the hydrogen-ion equilibrium in this state. In later subsections, we present the theoretical calculations of hydrogen ion equilibrium for the ground state and bleached states of both the wild-type and mutants. Experimental data from the present work, and from previous work, including the analysis of the recovery kinetics, are presented together with the corresponding calculations for comparison.

Reinterpretation of pH Profiles of Recovery Kinetics. Measurements of the pH-dependence of the rate of recovery of the ground state from the I_2 (bleached) state using time-resolved optical spectroscopy had previously been reported by Genick et al. (*1*). The recovery rate vs pH curve of the E46Q mutant clearly showed a single titration with a pK_a value near 8.0, while the pH profile of the wild-type was more complex, having a maximum at pH 8.0, and falling sharply under more alkaline or acidic conditions, indicating the involvement of two ionizable groups (Figure 1). In the original work, the E46Q data were analyzed using the expression,

$$k_{obs} = \frac{k_{max}}{1 + 10^{pK_1 - pH}} \quad (3)$$

with the fitted parameters $k_{max} = 280\ s^{-1}$ and p$K_1 = 8.0$ (See Figure 1A. We have dropped a k_{min} term that did not affect the analysis significantly.) The recovery data for wild-type PYP were originally analyzed using the expression,

$$k_{obs} = \frac{k_{max}}{(1 + 10^{pK_1 - pH})(1 + 10^{pH - pK_2})} \quad (4)$$

with the parameters k_{max}, pK_1, and pK_2 fit to 6.3 s^{-1}, 6.4, and 9.4, respectively. No relationships were assumed between the parameters pertaining to wild-type PYP and the E46Q mutant.

We begin our reinterpretation by accepting the original analysis of the simpler E46Q case, which implies that recovery from the bleached to the ground state requires that an unknown group (group 1) with a pK_a value of 8.0 must be deprotonated before recovery can proceed, and that for this deprotonated species the recovery rate is 280 s^{-1}. Recovery in the wild-type apparently requires one group to be deprotonated and another (group 2) to be protonated, as in the original analysis. Since the requirement of a protonated group disappears in the E46Q mutant, here we suppose group 2 to be Glu46. Since Gln is a conservative substitution for the protonated form of Glu, we further assume that the mutation does not change the intrinsic pK of group 1, and that the value of k_{max} (the recovery rate for the "correct" protonation species) is unaffected by the E46Q mutation. These should be reasonable assumptions if the rate-determining step for the protonation-state sensitivity of the recovery rate is essentially the same in the wild-type and the E46Q mutant.

If there is significant electrostatic site−site interaction between group 1 and Glu46, then the appropriate formula for fitting will be a specialization of the multiple interacting site formalism (*23*) involving the parameters k_{max} and pK_1,

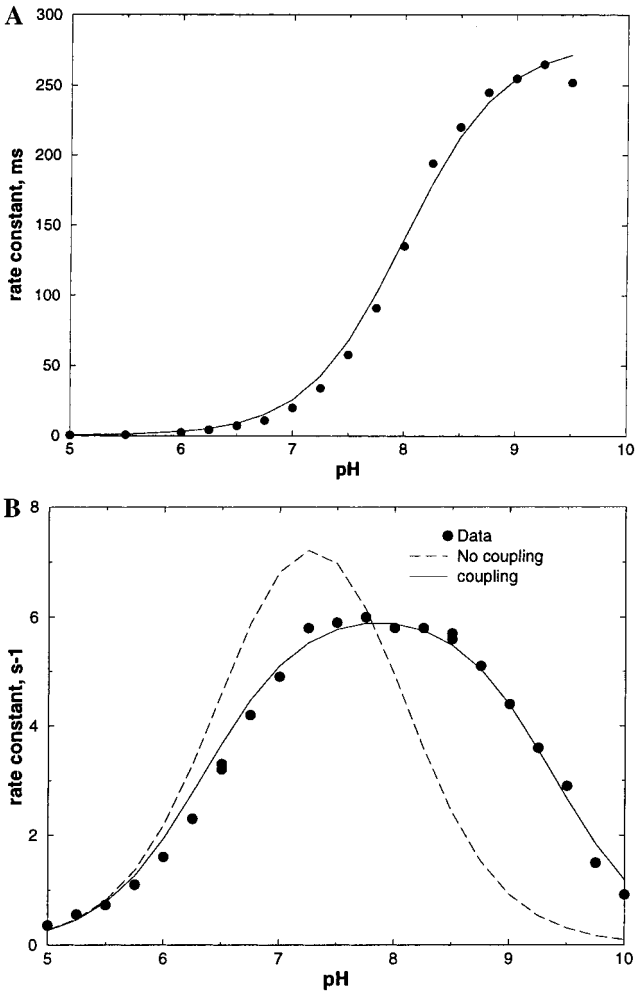


FIGURE 1: Rate of recovery of ground state from I_2 vs pH. Original data of Genick et al. (*1*) shown as points. (A) The E46Q mutant. Curve is eq 3 with parameters $k_{max} = 280\ s^{-1}$ and p$K_1 = 8.0$, by nonlinear least-squares fit to data points. (B) The wild-type. Solid curve: eq 5 with p$K_2 = 6.36$ and $W/2.3RT = 1.362$. Dashed curve: eq 4 with p$K_2 = 6.56$. In both curves, k_{max} and pK_i are taken from fit of (A) and remaining parameters adjusted by nonlinear least-squares fitting.

which are fixed from the results of the E46Q analysis, and two undetermined parameters, pK_2 and a site−site coupling, W. The manner in which W enters the formula depends on whether group 1 is assumed to be anionic or cationic. The only nearby cationic site is Arg52, and the mutation R52A does not change the shape of recovery vs pH curve (*1*). Anionic sites in the region include the chromophore, Tyr42, Tyr94, and Asp97. Therefore, we have assumed that group 1 is anionic. Using a notation in which [A1,A2] indicates the concentration of species with group 1 deprotonated and group 2 protonated, and so on, the appropriate formula for the recovery rate is,

$$k_{obs} = \frac{k_{max}[A1,A2H]}{[A1,A2] + [A1,A2H] + [A1H,A2] + [A1H,A2H]} = \frac{k_{max}}{10^{pH - pK_2 - W/2.3RT} + 1 + 10^{pK_1 - pK_2} + 10^{pK_2 - pH}} \quad (5)$$

If $W = 0$ (no interaction between sites) this formula is equivalent to eq 4.

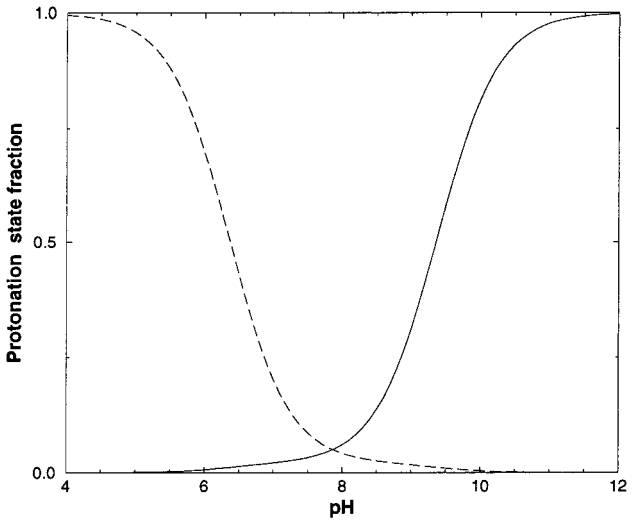


FIGURE 2: Fraction of protonation states required for recovery, as calculated with the parameters of Figure 1. Solid line: Fraction of group 1 in the deprotonated state. (Group 1 is presumed to be the chromophore, see Discussion) Dashed line: fraction of Glu46 (group 2) in the protonated state. The fraction of molecules with both groups in the state required for recovery lies below both curves, therefore, the maximum k_{obs} is much less than k_{max} (see eq 5).

A good fit to the experimental data is obtained by nonlinear least-squares fitting with the values pK_2 and $W/2.3RT = 1.362$ (solid curve in Figure 1B). The inclusion of a site−site interaction is essential to obtaining a good fit (compare the solid and dashed curves in 1B). In eq 5, pK_1 and pK_2 are formally analogous to the pK_{int} described earlier in the computational methods section, except that since groups 1 and 2 are the only sites explicitly treated as titrating, the interactions with all other ionizable groups in the protein are implicitly part of the "non-titrating" background. Apparent pK_2 or pK_{half} values corresponding to these groups must be obtained from the midpoint of the individual-site titration curves (Figure 2), which are calculated using a Boltzmann distribution over the four possible protonation states available to this two-site system (*23*). For Glu46 (group 2) the resulting pK_{half} is 6.37, a value nearly identical to the intrinsic pK_2 since group 1 is mostly neutral over the pH range in which Glu46 titrates. For group 1, the pK_{half} value is 9.35, which is equal to $pK_1 + W/2.3RT$, since Glu46 is mostly ionized over the pH range in which group 1 titrates. Examination of the titration curves of Figure 2 near pH 8, where the wild-type recovery rate is maximal shows that the fraction of group 1 that is deprotonated, and the fraction of Glu46 that is protonated are very small. This is why the maximal k_{obs} is approximately fifty times smaller than k_{max}.

In the Discussion section, we shall argue that group 1 is the chromophore, and that the titration behavior being probed by the pH-dependence of the kinetics is that of the bleached state. In that case, the results of the above interpretation of the kinetic data can be summarized as follows: in the bleached state of wild-type PYP, the chromophore titrates with a pK_{half} value of 9.35, and Glu46 titrates with a pK_{half} value of 6.37.

Calculated Charge Distributions. The RHF/6-31G** calculations for the neutral chromophore produced partial charges with significant differences from either the STO-3G calculations or the OPLS charges, as expected. The RHF/6-31G** calculations predict that upon deprotonation the

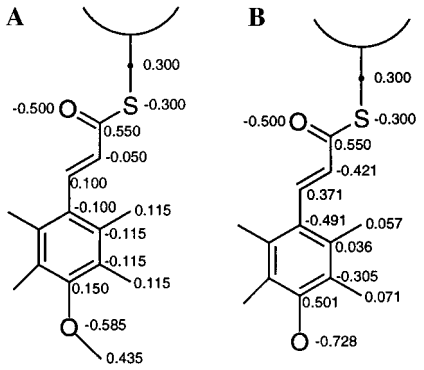
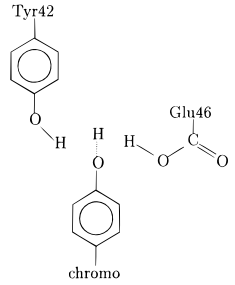


FIGURE 3: OPLS-compatible charges for the neutral, protonated form of the chromophore (A). Charges for the anionic form developed using RHF/6-31G** calculations of charge changes as described in text (B).

Scheme 1



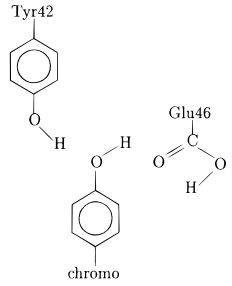
formal charge change of −1 is distributed as follows: phenyl group, −0.84; vinyl, −0.07; thioester linkage, including carbonyl and CH_2 groups, −0.03; backbone −0.06. The net propagation of the negative charge into the thioester linkage is quite small, while propagation into the backbone is slightly larger. In the preparation of OPLS-style charge parameters for the pK_a calculations, propagation of titrating charge into the backbone was disallowed to keep backbone charges consistent with the OPLS model. Propagation of titrating charge into the thioester linkage was neglected because it was small in the RHF/6-31G** calculations. The only atoms whose charges were considered to change upon ionization were those of the phenol and vinyl groups. The OPLS neutral-state charges, and the ionized-state charges that were prepared by distributing the 6-31G** charge changes using the procedure described in the Computational Methods section are shown in Figure 3.

Atomic partial charges for the side chain of tyrosine in the deprotonated state were developed by similar methods. The resulting charges were: C^{γ}, 0.041; C^{γ}, −0.290; C^{δ}, −0.071; H^{δ}, 0.067; C^{ϵ}, −0.263; H^{ϵ}, 0.066; C^{ς}, 0.405 and O^{η}, −0.754. The partial charges of tyrosine residue atoms not listed here are taken from the OPLS charge set.

Ground-State Titration Calculations and Experiment. Calculations of the titration behavior of all ionizable groups were carried out for the ground-state structure of the wild-type and three active-site mutants. Because the results are sensitive to the placement of hydrogen atoms in the active site, two different H-placement schemes were used in the wild-type calculations.

Scheme 1 favors a negative charge on the chromophore because the hydrogen atoms on Glu46 and Tyr42 are positioned to act as hydrogen bond donors to the chromophore's hydroxyl oxygen atom. This leaves the chro-

Scheme 2



mophore's hydroxyl hydrogen atom without a hydrogen-bond acceptor, so it is placed in a minimum-energy orientation below the plane of the figure.

H-placement Scheme 2 favors a negative charge on Glu46 rather than the chromophore. The chromophore's hydroxyl group is positioned to act as a hydrogen-bond donor to Glu46. This leaves the carboxylic acid hydrogen atom of Glu46 without a hydrogen-bond acceptor, so it is placed in a syn orientation on the oxygen atom farthest from the chromophore. Placement of this hydrogen atom in the anti position leads to results (not shown) that are nearly identical to those obtained from the syn placement. Calculations using a model in which this hydrogen is placed on the oxygen atom nearest the chromophore show that this placement is less favorable than placement on the far oxygen atom (results not shown). In both H-placement Schemes 1 and 2, the Tyr42 hydroxyl group was oriented so as to donate a hydrogen bond to the chromophore hydroxyl group, which is the only acceptor available to it.

In what follows, we first report calculations on single conformers, then we report a calculation in which the wild-type protein was allowed equilibrate between H-placement Schemes 1 and 2. Mutant calculations were done using only Scheme 1. The effects of assumptions concerning the dielectric properties of the protein were explored by carrying out calculations using several different ϵ_{in} values. Finally, we present the experimental measurements of chromophore titration and comparisons between theory and experiment.

Wild-type Calculations. The ordering of titration events in the Scheme 1 wild-type active site changes according to whether the interior dielectric constant is low or high (Table 1). For the lower values of ϵ_{in}, the chromophore titrates in the neutral to acid range (pK_{half} < 7), while Tyr42 and Glu46 have extremely high calculated pK_{half} values. Because of the lack of conformational flexibility in calculations of this type, these extreme calculated values should be interpreted as predictions that Tyr42 and Glu46 cannot be deprotonated without a significant rearrangement of the active-site geometry (*15*, *17*). This reflects the energetic difficulty of creating additional negative charges in the buried active-site region once the negative charge of the chromophore is present. The pK_{half} value of Arg52 is also perturbed upward by its interaction with this negative charge. As the dielectric constant of the protein is increased, the chromophore pK_{half} value shifts higher. When the interior dielectric constant is raised to 6.0, the θ vs pH curves (not shown) indicate that both the chromophore and Glu46 are protonated up to a pH of about 7.0, at which point one proton is lost and the two groups compete for the remaining proton, which is lost at pH 19. At higher values of the dielectric constant, the order of titration reverses as pK_{half} values tend toward their corresponding pK_{mod} values so that Glu46 becomes ionized more easily than the chromophore, and its negative charge shifts the chromophore pK_{half} to higher values.

Table 1: pK_{half} Values in the Ground State of the Wild-Type

interior dielectric (ϵ_{in})	calculated values			
	Tyr42	Glu46	Arg52	chromo
	H-placement Scheme 1			
1.7	44.0	68.8	18.9	−3.8
2	39.3	59.6	18.3	−2.1
3	30.4	42.3	17.1	2.6
4	33.8	25.7	16.6	5.9
5	29.7	21.7	16.2	6.6
6	26.9	16.3	15.7	19.2[a]
78.54	11.8	3.8	15.6	10.1
	H-placement Scheme 2			
2	38.6	17.1	12.2	59.5
4	25.6	8.3	14.9	36.0
6	21.2	6.9	14.8	28.1
78.54	11.8	3.3	15.7	10.5

[a] At $\epsilon_{in} = 6$, the titration curve of the chromophore exhibits a reversal of direction. The indicated pK_{half} is the highest-pH crossing of the $\theta = 0.5$ line.

The results are quite different for H-placement Scheme 2. In this case, Glu46 always has a lower pK_{half} than the chromophore, and the chromophore never becomes negatively charged in the neutral or acid pH range, no matter what dielectric constant is chosen. For low values of the protein dielectric constant (*2*−*4*), even the lowest pK_{half} within the active site is above 8.0, so there is no negative charge in the active site in the neutral to acid pH range.

We have also carried out calculations in which the pH-integration formalism described in the Computational Methods section is used to account for the freedom of the system's protonation states to follow either H-placement Scheme 1 or 2. Equation 2 is applied to the Q(pH) corresponding to the calculated titration curves for Schemes 1 and 2. The free energy difference obtained is used to calculate Boltzmann weights for averaging the individual-site titration curves. The reference state is chosen to be a hypothetical extreme high-pH state in which all sites are deprotonated. In this state, Schemes 1 and 2 have the same free energy because the protons that differentiate them are absent. In terms of eq 2, choosing pH_O to be extremely high, allows setting $\Delta G_{AB}(pH_O) = 0$.

The results of such a calculation using an interior dielectric constant of 2.0 are shown in Figure 4. Above pH 6.8, the free energy difference between H-placement Schemes 1 and 2 favors Scheme 1 and the chromophore is negatively charged. Below this pH, H-placement Scheme 2 is favored and the chromophore becomes protonated. Glu46 remains protonated throughout the 0−14 pH range. In other words, the calculation predicts that above neutral pH, there is one negative charge in the active site, and it resides on the chromophore where it is stabilized by the hydrogen bonds provided by H-bonding Scheme 1. Even though an alternative H-bonding arrangement that is more favorable to a negative charge residing on Glu46 is available (Scheme 2), Glu46 never becomes negatively charged. However, Scheme 2 appears to be better able than Scheme 1 to accommodate the addition of a proton that neutralizes the negative charge. Therefore, the system responds to decreasing pH by a switch

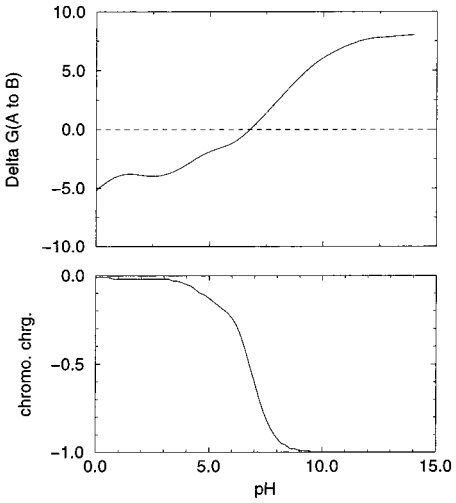


FIGURE 4: Results of multi-conformer calculation with $\epsilon_{in} = 2.0$. Top panel: Free energy difference between H-placement Scheme 1 and 2 calculated from eq 2. Bottom panel: Charge on the chromophore, from a Boltzmann-weighted average of A and B titration curves.

Table 2: Calculations Allowing H-placement Change

interior dielectric (ϵ_{in})	calculated values			
	Tyr42	Glu46	Arg52	chromo
1.7	44.2	68.8	19.0	6.2
2	39.4	59.5	18.4	6.8
4	25.6	14.9	15.3	35.7[a]
6	21.3	6.9	14.9	27.7
78.54	11.8	3.5	15.7	10.3

[a] Titration curve of the chromophore exhibits a reversal of direction. Indicated pK_{half} is highest-pH crossing.

from Scheme 1 to Scheme 2. Results for different values of the protein dielectric constant are shown Table 2. As in the single-conformer calculations for Scheme 1, increasing the dielectric constant reverses the order of titration for the chromophore and Glu46.

Mutant Calculations. The active-site pK_{half} values from calculations for the mutant structures are presented in Table 3. The mutation, E46Q, has very little effect on the chromophore or any other residues in the lower dielectric calculations, since the amide is an electrostatically conservative substitution for the protonated form of a carboxylic acid group. The mutations, Y42F and R52A, produce significant increases in the calculated pK_{half} values of the chromophore partly as a result of losing stabilizing interactions between the negative charge of the chromophore and the mutated side chain, and partly as an indirect effect involving interactions with other residues, particularly Asp97 (see below). The reversal of the predicted titration order of the chromophore and Glu46 at higher dielectric constants that was found in the wild-type calculations also occurs in the Y42F and R52A mutants.

The calculated pK_a values for all titrating groups within 15 Å of the active site, using an ϵ_{in} value of 2.0 and H-placement Scheme 1 are shown in Table 4. Other than the immediate active-site residues discussed above, Asp97 is the most notable for its effect on the chromophore. The W_{ij} value corresponding to its interaction with the chromophore in the wild-type is equivalent to 2.27 pK units. Therefore, in its negatively charged state, it significantly

Table 3: pK_{half} Values in the Ground State of Mutants

interior dielectric (ϵ_{in})	calculated values			
	Tyr42	Glu46	Arg52	chromo
		Y42F		
1.7		51.1	18.9	4.4
2		44.5	18.3	5.9
3		32.1	17.1	7.2
4		25.5	16.4	7.7
5		15.8	15.6	22.7[a]
6		14.0	15.0	21.0[a]
78.54		3.8	15.3	10.1
		E46Q		
1.7	46.4		18.9	−3.7
2	41.3		18.3	−2.0
3	31.8		17.1	3.0
4	26.9		16.6	5.9
5	24.0		16.3	6.6
6	22.1		16.1	7.1
78.54	11.1		15.3	9.2
		R52A		
1.7	40.5	65.0		4.2
2	36.4	56.5		5.8
3	28.6	40.3		7.6
4	31.5	25.3		8.2
5	28.4	21.1		8.6
6	25.9	18.4		8.9
78.54	12.3	4.3		10.7

[a] Titration curve of the chromophore exhibits a reversal of direction. Indicated pK_{half} is highest-pH crossing.

Table 4: pK_{half} of Sites within 15 Å of Chromophore (Ground State)

site	wild-type	Y42F	E46Q	R52A
Asp24	3.0	2.5	3.1	2.6
Asp34	−3.4	−3.3	−3.4	−3.3
Tyr42	39.3		41.3	36.4
Glu46	59.6	44.5		56.5
Asp48	5.7	5.4	5.8	5.7
Arg52	18.3	18.3	18.3	
Asp53	2.5	2.1	2.5	2.6
Lys55	12.6	12.6	12.6	12.7
Lys60	10.4	10.4	10.4	11.2
Asp65	2.0	1.5	2.0	1.8
chromo	−2.1	5.9	−2.0	5.8
Asp71	5.7	5.3	5.6	5.5
Tyr94	22.5	22.6	22.4	22.3
Asp97	2.6	1.2	2.5	1.8
Tyr98	13.0	13.0	13.0	13.6
Arg124	13.5	13.5	13.4	13.6

disfavors the negatively charged state of the chromophore. In the wild-type and in E46Q, the chromophore titrates in a lower pH range than Asp97, so this influence is not felt. But in the Y42F and R52A mutants, Asp97 is deprotonated over most of the pH range in which the chromophore titrates, and its negative charge shifts the chromophore titration to a higher pH range. The unusual shapes of the low pH tails of the Y42F and R52A curves in Figure 5 can be understood by comparison with calculations in which Asp97 is fixed in either the deprotonated or protonated form. Using $\epsilon_{in} = 2.0$, the chromophore pK_{half} value is calculated to be 3.3 if Asp97 is treated as fixed in its neutral state, and 6.0 if Asp97 is fixed in its charged state; and in both cases the curves have a conventional Langmuir-type form. The curves in Figure 5 reflect a shift from the Asp-deprotonated curve to the Asp-protonated curve as the pH falls below 2.0. Over most of the pH range in which the chromophore titrates in these

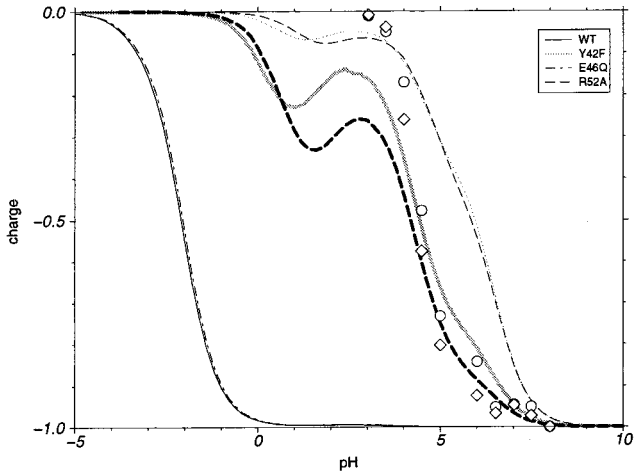


FIGURE 5: Chromophore titration in the ground state. Thin lines: calculations using $\epsilon_{in} = 2.0$; Thick lines: calculations using $\epsilon_{in} = 1.7$. Data points are from the absorption measurements described in Methods. Circles: Y42F data; Diamonds: R52A data. The unusual shape of the calculated curves is caused by interactions with Asp97 (see text).

mutants, the calculations with no special treatment of Asp97 are similar to the calculations treating Asp97 as charged.

Comparison of Calculations to Experiment. As for experimental pK_a values for the chromophore, the 446 nm absorption band of the wild-type is lost in very acidic solutions with an apparent pK_a value of 2.7 (*5*). This may reflect the acid-induced denaturation of the protein rather than the titration of the chromophore in an intact protein, so calculated results with pK_a values approximately equal to or less than 2.7 can be regarded as consistent with experiment. By this criterion, the wild-type calculations are in quantitative agreement with experiment only if $\epsilon_{in} \leq 3.0$ and H-placement Scheme 1 is used. The multi-conformer calculations using a low ϵ_{in} value are consistent with experimental evidence that the chromophore's pK_a in the protein is reduced relative to that of the free chromophore so that it is negatively charged at neutral pH, but the predicted pK_a value for the chromophore remains too high. As the ϵ_{in} value in the multi-conformer calculations is increased, even the qualitative agreement is lost.

Our initial calculations suggested that titration in the Y42F and R52A mutants may occur at significantly higher pH than in the wild-type. This motivated photoabsorption measurements to provide a test of the calculations. In the present work, experimental chromophore pK_a values near 4.5 were determined for the Y42F and R52A mutants (Figure 5). The calculations with $\epsilon_{in} = 2.0$ predict pK_{half} values approximately 1.5 pK units higher than experiment. A better fit to experiment for these mutants can be obtained by choosing an even lower interior dielectric constant $\epsilon_{in} = 1.7$ (Figure 5). Calculations with $\epsilon_{in} > 5$ fail to predict the correct protonation state for the wild-type even in the neutral pH range, while calculations with ϵ_{in} greater than 3 or 4 fail to predict the chromophore protonation state of the mutants at neutral pH. In general, rather good agreement with available experimental data for both the mutant and wild-type can be obtained in single conformer calculations, but only if a low value of ϵ_{in}, and a H-placement that favors a negatively charged chromophore is used.

Bleached-state Titration Calculations. Calculated pK_{half} values for all titratable groups within 15 Å of the active site in the bleached state are presented in Table 5. The calculations were done using $\epsilon_{in} = 2.0$, and so are directly comparable to the ground-state pK_{half} values presented Table 4. The only pK shift upon bleaching that would cause a change of protonation state at neutral pH is that of the chromophore, whose pK_{half} shifts upward by an amount that is more than sufficient to cause it to take up a proton. This occurs despite the close proximity of the positively charged guanidinium group of Arg52 because the chromophore's OH group can rotate away from the positive charge and toward the solvent. The pK_{half} values of Glu46, Tyr42, and Arg52 are all shifted downward, as would be expected from loss of interaction with the negatively charged form of the chromophore, but the shifts are not large enough to cause these residues to become deprotonated in the neutral range. In particular, Glu46 remains protonated because it is remains deep within the active-site pocket where solvent access is limited and the protein's hydrogen bonding network does not stabilize a carboxylate group in this position. Other sites have much smaller pK_{half} changes upon going from the ground state to the bleached state. The localization of significant calculated pK_{half} changes to the immediate environment of the chromophore is not surprising, given the localized character of the structural change upon bleaching, as determined by crystallographic methods (*12*). The above observations apply to the $\epsilon_{in} = 2.0$ calculations for wild-type PYP and all three mutants. In Table 6, results of calculations for several different values of ϵ_{in} are summarized for the active site residues. Increasing the protein dielectric constant decreases the calculated pK_{half} values, with the Glu46 pK_{half} falling into the neutral or acid range at $\epsilon_{in} = 6.0$ or 78.5, respectively.

The calculations are consistent with the experimental finding that the chromophore is protonated in the bleached conformation. However, recent proton uptake data (*43*) and our reanalysis of the kinetic data presented here suggest that in the bleached state of the wild-type, the chromophore pK_a

Table 5: pK_{half} of Sites within 15 Å of Chromophore (Bleached State)

site	wild-type	Y42F	E46Q	R52A
Asp24	2.2	2.2	2.4	2.3
Asp34	−3.4	−3.3	−3.4	−3.2
Tyr42	38.9		25.1	35.1
Glu46	25.1	25.4		24.3
Asp48	4.9	4.6	5.0	5.1
Arg52	13.9	14.3	13.8	
Asp53	2.0	2.0	2.0	2.6
Lys55	12.4	12.4	12.4	12.5
Lys60	8.7	8.7	8.6	10.8
Asp65	1.1	1.3	1.1	1.9
chromo	18.4	19.2	17.5	15.4
Asp71	5.1	5.2	5.1	5.3
Tyr94	22.4	22.5	22.3	22.4
Asp97	1.8	1.8	1.7	1.8
Tyr98	13.1	13.0	13.3	13.4
Arg124	12.5	12.3	12.6	12.6

Table 6: Calculated pK_{half} Values for Active Site of Bleached State

interior dielectric	calculated pK_{half}			
	Tyr42	Glu46	Arg52	chromo
2	39.1	25.1	13.9	18.4
4	26.2	10.2	15.1	18.0
6	21.7	8.0	16.3	14.0
78.54	11.2	3.7	15.5	9.9

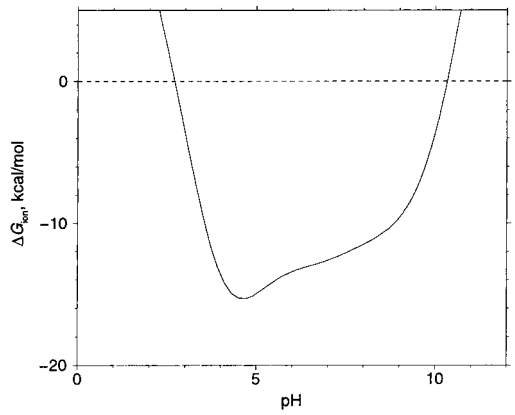


FIGURE 6: Calculated ΔG of folding, as a function of pH.

is near 9.4 and the Glu46 pK_a is between 6 and 7 (see Discussion). By this interpretation of experiment, the pK_{half} values for the chromophore and Glu46 calculated using a low interior dielectric constant are too high. Use of higher ϵ_{in} values brings the calculated pK_{half} of Glu46 closer to the experimental range.

Energetics of Unfolding and Photocycle. The pH-dependent free energy of folding from the ground state was calculated by using eq 2 and the pK_{int} and W_{ij} values from the ground-state calculations with $\epsilon_{in} = 2$ and H-bonding Scheme 1, described above, and assuming that in the unfolded state, the pK_{int} values revert to the corresponding model compound pK_a values, and the site−site interactions are negligible ($W_{ij} = 0$). The free energy scale was set by putting $\Delta G_{fold}(pH_O)$ to zero at $pH_O = 2.7$, which is the experimental point of acid-induced bleaching (*5*). This assumes that the acid-induced bleaching is caused by the unfolding of the protein coupled to the protonation of the chromophore, rather than by the protonation of the chromophore in an otherwise intact protein. The resulting plot of ΔG_{fold} vs pH (Figure 6) predicts that the protein will be stable over the mildly acidic to mildly alkaline range, but will undergo alkaline-induced unfolding in the pH 10−11 range. PYP does indeed show signs of denaturation in the alkaline range; for example, the 446 nm absorbance band is replaced by a 410 nm band at pH 11.7 (*42*).

At pH 7.0, the calculated value of ΔG is −12.6 kcal/mol. This calculated value can be compared to the experimental estimate of −10.8 kcal/mol which was obtained by spectral monitoring of PYP light absorption under urea denaturation conditions (*44*). At pH 3.4, the calculated value of ΔG is −8.4 kcal/mol, but the pH dependence is very steep in this part of the curve. An experimental estimate of unfolding free energy based on temperature-induced denaturation was $\Delta G(pH = 3.4) = -4.2 \pm 5.9$ kcal/mol (*45*).

DISCUSSION

A key feature of photoactive yellow protein is the presence of a buried negative charge in the active site. A computational study of such a situation often involves small differences between relatively large energetic terms. For example, the Born formula leads to the conclusion that the free energy of transferring a simple, monatomic, monovalent ion from a medium of dielectric constant 80 to a medium of dielectric constant 2.0 is on the order of 40 kcal/mol, which is equivalent to approximately 29 pK units (units of 2.303RT). The somewhat more extended geometry and charge distribution of charged groups in proteins, and proximity to the exterior, even for relatively buried sites, typically results in less extreme energy terms in MEAD-type calculations but $\Delta\Delta G_{Born}$ terms equivalent to 10 or more pK units are not uncommon for buried sites in proteins. This implies that stabilization of a buried charge requires favorable $\Delta\Delta G_{Born}$ terms or site−site interactions of a similar magnitude.

Given the approximate nature of the MEAD model, and the sensitivity of individual terms to conformational details, the extent of agreement between calculations and experiment obtained in the present study is encouraging. Further insight into the factors responsible for the stabilization of functionally important protonation states can be gained by a more detailed examination of the calculated energy terms.

The Ground State. The 4-hydroxycinnamic acid chromophore of PYP is known by experiment to be in its deprotonated form and to have a pK_a of 2.7 or less in the ground state even though the measured pK_a of this group in the liberated chromophore or the unfolded protein is 9.0 (*3*). Examination of the crystallographically determined structure suggested a qualitative explanation that the stabilization of this negatively charged form of the chromophore is provided by hydrogen-bonding interactions with Tyr42 and Glu46, and charge−charge interactions with Arg52. This picture not only requires a very large shift of the chromophore pK_a, it also requires that Glu46 remain protonated well into the alkaline range. This raises two quantitative questions: can sufficient favorable interactions be found to stabilize a negative charge in the active site, and if so, why does this charge reside on the chromophore, rather than the glutamic acid, which would normally have a lower pK_a?

The calculations give a positive answer to the first question when a low ϵ_{in} value is used, and when the H-placement scheme favors H-bonding to the chromophore oxygen atom (H-placement Scheme 1). The pK_{half} value calculated for the chromophore is below 2.7, while Glu46 and Tyr42 are predicted to be protonated. Insight into the second question is provided by calculations using an alternative H-placement scheme that favors H-bonding to the Glu46 carboxylate oxygen atom (H-placement Scheme 2). When using a low ϵ_{in} value, the Scheme 2 calculations predict that no negative charge can be formed in the active site in the normal pH range. Calculations that allowed a dynamic equilibrium between the two H-placement schemes with no built-in bias toward either provide insights into both questions. In the neutral to alkaline range, a negative charge can be sustained, this charge resides on the chromophore, and the favored H-placement is Scheme 1. However, in the acid range, the favored H-placement scheme changes to Scheme 2, and the chromophore becomes protonated.

Evidently, H-placement Scheme 1 provides more stabilization for a charge on the chromophore than Scheme 2 provides for a charge on Glu46. A more detailed examination of the terms going into the calculation of the pK_{int} values of these two sites confirms this, and also suggests a charge delocalization effect. Table 7 compares contributions to the pK_{int} value of the chromophore in H-placement Scheme 1, with contributions to the pK_{int} value of Glu46 in H-placement Scheme 2. The Born desolvation penalty for burying the chromophore charge is significantly less than that for the burial of the glutamate charge, apparently because of the

Table 7: Active-site pK_{int} Terms from $e_{in} = 2$ Calculations for the Wild-type Ground State

term	chromophore (H-plac. A)	Glu46 (H-plac. B)
pK_{mod}	9.0	4.5
ΔpK^a_{Born}	10.3	17.6
	Contrbs. to ΔpK^a_{back}	
Gly29	−0.4	−2.4
Tyr42	−3.8	−0.4
Asn43	1.1	2.6
Glu46	−5.9	−0.3
Ile49	−0.8	−2.1
Thr50	−2.5	−2.1
chromo	−0.5	−3.4
others	−0.9	−2.2
subtotal	−13.7	−10.3
ΔpK_{back}	−13.7	−10.3
pK_{int}	5.6	11.6

[a] The ΔpK correspond to the $\Delta\Delta G$ terms of eq 1, and are given in pK units ($-2.303RT$, where $T = 300$ K). ΔpK_{Born} is the effect of desolvation and ΔpK_{back} involves the difference between a titrating site's charged and neutral forms' interactions with the dipoles of other residues in the protein. (see Computational Methods). Contrbs. to ΔpK_{back} are interactions of named residues with chromophore or Glu46.

active site geometry, and because the negative charge of the chromophore is spread over a larger number of atoms. (A more delocalized charge, which approximately corresponds to a larger ion radius in the Born formula, costs less free energy to desolvate.) The background term of eq 1 is decomposed into contributions from the nontitrating or neutral-form charge distributions of selected residues in the "contrbs. to ΔpK_{back}" part of Table 7. This analysis shows that the negative charge of the chromophore is indeed stabilized by H-bonding interactions from Glu46 and Tyr42, and also by Thr50. For stabilization of a negative charge on Glu46, the largest interaction is provided by the chromophore in its protonated state, while Gly29, Ile49, and Thr50 also provide significant stabilization. Asn43 tends to destabilize negative charge at either site. The net effect is that the background terms can provide stabilizing effects on the order of 10 pK units for a negative charge at either site, but more stabilizing interactions are available for the chromophore.

Although the multi-conformer calculations account for the preference of the chromophore as the charged group, their prediction of the titration of the chromophore in the pH 6−7 range in is not in quantitative agreement with the experimental finding that the chromophore pK_a is 2.7 or less. It may be that a higher level of theory, such as quantum chemistry rather than classical electrostatics, is needed to accurately calculate the strength of the charged hydrogen-bonding interactions, or the effects of charge delocalization that may occur either across the hydrogen-bond with Glu46 or through interactions of the chromophore carbonyl group with the protein backbone.

The absorbance vs pH measurements presented here show that in the mutants Y42F and R52A, the chromophore pK_a is shifted upward relative to the wild-type, reflecting a loss of stabilization of the negative charge from these residues. On the other hand, the E46Q mutant does not shift the chromophore pK_a into the measurable range, presumably because glutamine is an electrostatically conservative substitution for the protonated form of glutamic acid.

The low-ϵ_{in} calculations for the mutants are in agreement with the experimental results. The pH-dependent way in which Asp97 interacts with the chromophore tends to amplify the effects of mutation on the titration of the chromophore. In the wild-type, the calculated titration of the chromophore (or perhaps the actual global unfolding of the protein) occurs in a low pH range in which Asp97 is neutral. If a mutation removes enough chromophore charge stabilization to move the chromophore titration into the pH range in which Asp97 is charged, then the charge−charge interaction will further increase the effective pK_a value of the chromophore. This theoretical prediction of the role of Asp97 in modulating the effects of other mutations could be tested by experiments on Asp97 mutants and double mutants involving Asp97 and either Tyr42 or Arg52.

The Bleached State and Recovery to Ground. In the initial report of measurements of the pH-dependence of the rate of the bleached state's return to the ground state, the data were analyzed using conventional formulas involving three parameters for the wild-type, and an independent set of two parameters for the E46Q mutant. Thus, there were two distinct physical models and five fitting parameters. In the present reinterpretation of these data, we use a somewhat more complicated formula, but it corresponds to a single physical model for both the wild type and E46Q and has a total of only four independent parameters. The new physical model suggests assignments of the titrating groups involved, their intrinsic and observable pK_a values, and the state of photocycle in which these groups have these pK_a values.

Specifically, the reinterpretation suggests that there is an intermediate conformer along the recovery pathway in which Glu46 must become protonated and another group, group 1, must become deprotonated before recovery can continue. The obvious candidate for group 1 is the chromophore, since it must certainly become deprotonated at some point along the recovery pathway, and it can easily be close enough to Glu46 to have the strong electrostatic interaction implied by the interaction parameter value, $W = 1.83$ kcal/mol. The most obvious candidate for the intermediate conformer that leads to the pH-dependent rate limitations, is then the bleached state itself. These two interpretations imply that in the bleached state, the chromophore and Glu46 have observable pK_a values of 9.35 and 6.37, respectively. The high pK_a value for the chromophore is consistent with the experimental finding that photobleaching under mildly acidic conditions is accompanied by the uptake of one proton, and the chromophore is the site that receives this proton (*3*, *46*). The pK_a of 6.37 for Glu46 in the bleached state is consistent with the recent finding that there is no proton uptake in the mildly alkaline pH range, and the interpretation of this as a net transfer of a proton from the chromophore to Glu46 (*43*). (This finding became known to us after the reinterpretation detailed above.) It is also consistent with the suggestion, based on the disappearance of a C═O stretching band attributed to Glu46 in the bleached state, that a proton is transferred from the chromophore to Glu46 (*47*).

The reinterpretation of the kinetic data presented here implies a mechanism for the control of the recovery rate in the photocycle by electrostatically modulated changes in the titration properties of the bleached state. This is of biological significance, since the bleached state is believed to act as

the "signaling" state in phototaxis. The structural changes between the ground and bleached states, which must be responsible for the change of titration properties, have been characterized crystallographically (*12*). In the bleached state, the hydrogen bonds between the chromophore and Glu46 and Tyr42 are broken as the chromophore swings out of its pocket and becomes more exposed to solvent. It is to be expected that the pK_a values of both the chromophore and Glu46 should move closer to those of the analogous free compounds in solution, so the protonation of the chromophore and the deprotonation of Glu46 in the bleached state can be understood in these terms.

The MEAD calculations for the bleached-state conformer using a low ϵ_{in} value predict that upon photobleaching, the pK_a value of the chromophore shifts into the extreme alkaline range. No other sites are predicted to have large pK shifts within the experimentally accessible pH range. Although the pK_a value calculated for Glu46 is shifted downward relative to the ground state, as one one would expect from the loss of interaction with the negative charge of the chromophore, it remains in the extreme alkaline range. The prediction that the chromophore becomes protonated is correct, but the calculated pK_a values for the chromophore and Glu46 are not in good agreement with the interpretations of the experimental results given above. A possible cause of the discrepancy is that the bleached state has more flexibility than the ground state and can undergo small, but energetically significant conformational changes coupled to changes in protonation. For example, a more open conformation of the pocket vacated by the chromophore would allow greater solvent access to Glu46, lowering its pK_a. The crystal structure, which was determined under acidic conditions (*12*), may then reflect a conformer more favorable to the protonated state of Glu46. Similarly, the chromophore, which has more conformational freedom in the bleached state, may change its position under alkaline conditions.

Stability and pH. On the basis of the calculated titration curve of the ground state, a simple model for the unfolded state, and an experimental pH point for acid denaturation, we have calculated the pH-dependence of the free energy of folding for ground-state PYP. The resulting stability curve (Figure 6) is in reasonable agreement with experimental estimates of stability at pH 3.4 and pH 7.0. The calculated curve also predicts alkaline denaturation in the pH range 10−11. There is some experimental evidence for alkaline denaturation, albeit at a slightly higher pH (*42*).

The Dielectric Constant of the Protein. In contrast to earlier spherical models (*22*, *48*), the MEAD model of protein electrostatics can be quite sensitive to the choice of the dielectric constant for the protein interior. This occurs because the location of partial charges and the protein−solvent boundary in MEAD are taken directly from the coordinates rather than assuming positions near the surface of a sphere, and because the $\Delta\Delta G_{Born}$ and $\Delta\Delta G_{back}$ terms are computed directly from the electrostatic model. In contrast, the earlier calculations typically used their model only for the site−site interaction terms W_{ij}, which were usually dominated by solvent dielectric screening since sites were placed near the surface. When a low value of ϵ_{in} is used in the MEAD model, it becomes unfavorable to place charges in buried locations, and the $\Delta\Delta G_{Born}$ terms for groups such as carboxylic acid and phenolic groups tend to shift their pK values upward. On the other hand, the low ϵ_{in} may make the hydrogen-bonding network around a buried charge much more favorable than a high ϵ_{in} value because the screening of these relatively short-range interactions is dominated by ϵ_{in} rather then the solvent dielectric.

It has been pointed out that a higher value of ϵ_{in} leads to much less sensitivity to details, and statistical comparisons of calculated and experimental pK_a values have shown that overall agreement is better with $\epsilon_{in} \approx 20$ than with low values of ϵ_{in} (*33*, *49*−*51*). Since most titrating sites in proteins are near the surface and exposed to solvent, and their pK_a values are near those of corresponding model compounds, models that tend to predict small magnitudes for the individual terms contributing to titration behavior will tend to perform well overall. However, the active sites of proteins often feature buried ionizable residues with highly perturbed pK_a values. A cautious use of a low-dielectric model may therefore be preferable to the use of a high-dielectric model that, while achieving good accuracy for ordinary sites, misses the special features of the active site (*33*). The present case is a dramatic illustration of this. The unusual protonation states of the chromophore and Glu46 in the ground state, which are central to the function of PYP, are predicted correctly only when a low value of ϵ_{in} is used.

The idea of a protein dielectric constant is problematic given the inhomogeneous and semi-flexible nature of proteins (ref *52* and references therein), particularly when one is interested in the electrostatic potentials felt by individual atoms. Even so, one can draw a qualitative correspondence between increasing protein flexibility and a higher ϵ_{in} value, and it is reasonable to prefer choices of ϵ_{in} that are more consistent with other theoretical or experimental determinations that also focus on the potentials felt by atoms. In comparisons of a MEAD-type model to molecular dynamics calculations for the effects of electronic and dipolar relaxation upon the energetics of charge interactions in cytochrome *c*, the best correspondence between the two theories was found for ϵ_{in} in the range 2−4, and relaxation effects were found to be smallest near the active site (*53*). This result is consistent with our present finding that calculations for the ground state give the best agreement with experiment when ϵ_{in} is low, particularly near the active site.

On the other hand, calculations using a higher ϵ_{in} value give better agreement with experiment for the bleached state. This could be interpreted as an indication of greater flexibility, thus a higher local dielectric constant, in the bleached state than in the ground state. However, the interpretation of experiment is less firm for the bleached state, and there is a possibility of subtle differences between the bleached-state structure determined by Laue crystallographic techniques at −12 °C and pH 4.8 (*12*), and the solution structure at room temperature. If so, multi-conformer calculations incorporating structural changes that stabilize the Glu46 charge might be more appropriate than raising the interior dielectric constant.

Concluding Remarks. The electrostatic calculations presented here provide a quantitative account of how a negative charge can be stabilized in the active site of ground-state PYP, and why that negative charge resides on the chromophore, rather than on Glu46. Charged hydrogen bonds, and charge delocalization emerge as important mechanisms

of stabilizing this charge; and quantum mechanical studies of these effects in PYP would be a logical next step. The necessity of a low protein dielectric constant to obtain correct predictions for the ground state supports the general practice of treating protein interiors as low-dielectric media, while remaining alert to the possibility of conformational changes whose neglect may lead to error. The effect of Asp97 on the outcome of the calculations for the wild type and mutants in the ground state suggests that mutation of Asp97 would have more significant effects on the chromophore in the mutants than in the wild type. Studies of Asp97 mutants and double mutations involving Asp97 would therefore be of interest. The reinterpretation of the bleached to ground-state recovery data presented here leads to proposed pK_a and site−site interaction values for the chromophore and Glu46 in the bleached state, that can be compared to theoretical calculations and to other experimental probes of ionization states along the photocycle.

ACKNOWLEDGMENT

We thank Nidhi Arora for help with figures, Velin Spassov for providing a computer program for multi-conformer titration calculations, and Jonathan Hirst and Zhongxiang Zhou for helpful discussions.

REFERENCES

1. 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.
2. Van Beeumen, J. J., Devreese, B. V., van Bun, S. M., Hoff, W. D., Hellingwerf, K. J., Meyer, T. E., McRee, D. E., and Cusanovich, M. A. (1993) *Protein Science 2*, 1114−1125.
3. Baca, M., Borgstahl, G. E. O., Boissinot, M., Burke, P., Williams, D. R., Slater, K. A., and Getzoff, E. D. (1994) *Biochemistry 48*, 14369−14377.
4. Hoff, W. D., Düx, P., Hård, K., Devreese, B., Nugteren-Roodzant, I. M., Crielaard, W., Boelens, R., Kaptein, R., Van Beeumen, J., and Hellingwerf, K. J. (1994) *Biochemistry 33*, 13959−13962.
5. Meyer, T. E. (1985) *Biochim. Biophys. Acta 806*, 175−183.
6. Borgstahl, G. E. O., Williams, D. R., and Getzoff, E. D. (1995) *Biochemistry 34*, 6278−6287.
7. Kim, M., Mathies, R. A., Hoff, W. D., and Hellingwerf, K. J. (1995) *Biochemistry 34*, 12669−12672.
8. Meyer, T. E., Tollin, G. Hazzard, J. H., and Cusanovich, M. A. (1989) *Biophys. J. 56*, 559−564.
9. Hoff, W. D., van Stokkum, I. H., van Ramesdonk, H. J., van Brederode, M. E., Brouwer, A. M., Fitch, J. C., Meyer, T. E., van Grondelle, R., and Hellingwerf, K. J. (1994) *Biophys. J. 67*, 1691−1705.
10. Imamoto, Y., Kataoka, M., and Tokunaga, F. (1996) *Biochemistry 35*, 14047−14053.
11. Genick, U. K., Soltis, S. M., Kuhn, P., Canestrelli, I. L., and Getzoff, E. D. (1998) *Nature 392*, 206−209.
12. 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. (1997) *Science 275*, 1471−1475.
13. Hellingwerf, K. J., Hoff, W. D., and Crielaard, W. (1996) *Mol. Microbiol. 21*, 683−693.
14. Bashford, D. (1998) In *The Encyclopedia of Computational Chemistry* (Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer III, H. F., and Schreiner, P. R., Eds.), Vol. 3. pp 1542−1548 John Wiley and Sons, Chichester.
15. Bashford, D., and Karplus, M. (1990) *Biochemistry 29*, 10219−10225.
16. Yang, A.-S., Gunner, M. R., Sampogna, R., Sharp, K., and Honig, B. (1993) *Proteins 15*, 252−265.
17. Bashford, D., and Gerwert, K. (1992) *J. Mol. Biol. 224*, 473−486.
18. Sampogna, R. V., and Honig, B. (1994) *Biophys. J. 66*, 1341−1352.
19. Engels, M., Gerwert, K., and Bashford, D. (1995) *Biophys. Chem. 56*, 95−104.
20. Beroza, P., Fredkin, D. R., Okamura, M. Y., and Feher, G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5804−5808.
21. Connolly, M. L. (1983) *Science 221*, 709−713.
22. Tanford, C., and Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 79*, 5333−5339.
23. Bashford, D., and Karplus, M. (1991) *J. Phys. Chem. 95*, 9556−9561.
24. Gilson, M. K. (1993) *Proteins 15*, 266−282.
25. Spassov, V. Z., and Bashford, D. (1999) *J. Comp. Chem. 20*, 1091−1111.
26. Tanford, C. (1970) *Adv. Protein. Chem 24*, 1−95.
27. Schellman, J. A. (1975) *Biopolymers 14*, 999−1018.
28. Yang, A.-S., and Honig, B. (1994) *J. Mol. Biol. 237*, 602−614.
29. Alexov, E. G., and Gunner, M. R. (1997) *Biophys. J. 74*, 2075−2093.
30. Bashford, D. (1997) In *Scientific Computing in Object-Oriented Parallel Environments* (Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W., and Tholburn, M., Eds.), Vol. 1343 pp 233−240, of *Lecture Notes in Computer Science*. ISCOPE97, Springer, Berlin.
31. Jorgensen, W. L., and Tirado-Rives, J. (1988) *J. Am. Chem. Soc. 110*, 1657−1666.
32. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) *J. Phys. Chem. 98*, 1978−1988.
33. Demchuk, E., and Wade, R. C. (1996) *J. Phys. Chem. 100*, 17373−17387.
34. Jorgensen, W. L., Chandrashekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−935.
35. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem. 7*, 230−252.
36. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comp. Chem. 4*, 187−217.
37. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, Jr., K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
38. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, J. R., Keith, T. A., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakrzewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S., Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C., and Pople, J. A. (1994). Gaussian 94. Gaussian, Inc., Pittsburgh, PA.
39. Singh, U. C., and Kollman, P. A. (1984) *J. Comp. Chem. 5*, 129−145.
40. Besler, B. H., Merz, Jr., K. M., and Kollman, P. A. (1990) *J. Comp. Chem. 11*, 431−439.
41. Jorgensen, W. L., and Severance, D. L. (1990) *J. Am. Chem. Soc. 112*, 4768−4774.
42. Hoff, W. D., Devreese, B., Fokkens, R., Nugteren-Roodzand, I. M., Van Beeumen, J., Nibbering, N., and Hellingwerf, K. J. (1996) *Biochemistry 35*, 1274−1281.
43. Hendriks, J., Hoff, W. D., Crielaard, W., and Hellingwerf, K. J. (1999) *J. Biol. Chem. 274*, 17655−17660.
44. Meyer, T. E., Yakali, E., and Cusanovich, M. A. (1987) *Biochemistry 26*, 418−423.
45. Van Brederode, M. E., Hoff, W. d., Van Stokkum, I. H. M., Groot, M.-L., and Hellingwerf, K. J. (1996) *Biophys. J. 71*, 365−380.

46. Meyer, T. E., Cusanovich, M. A., and Tollin, G. (1993) *Arch. Biochem. Biophys. 306*, 515−517.
47. Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N., and Yoshihara, K. (1997) *J. Biol. Chem. 272*, 12905−8.
48. Linderstrøm-Lang, K. (1924) *C. R. Trav. Lab. Carlsberg 15*, 1−29.
49. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994) *J. Mol. Biol. 238*, 415−436.
50. Antosiewicz, J., and Gilson, M. K. (1996) *Biochemistry 35*, 7819−7833.
51. Juffer, A. H., Argos, P., and Vogel, H. J. (1997) *J. Phys. Chem. 101*, 7664−7673.
52. Löffler, G., Schreiber, H., and Steinhauser, O. (1997) *J. Mol. Biol. 270*, 520−543.
53. Simonson, T., and Perahia, D. (1995) *J. Am. Chem. Soc. 117*, 7987−8000.

BI991513P