Molecular Basis for pH Sensitivity and Proton Transfer in Green Fluorescent Protein: Protonation and Conformational Substates from Electrostatic Calculations

C. Scharnagl, R. Raupp-Kossmann, and S. F. Fischer
Institut für Theoretische Physik T38, Technische Universität München, D-85747 Garching

ABSTRACT We performed a theoretical study to elucidate the coupling between protonation states and orientation of protein dipoles and buried water molecules in green fluorescent protein, a versatile biosensor for protein targeting. It is shown that the ionization equilibria of the wild-type green fluorescent protein-fluorophore and the internal proton-binding site E222 are mutually interdependent. Two acid-base transitions of the fluorophore occur in the presence of neutral (physiologic pH) and ionized (pH $>$ 12) E222, respectively. In the pH-range from $\approx$8 to $\approx$11 ionized and neutral sites are present in constant ratio, linked by internal proton transfer. The results indicate that modulation of the internal proton sharing by structural fluctuations or chemical variations of aligning residues T203 and S65 cause drastic changes of the neutral/anionic ratio—despite similar physiologic fluorophore pK_a s. Moreover, we find that dipolar heterogeneities in the internal hydrogen-bond network lead to distributed driving forces for excited-state proton transfer. A molecular model for the unrelaxed surrounding after deprotonation is discussed in relation to pathways providing fast ground-state recovery or slow stabilization of the anion. The calculated total free energy for excited-state deprotonation ($\approx$19 k_BT) and ground-state reprotonation ($\approx$2 k_BT) is in accordance with absorption and emission data ($\leq$5000 cm^{-1} or 24 k_BT).

INTRODUCTION

Green fluorescent protein (GFP), a bioluminescent protein from the jellyfish *Aequorea victoria*, is an exceptionally versatile and useful tool in cell biology and biotechnology (e.g., Misteli and Spector, 1997; Tsien, 1998). Fluorescence is emitted from the 4-hydroxybenzylidene-imidazolinone chromophore (referred to as Y66 in the following), which is formed by the cyclization of the internal tripeptide S65–Y66–G67 (amino acids are referred to by their one-letter code), followed by the 1,2-dehydrogenation of the tyrosine (Niwa et al., 1996). Wild-type GFP at room temperature has two major absorption maxima at $\approx$390 nm and $\approx$480 nm and a single emission peaking around 505 nm (Ward, 1981). The photophysical pattern has been explained by the existence of two ground-state conformations with different protonation of the fluorophore and excited-state proton transfer (ESPT) from the tyrosyl hydroxyl group to an internal acceptor, leading to the deprotonated chromophore from which actual emission occurs (Chattoraj et al., 1996; Lossau et al., 1996). A variety of mutants provides a broad range of altered spectroscopic properties (e.g., Heim et al., 1995; Ehrig et al., 1995; Heim and Tsien, 1996; Yang et al., 1998). Already, early studies (Ward, 1981; Bokman and Ward, 1981; Ward et al., 1982) noted the pH-sensitive spectral changes in the physiologic range (with a pK_a near 6), and recent investigations (Patterson et al., 1997; Kneen et al., 1998) suggest that pH shifts the equilibrium between neutral and anionic fluorophore. Single molecule spectroscopy (SMS) lead to the recognition (Dickson et al., 1997; Jung et al., 1998; Haupts et al., 1998) that, even in an equilibrium state, the protonation state of the fluorophore is not fixed over time but changes in response to the intrinsic dynamics of the protein. Crystallographic studies confirmed the existence of two ground-state conformations (Brejc et al., 1997; Palm et al., 1997): based on the correlation of structural elements with spectral characteristics of GFP mutants, it was rationalized that, in mutants with excitation maximum at $\approx$390 nm, the phenol in Y66 is uncharged, whereas it is in the charged phenolate form in mutants with excitation maxima at $\approx$475 nm. The side chain of Y66 is highly protected from solvent by the surrounding β-barrel and part of an extended hydrogen-bond network, that could act as charge relay (Brejc et al., 1997).

Photoactive proteins, like GFP and bacteriorhodopsin (e.g., Lanyi, 1997), have important scientific and technological applications. Despite the use of GFP as noninvasive pH-sensor (Kneen et al., 1998; Robey et al., 1998) in the living cell, the mechanism of GFP pH-sensitivity has not been investigated. It is therefore important to understand the structural basis of the response to charge displacements induced upon excitation, the pathway of the transferred proton, the role of protein heterogeneities, and their sensitivity to external pH. Our work reports the first theoretical study of the coupling between conformational and protonation substates in GFP basing on the recently solved x-ray structure for the monomeric wild-type protein (Brejc et al., 1997). The state energies at a given pH are obtained with numerical methods that permit a reliable description of protein energetics: a continuum model with atomic level of detail (Honig and Nicholls, 1995) was used for the evaluation of electrostatic energies, bonded and steric interactions

Received for publication 22 March 1999 and in final form 28 June 1999.
Address reprint requests to Dr. Christina Scharnagl, Technische Universität München, Institut für Theoretische Physik T38, James-Franck-Strasse 12, D-85747 Garching, Germany. Tel.: +49-89-2891-2731; Fax: +49-89-2891-2444; E-mail: scharnag@jupiter.t30.physik.tu-muenchen.de.

0006-3495/99/10/1839/19 $2.00

are calculated with an empirical force field (Karplus and Petsko, 1990). Conformational sampling has been restricted to hydroxyl dipole and water reorientations in the hydrogen-bonded network localized around the chromophore binding site. The calculations provide information at two levels: the statistical analysis yields ensemble averages like conformer occupancies and titration curves, the microscopic analysis reveals charge and dipole distributions in the available substates characterizing the protein's heterogeneity. The results quantify the relative population of neutral and anionic fluorophore, the dependence of the chemical equilibrium on environmental structural factors, and the presence of an internal proton acceptor, the driving forces and energetics of initial and final states in the deprotonation process, and point out pathways for equilibration after ESPT. The characterization of the protein reorganization is a prerequisite for subsequent analysis of barriers and time scales of the relaxation processes.

METHODS

Calculation of acid-base equilibria in proteins focus on how the protein environment alters the electrostatic potential at titrating sites. Several methods have been developed to include structural relaxation of the protein upon ionization changes and improve pK_a calculations. The adopted formalisms treat flexibility of polar and/or charged side-chain conformations in a single calculation of free energies and statistical averages (You and Bashford, 1995; Scharnagl and Fischer, 1996; Beroza and Case, 1996; Ripoll et al., 1996; Alexov and Gunner, 1997) or by computing averages for ensembles of structures (Antosiewicz and McCammon, 1996; Van Vlijmen et al., 1998), thereby neglecting probability weighting for the different conformations of each site (Beroza and Case, 1996). We apply the first formalism for multiconformational calculations and use predetermined likely minimum energy conformations for the residues under investigation, evaluate the energies for all accessible states of the protein, and determine the Boltzmann distribution of states for different pH values. Due to the large number of states that must be evaluated, conformational sampling has to be restricted. Hydroxyl reorientations have been shown (Alexov and Gunner, 1997) to be—due to their low barriers—statistically most important. Dependent on the ionization state of nearby residues, they can act as hydrogen-bond donors or acceptors and are able to align in response to changes in the local electrostatic field.

Our calculations base on the x-ray structural model obtained for monomeric wild-type GFP at pH 3.8 (Brejc et al., 1997; protein data bank entry: 1EMB). Only buried waters (criterium: solvent accessible surface <2 Å^2, probe radius 1.4 Å) are treated in atomic detail (crystal water no. 1–8, 10, 12–17, 19, 21–24, 27, 37, 51, 53, 54, 61, 70, 74), surface waters are treated as continuum solvent. We focus on the immediate surrounding of the fluorophore, therefore, only the tyrosine side chain of the 4-hydroxybenzylidene-imidazolinone chromophore (Y66) and the carboxylate of E222 are included as ionizable. No other titratable residues are found within Debye screening distance (8 Å). Y66 and E222 are part of an extended hydrogen-bonding network, including water W22, S205 as well as the aligning polar sites T203, S65 and waters W12, W19, and W27. They have been included as flexible sites. Conformational changes that alter heavy atom positions have only been included for T203. For this residue, the x-ray structure resolved two conformations in the protein with different OG1/CG2 position (dominant population A [$\chi 1 = 70°$], minor population B [$\chi 1 = -60°$] [Brejc et al., 1997]). With $\approx$5 kT (force field parameters from CHARMM [Brooks et al., 1983], 21-parameter set) the torsional barrier for the CA–CB bond is in the upper limit of the energy barries for the rotation of polar hydrogens (Alexov and Gunner, 1997). Because T203 is buried in the protein core, its rearrangement will not alter the protein shape. The hydrogen positions of neutral Y66 and E222, selected polar residues, and water molecules were assigned by a method focusing on the formation of favorable hydrogen bonds to any possible acceptor (distance <3.5 Å) using standard bond lengths and angles. All other polar hydrogens not included in the flexible set were constructed with the HBUILD routine of CHARMM (Brooks et al., 1983) and a subsequent 1000-step conjugated gradient minimization.

The state of each of the K multiconformation residues (titrating or polar) is represented by a subset of m_K elements $\vec{r} = (r_1, \ldots, r_{m_K})$, only one of which can be nonzero, indicating the occupied conformer. A statistical state $\vec{x}$ of the protein describes conformers differing in protonation state as well as polar hydrogen orientations: $\vec{x} = (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_K)$. For the inclusion of the A/B orientations of T203, care has been taken to combine orientations only if the acceptor state is also populated. The actual conformers are discussed in Results (see Table 2). The total number $m = \Sigma_{k=1}^{K} m_K$ of conformation was 39 for titration calculations (Y66: 3, E222: 5, T203: 6, S65: 2, S205: 2, W22: 8, W12: 6, W19: 2, W27: 5) making up $N = 64{,}800$ protein states. For the discussion of proton transfer reactions, five additional conformations for protonated W22 and one conformation for protonated S205 have been included ($N = 153{,}900$ protein states).

The free energy of a given protein state $\vec{x}^{(n)}$ is given by (Alexov and Gunner, 1997),

$$\Delta G(\vec{x}^{(n)}) = \sum_{i=1}^{m} x_i^{(n)} \left\{ \begin{array}{l} \gamma(i)[2.3\, k_B T(pH - pK_{solv}(i)] \\ + (\Delta G_{rxn}(i) + \Delta G_{pol}(i) + \Delta G_{nonel}(i) + \Delta G_{entr}(i)) \end{array} \right\} + \sum_{i=1}^{m} x_i^{(n)} \sum_{j=i+1}^{m} x_j^{(n)} \Delta G_{ij}. \quad (1)$$

The reference states for the calculation of energy contributions is the residue side chain in aqueous solvent with the

corresponding acid-base equilibrium described by $pK_{solv}(i)$. $\gamma(i)$ indicates the charge state (1 for bases; −1 for acids; 0 for nonionized and neutral sites).

Electrostatic contributions to the state energy (1 k_BT = 2.48 kJ/mol) are:

ΔG_{rxn} difference between reaction field energy in the protein and solution;
ΔG_{pol} interaction with polar and charged groups not included in the set;
ΔG_{ij} site–site interactions.

Conformation-dependent nonelectrostatic energy terms ΔG_{nonel} include Lennard–Jones steric interactions and bond angle (for water molecule) or dihedral angle (for amino acids) energies. The electrostatic interaction energies are finite difference solutions of the Poisson–Boltzmann equation using the program DelPhi (Sharp and Honig, 1990). The numerical error is in the order of 5–10% (Gilson et al., 1987). The dielectric boundary between protein and solvent was assumed to be the molecular surface. Regions inside were assigned a dielectric constant of 4, the solvent was modeled as continuum with a dielectric constant of 80, physiological ionic strength (0.15 M) and an ion exclusion layer of 2 Å. Two focusing steps resulted in a final resolution of 2.14 grid points/Å. Atomic radii have been taken from the PARSE parameter set (Sitkoff et al., 1994), atomic partial charges for protein residues form the CHARMM force field (21-parameter set). The partial charges for ground and excited state of the neutral and anionic 4-hydroxybenzylidene-imidazolinone were obtained from Mulliken charge analysis applying the semiempirical INDO/S SCF-CI method (quantum chemical program ARGUS [Thompson and Zerner, 1991]) and are summarized in Table 1.

The calculation of the excited state solvation energy for the chromophore in the protein was carried out with the ground-state reaction field. Due to the buried nature of the fluorophore, it showed only small variations as compared to the fully relaxed surrounding. Lennard–Jones parameters and angle bending potentials have been taken from the CHARMM force field (21-parameter set), and torsional potentials from Alexov and Gunner (1997). Especially, the 4.6 k_BT preference of the *syn*- over the *anti*-isomer for the neutral carboxylic acid of E222 has been included. An entropy correction $\Delta G_{entr} = k_BT \ln(n_I)$ has been introduced for ionizable residues, with the number n_I of states for each ionization state (Alexov and Gunner, 1997), thereby neglecting the iterative evaluation of the actual occupancy and introducing a maximal error in the order of ≈0.5 *pK* units.

The solution pK_a for glutamate was taken to be 4.3 (Stryer, 1988). For the 4-hydroxybenzylidene-imidazolinone chromophore in unfolded wild-type GFP, a value of 8.1 has been reported (Ward and Bokman, 1982). Drastic changes in the photophysical behavior of the peptide as compared to the intact protein indicate additional determinants, e.g., conformational equilibria. Therefore, rather than taking this value, we adjusted pK_{solv} for the fluorophore by a series of test titration calculations. A value of 6.5 is appropriate to correctly reproduce both, the experimentally determined ground-state equilibrium between protonated and unprotonated fluorophore ([YH]/[Y-] = 6 at pH 6.5 [Chattoraj et al., 1996; Haupts et al., 1998]) as well as the concentration of the major component of T203 ([T203A] = 85% at pH = 3.8 [Brejc et al., 1997]). This shift to a lower value is in line with values determined for comparable titratable side chains. The solution pK_a of tyrosine is 10.9 (Stryer, 1988), the extension of the phenolic conjugation system by an additional double bond in the cleaved 4-hydroxycinnamide chromophore of photoactive yellow protein reduces the pK_a to a value of 9.0 ± 0.3 (Baca et al., 1994).

TABLE 1 Partial charges* (in units of elementary charge e) for the 4-hydroxybenzylidene-imidazolinone chromophore (Y66)

atom	Y66-h#	Y66-h*#	Y66-#	Y66-*#
C1	0.30	0.20	0.25	0.10
N2	−0.45	−0.35	−0.44	−0.50
C2	0.52	0.22	0.49	0.35
N3	−0.24	−0.19	−0.24	−0.24
O2	−0.60	−0.42	−0.68	−0.71
CA2	0.11	0.15	0.08	0.16
CB2	−0.02	−0.17	0.00	−0.30
CG2	0.03	0.09	−0.05	0.17
CD1	−0.05	−0.06	−0.08	−0.11
CD2	−0.05	−0.06	−0.08	−0.11
CE1	−0.05	−0.02	−0.10	0.00
CE2	−0.05	−0.02	−0.10	0.00
CZ	0.15	0.18	0.21	0.24
OH	−0.50	−0.45	−0.86	−0.65
HH	0.30	0.30		

*Charges were evaluated using semiempirical quantum chemical INDO/S SCF-CI calculations and Mulliken analysis (program ARGUS [Thompson and Zerner, 1991]) separately for both protonated and unprotonated fluorophore in the x-ray geometry (Brejc et al., 1997). Charges for atoms CA and CA3 are 0.10, charges for the hydrogens (attached to CD1, CD2, CE1, CE2, CB2) are 0.08 (see Fig. 1).
#Y66-h and Y66-h* (Y66- and Y66-*) refer to the neutral (anionic) charge state in electronically ground and excited state, respectively.

The Förster cycle (Stewart, 1985) was used for the determination of the free energy difference between electronically excited neutral and anionic fluorophore in aqueous solvent. The 4200-cm^{-1} bathochromic shift of the absorption spectrum upon ionization corresponds to an increase in acidity of the electronically excited phenolic compound by 8.8 units, a value comparable to the 6 pK_a units determined by flash photolysis measurements of phenol (Stewart, 1985).

The charge and geometry model for protonated R–OH_2^+ compounds (water W22, serine S205) have been determined previously by quantum chemical calculations (Cometta-Morini et al., 1993) (partial charges $\delta_{q_O} = 0.58$, $\delta_{q_{H(R)}} = 0.14$, apex angle 62.2°, $pK_{solv} = -1.74$). They are used for the determination of free energy differences in the proton translocation steps.

pH-Dependent site occupancies were determined as statistical averages over all protein states,

$$\langle x_i \rangle = \sum_{n=1}^{N} x_i^{(n)} g(n) \Big/ \sum_{n=1}^{N} g(n) \qquad (2)$$

$$\text{with} \quad g(n) = \exp(-\Delta G(\vec{x}^{(n)})/k_B T).$$

The pK_a value of a titratable residue is the pH with an occupancy of 0.5 for the ionized form. Besides the calculation of these macroscopic properties, we sort the state vectors according to increasing energy and analyze further individual state properties (e.g., conformational dependencies, absorption spectra, energy gaps for proton transfer) of the lowest energy states.

UV-/Vis-absorption spectra were calculated with the semiempirical INDO/S SCF-CI method (quantum chemical program ARGUS [Thompson and Zerner, 1991]). Excited states are described as singly excited electron configurations using a set of 20 occupied and 20 virtual molecular orbitals. The chromophore has been truncated at CA and CA3 (see Fig. 1), which were modeled as methyl groups. As pointed out already by Wachter et al. (1998), there is no experimental evidence for the heterocyclic ring to be protonated in any of the structures solved to date. Several other findings also contradict the presence of a second proton on the chromophore: pH-titration shows only one isosbestic point (e.g., Terry et al., 1995; Patterson et al., 1997) and the amide hydrogen of the Y66 peptide is removed upon dehydration in the cyclization process leading to fluorophore formation (Niwa et al., 1996). In accordance with these findings, nitrogen N2 remains unprotonated in our analysis. This is at variance with indirect conclusions from quantum chemical calculations (Voityuk et al., 1998), which are, however, based on a modeled chromophore structure. Our calculated longest wavelength absorptions are at 373 nm (oscillator strength $f = 0.38$) for the isolated neutral fluorophore and 468 nm ($f = 0.83$) for the isolated anion—both calculated with the x-ray geometry. Within the error limits of the method, these values correspond to the experimental values,

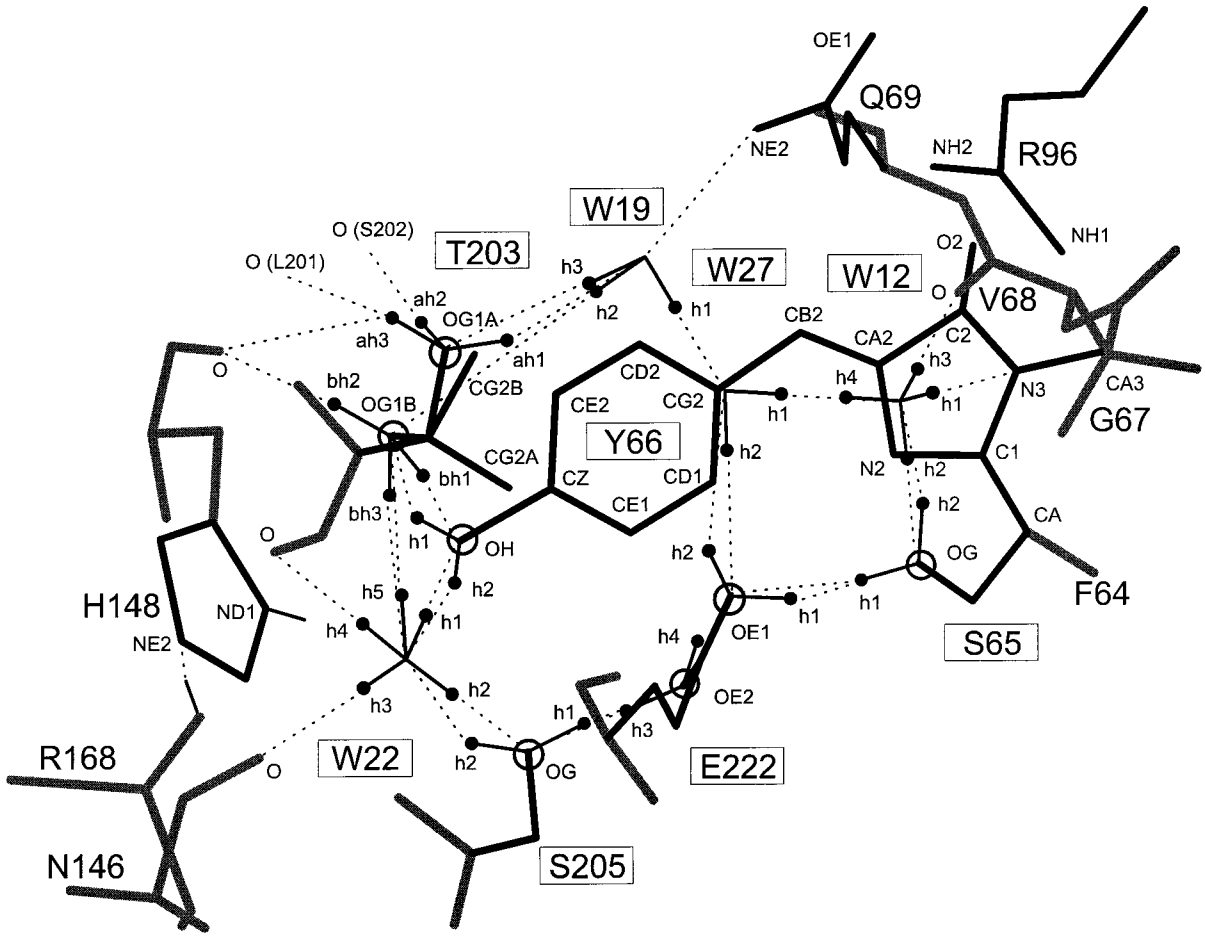

FIGURE 1 Hydrogen-bond net around the fluorophore in monomeric wild-type GFP. The figure shows the heavy atom positions as revealed by x-ray analysis (Brejc et al., 1997; protein data bank entry: 1EMB) and the polar hydrogen positions modeled to optimize hydrogen-bonding to any possible acceptor within 3.5 Å (compare also Methods and Table 2). Conformations for T203 differ in both position of OG1/CG2 (A and B) and orientations of polar hydrogens. Hydrogen bonds are marked with dashed lines.

especially the experimentally determined intensity ratio of 0.5 (Chattoraj et al., 1996) or 0.38 (Haupts et al., 1998) is reproduced (calculated 0.46)—without invoking an additional protonation of the heterocyclic ring. To calculate influences from surrounding amino acids, the polar residues have been included in a supermolecule approach. The conformations were taken from the accessible states.

RESULTS

Characterization of the local hydrogen bond network

The 4-hydroxybenzylindene-imidazolinone chromophore of GFP (referred to as Y66) and the carboxylate group of E222 are highly protected from solvent inside the hydrophobic core of the protein (e.g., Brejc et al., 1997). Within Debye screening distance (≈8 Å), there are no other titratable amino acid side chains. Additionally, x-ray atomic models for the protein at different pH (Yang et al., 1996; Ormö et al., 1996; Brejc et al., 1997; Palm et al., 1997; Wachter et al., 1997) as well as circular dichroisms (Kneen et al., 1998) and Fourier transform infrared (FTIR) spectroscopy (Van Thor et al., 1998) indicate, that protein conformational changes upon ionization of the chromophore are localized to its binding region. Therefore, we restrict our analysis to the coupling of the ionization states of Y66 and E222 and the concomitant side chain reorientations in the internal hydrogen-bond coupled network (W22, S205, S65, T203, W12, W19, W27). For the inclusion of side chain conformational flexibility in the calculation of the protonation equilibria of Y66 and E222, likely minimum energy hydrogen positions for hydroxyl groups (threonine, serine, and tyrosine side chains), the carboxylic acid of E222 and water molecules have been predetermined (see Methods). The selected positions are summarized in Table 2 and Fig. 1.

As revealed by the x-ray structure, the side chain of T203 has two conformations: A, $\chi 1 = 70°$ and B, $\chi 1 = -60°$. Therefore, conformers for T203 differ in both, position of OG1/CG2 and orientation of polar hydrogens. Water W19 and the backbone oxygens of S202 and L201 are possible hydrogen-bond acceptors for OG1A (hydrogens ah1, ah2, and ah3), whereas the phenolic oxygen of Y66, the backbone oxygen of H148, and water W22 are acceptors for OG1B (bh1, bh2, and hb3). In contrast, OG1A can accept a hydrogen bond from W19, OG1B from Y66, and water W22. The hydrogen-bond pattern for Y66 involves interactions to T203B (h1) and water W22 (h2). S205 is linked either to E222 (h1) or water W22 (h2). In a similar way, S65 can act either as hydrogen-bond donor or acceptor with E222 (h1) and water W12 (h2). The carboxylic acid side chain of E222 can be protonated either on OE1 (h1, h2) or OE2 (h3, h4) with the energetically favored *syn*-isomers h2 and h4. In the monomeric wild-type structure, NE2 of the side chain of HIS148 is a hydrogen-bond acceptor from the backbone amide of R168. This fixes the protonation site to be ND1, which is—due to its long distance (3.4 Å)—no hydrogen-bond partner to Y66.

TABLE 2 Residues and predefined polar hydrogen positions h_i for multiconformation calculations

Donor D*	Acceptor A*	Hydrogen h_i	Distance $h_i \rightarrow$ A/Å
Y66-OH	T203-OG1B	h1	2.007
	W22-OH2	h2	1.617
T203-OG1A	W19-OH2	ah1	2.497
	S202-O	ah2	2.971
	L201-O	ah3	2.946
T203-OG1B	Y66-OH	bh1	1.804
	H148-O	bh2	1.899
	W22-OH2	bh3	2.465
W22-OH2#	Y66-OH	h1	1.603
	S205-OG	h2	1.635
	N146-O	h3	2.201
	T203-O	h4	2.022
	T203-OG1B	h5	2.423
S205-OG	E222-OE2	h1	1.645
	W22-OH2	h2	1.781
E222-OE1	S65-OG	h1	2.128
	W27-OH2	h2	2.561
E222-OE2	S205-OG	h3	1.643
	E222-OE1	h4	2.318
S65-OG	E222-OE1	h1	2.108
	W12-OH2	h2	1.748
W12-OH2#	G67-N3	h1	2.603
	S65-OG	h2	1.708
	V68-O	h3	2.474
	W27-OH2	h4	2.067
W19-OH2	W27-OH2	h1	1.641
	T203-OG1B	h2	2.925
	T203-OG1A	h3	2.447
W27-OH2	W12-OH2	h1	2.067
	E222-OE1	h2	2.394
	W19-OH2	h3	1.641

*D and A denote the hydrogen-bond donor and acceptor atoms (compare Fig. 1), respectively.
#The water orientations selected among the listed hydrogens are shown in Fig. 2.

Among the buried waters, W22 develops the largest number of hydrogen bonds. The connections to Y66 (h1), S205 (h2), and T203B (h5) can act either as donor or acceptor relations. Additionally, the backbone oxygens of T203 and N146 can accept a hydrogen bond from W22 (hydrogens h3, h4). Among the hydrogens listed above, the water molecule orientations that develop at least two hydrogen bonds with neighbors have been selected. Water W12 also shows a high level of connectivity. Four polar hydrogens have been placed according to the four partially negative charged atoms in the neighborhood: h1 points toward N3 of the chromophore; h2 and h4 donate hydrogen bonds to S65 and

water W27, respectively; h3 realizes a hydrogen bond to the backbone oxygen of V68. Selected water molecule orientations for W22 and W12 are summarized in Fig. 2.

The hydrogen-bond pattern of water W27 involves donor/acceptor relationships to water W12 (h1), E222 (h2), and water W19 (h3). The fixed orientation of the side chain of glutamate Q69 forces buried water W19 to accept the hydrogen bond from NE2 and to orient its two hydrogens to develop hydrogen bonds to the oxygen of water W27 (h1) and the side chain hydroxyls of T203A (h2) and T203B (h3). The positions for h2 and h3 differ only by 0.32 Å, but show different distances to their acceptors. As a consequence of the restricted W19 orientation, the orientation of W27 is also locked as shown in Fig. 1 and confirmed by the sampling procedure.

Ground-state protonation equilibria

The calculation of statistical averages over the protonation states of Y66 and E222 in the GFP protein has been carried out including conformational flexibility of the protons in the hydrogen-bond coupled cluster around the chromophore binding site. Table 3 summarizes the conformational-dependent energy contributions to the state energies $\Delta G(\vec{x})$ (compare Eq. 1). The electrostatic interaction energy ΔG_{pol}, with polar sites not included in the cluster, is dominated by the interaction with the positively charged side chain of R96, which is largely favorable for ionized chromophore and E222 and discriminates between the orientations of W12

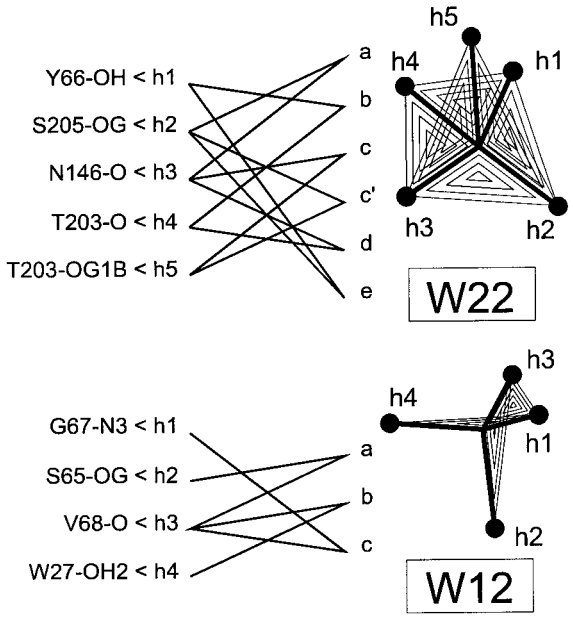

FIGURE 2 Orientations of water molecules W22 (a–e) and W12 (a–c) sampled in the multiconformation calculations (compare also Fig. 1 and Table 2). Short hydrogen bonds connect W22 either to the fluorophore (Y66) (orientations W22-b/e) or to S205 (conformations W22-a/c′/e). Longer hydrogen bonds develop to the backbone (W22-d) and the side chain of T203 (W22-c/c′). W12 flips between perpendicular (W12-a) and parallel (W12-b/c) orientation relative to the chromophore plane.

and the hydroxyl positions of S65. Due to the buried nature of the sites, the reaction field energy ΔG_{rxn} is dominated by the large desolvation penalty, reflecting the reaction field stabilization in solution. The steric and bonding contributions ΔG_{nonel} to the state energies reveal the unfavorable Lennard–Jones repulsion, that destabilizes the hydroxyl positions with the shortest hydrogen bonds, e.g., between E222 and S205 or orientations of W22 forming a hydrogen bond to Y66. The torsional energy for the proton orientations to E222 includes also the 4.6 k_BT preference (Alexov and Gunner, 1997) of the *syn*- (h4, h2) over the *anti*-isomers (h3, h1). The largest stabilizing site–site interactions are hydrogen bonds between the anionic fluorophore and T203 (bh1) or water W22 (b, e).

Evaluation of the average occupancies for ionized Y66 and E222 in the pH-range from 1 to 15 results in the calculated titration curves shown in Fig. 3 *A*. Like the experimentally determined concentration of anionic fluorophore (Bokman and Ward, 1981), the calculated one deviates significantly from the simple case described by a Henderson–Hasselbalch equation. Two acid-base transitions are calculated to occur at pH values of 6.8 and 13.4 (as compared to the fluorescence experiment: ≈6 and ≈12). This titration behavior indicates a competition between the acidity of the E222 carboxylic acid and the phenolic core of the fluorophore. Each side chain shows two protonation equilibria with mutually interdependent pK_a values. The acid-base transitions of two interacting residues can be described by four equilibrium constants (Fig. 4). The four pK_a values are not independent, but rather bound by the state function property of the underlying free energies, $pK_{a1} + pK_{a2} = pK_{a3} + pK_{a4}$. The fractional occupancies of the four states are (with $x_i = \mathrm{pH} - pK_{ai}$),

$$f_{Y\text{-}/EH} = \frac{10^{x_1}}{1 + 10^{x_1} + 10^{x_3} + 10^{(x_1+x_2)}}, \tag{3a}$$

$$f_{YH/EH} = \frac{1}{1 + 10^{x_1} + 10^{x_3} + 10^{(x_1+x_2)}}, \tag{3b}$$

$$f_{YH/E\text{-}} = \frac{10^{x_3}}{1 + 10^{x_1} + 10^{x_3} + 10^{(x_1+x_2)}}, \tag{3c}$$

$$f_{Y\text{-}/E\text{-}} = \frac{10^{(x_1+x_2)}}{1 + 10^{x_1} + 10^{x_3} + 10^{(x_1+x_2)}}. \tag{3d}$$

The three independent pK_a values have been obtained by a fit of the calculated pH-dependent occupancies of the ionized forms of Y66 or E222,

$$f_{Y\text{-}} = f_{Y\text{-}/EH} + f_{Y\text{-}/E\text{-}}, \tag{4a}$$

$$f_{E\text{-}} = f_{Y\text{-}/E\text{-}} + f_{YH/E\text{-}}. \tag{4b}$$

In the nomenclature introduced here, protein states are characterized by the corresponding ionization states of the key residues Y66 and E222 averaged over all possible orientations in the hydrogen-bond net; e.g., Y-/EH denotes a state

TABLE 3 Energy terms (A) and selected fluorophore interactions (B)

(A)				(B)				
					$\Delta G_{ij}/k_BT$			
	$\Delta G_{pol}/k_BT$	$\Delta G_{rxn}/k_BT$	$\Delta G_{nonel}/k_BT$		Y66-	Y66-*	Y66-h1	Y66-h2
Y66-	−16.82	30.86	0.00	T203-ah3	1.62	1.08		
Y66-h1	−4.61	7.90	2.08	T203-bh1	−10.08	−6.42	1.39	−1.37
Y66-h2	−4.89	7.84	4.83	T203-bh2	4.29	2.65		
Y66-*	−23.08	24.84	0.00	T203-bh3	−6.67	−4.61	0.75	
Y66-h1*	−2.62	4.76	2.08	W22-a	5.06	3.42	1.13	−1.94
Y66-h2*	−2.90	4.71	4.83	W22-b	−9.84	−6.88	−2.52	8.05
T203-ah3	−1.64	3.39	0.19	W22-c	−2.19	−1.55		−0.81
T203-bh1	0.52	3.56	2.05	W22-c′	−0.72		0.61	
T203-bh2	−2.10	3.15	0.31	W22-d	3.94	2.55	1.18	−2.69
T203-bh3	0.01	3.40	0.58	W22-e	−8.72	−6.01	−2.57	8.79
W22-a	−0.78	4.91	1.52	S205-h1				
W22-b	−0.95	6.10	4.12	S205-h2	−1.22	−0.71		
W22-c	−1.58	5.13	−1.31	E222-	13.26	12.33		−1.62
W22-c′	0.72	4.62	1.45	E222-h3		1.26		
W22-d	−3.01	6.12	−1.31	E222-h4	−1.05			
W22-e	1.29	5.08	6.95	S65-h1	−1.46	−0.52	−1.40	−1.33
S205-h1	0.41	3.70	4.08	S65-h2		−1.46		
S205-h2	0.21	3.70	0.33	W12-a	1.08	1.81		
E222-	−14.97	29.10	0.00	W12-b	−0.50	0.75	−0.97	−0.99
E222-h3	−3.55	6.13	7.15	W12-c	0.50	−0.71	0.62	0.59
E222-h4	−2.23	4.44	−0.34					
S65-h1	0.40	3.64	1.73					
S65-h2	2.97	3.54	2.63					
W12-a	−4.36	4.29	0.06					
W12-b	3.06	4.70	−1.23					
W12-c	0.82	4.79	−1.70					

Energy values for conformations included in the calculation of state energies (see Eq. 1 and Table 2) are in k_BT (=2.48 kJ/mol). Only interaction energies larger than 0.5 k_BT are given.

Y66-*, Y66-h1*, Y66-h2* refer to the electronically excited fluorophore. Interaction energies for Y66-h1* and Y66-h2* differ from the ground state values by maximally 0.5 k_BT.

with simultaneous anionic fluorophore and neutral E222. In experimental work two other sets of names have been introduced. Chattoraj et al. (1996) use the terms A and B for equilibrium states with neutral and deprotonated fluorophore, respectively; Lossau et al. (1996) refer to these states as RH_{eq} and R^-_{eq}. Our analysis reveals that, below pH 12, the protein state RH_{eq} has pH-dependent contributions from YH/E- and YH/EH, whereas Y-/EH contributes to R^-_{eq}. The fourth state, Y-/E-, is populated only for pH > 12. RH_{eq} and R^-_{eq} are linked by intraprotein proton transfer via the states YH/E- and Y-/EH. The calculated acid-base constants (standard deviation ≤0.01 *pK*-units) are summarized in Table 4, together with the relative populations of the states at pH 4 and 8. The pK_a-values in the basic pH-range (>12) will depend on further acid-base transitions, e.g., the titration of arginine R96. Because they are not regarded in the current analysis, these values are less reliable.

Figure 3 *B* shows the intimate relation between titration and conformer occupancy for the mobile polar side chains. Two principal situations can be distinguished. While sites like W22, T203, and S205 occupy multiple conformations, the hydroxyl group of S65 and water W27 are localized in steep energy minima (S65-h1, W27-a). The side chains with the largest number of accessible states will provide easier proton rearrangements. Interestingly, they are distributed near the chromophore protonation site. The calculation confirms that the reorientation of T203 from the major A-position at low pH to the B-position at higher pH is linked to the protonation state of the chromophore. Mutagenesis experiments indicate that the energy of chromophore ionization is also coupled to dipole reorientation or polar-hydrophobic exchange at residue 65. In wild-type GFP, the side chain of S65 can adopt a conformation in which the hydroxyl donates a hydrogen-bond to E222 (hydrogen position h1 in our analysis). In the S65T mutant, the hydroxyl is hydrogen-bonded to the main chain carbonyl oxygen of V61 (Ormö et al., 1996) located above the chromophore plane. This conformational change, but also the substitution of S65 by nonpolar residues (G, A, C, V [e.g., Heim et al., 1995]), favor the anionic form of the chromophore. On the other side, replacement of T203 by hydrophobic chains (T203V [B. Steipe, Genzentrum der Universität München, 1998, personal communication] or T203I [Ehrig et al., 1995]) stabilize the neutral chromophore. We simulated these changes (referred to as hypothetical mutants) of the specific local surrounding either by fixation or removal of the hydroxyls without further relaxation of the protein environment.

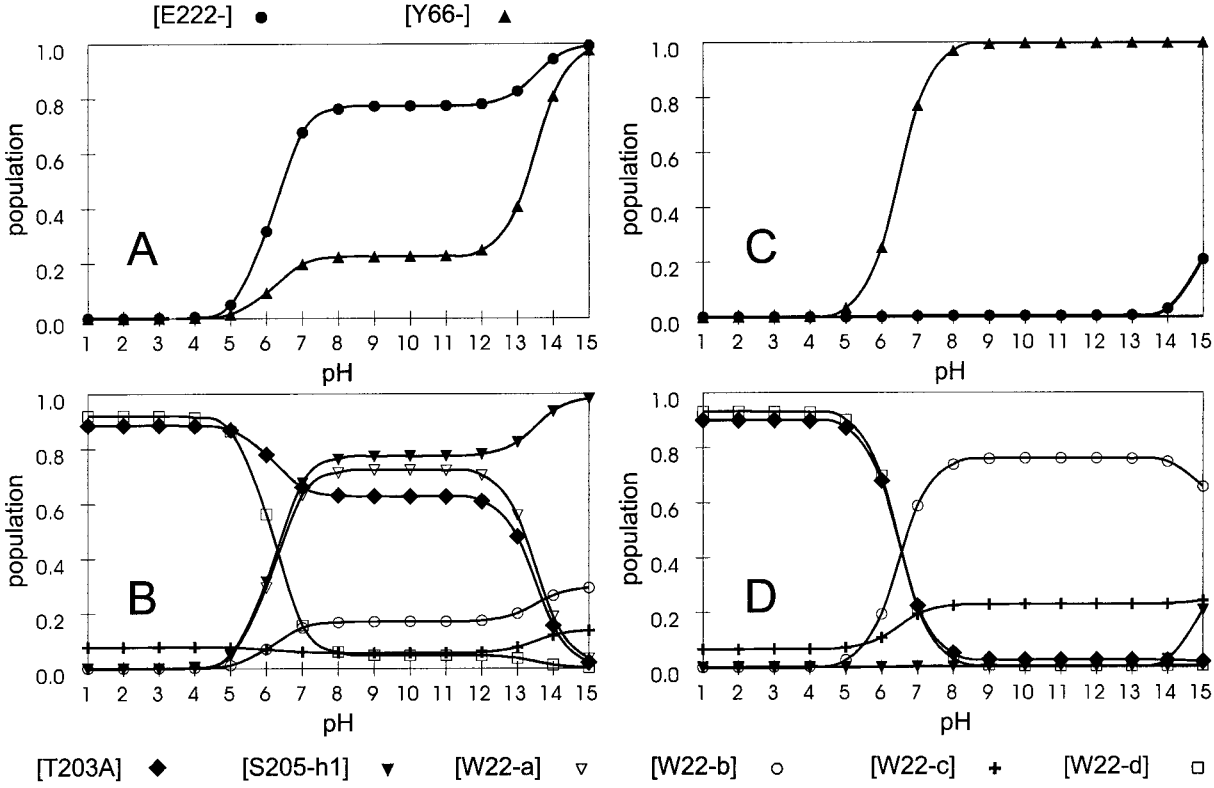

FIGURE 3 Site occupancies for wild-type GFP and the hypothetical mutant S65T. For the simulation of the mutant the hydroxyl orientation of S65 was fixed to position h2 (compare text and Fig. 1). (*A*, *C*) pH-Dependent population of ionized fluorophore Y66 and ionized E222 in wild-type (*A*) GFP and (*C*) S65T. (*B*, *D*) pH-Dependent fractional conformer occupancy for residues in the hydrogen-bonded net around the fluorophore in wild-type (*B*) GFP and (*D*) S65T.

The hypothetical mutant S65T was simulated by a restriction of the sampling to the h2-orientation of S65. Fig. 3 *C* confirms, that the removal of the hydrogen bond of residue 65 to E222 (the only change as compared to Fig. 3, *A* and *B*) shifts the acid-base equilibria for both, E222 and Y66, concomitant with a shift in conformer population (Fig. 3 *D*). Consistent with the x-ray structure for S65T (Ormö et al., 1996; Elsliger et al., 1999), the dominant orientation of T203 at pH 8 is the B position, correlated with the population of the ionized chromophore. Additional rearrangements include the flip of W12 from perpendicular (a) to parallel (b, c) orientation in respect to the imidazolinone ring and the freezing of the S205-h1 position. The inspection of the molecular model reveals that rearrangements occurring cooperatively with the reorientation of W12 but not sampled in the current treatment (e.g., twist of Q69) will contribute to the stabilization of the S65-h2 orientation. The hydrophobic side chains of the hypothetical mutants S65G, T203V and the double mutant T203V/S65G have been included in the Lennard–Jones interaction for the titratable sites and the removal of hydrogen bond possibilities (e.g., E222-h1/W12-h2 and Y66-h1/W22-h5). Additionally, we studied the connection of these protein changes with the presence of the buried water molecules, thereby modulating the polarity of the binding site.

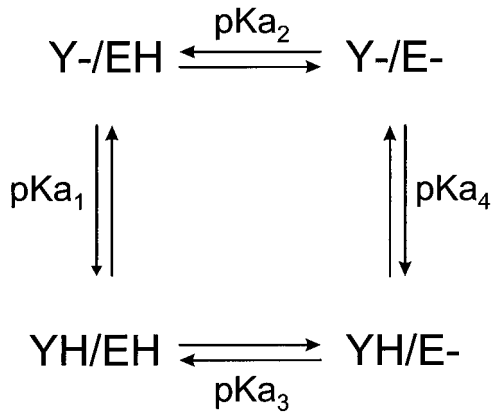

FIGURE 4 Scheme describing the equilibrium acid–base transitions of two interacting residues. YH and EH refer to equilibrium protein states with neutral Y66 and E222 residue, respectively. Y- and E- refer to protein states with negatively charged residues.

The results for the hypothetical mutants are summarized in Table 4. The analysis of the coupled acid-base equilibria for Y66 and E222 shows how the influences from neighboring sites modulate the coupling. If the surrounding shifts the pK_a so that they match each other (within a range of approximately 2 *pK* units), the states Y-/EH and YH/E- are populated in a nonvanishing constant ratio $f_{Y\text{-}/EH}/f_{YH/E\text{-}} =$

TABLE 4 Calculated equilibrium constants and relative occupancies describing the transitions of the interacting residues Y66 and E222

					[T203A]%*		$f_{Y-/EH}$%#		$f_{YH/EH}$%#		$f_{YH/E-}$%#	
	pK_{a1}	pK_{a2}	pK_{a3}	pK_{a4}	pH = 4	pH = 8	pH = 4	pH = 8	pH = 4	pH = 8	pH = 4	pH = 8
wild type												
A+B§	6.8	12.9	6.3	13.4	88.4	63.0	0.2	23.7	99.3	1.5	0.5	74.8
A§	(8.4)	(13.5)	6.3	15.6	100.0	100.0	0.0	0.8	99.5	1.9	0.5	97.3
B§	5.9	12.9	6.0	12.8	0.0	0.0	1.2	55.5	97.8	0.4	1.0	44.1
S65T												
A+B§	6.5	15.5	(8.9)	(13.1)	89.7	4.8	0.3	96.5	99.7	3.1	0.0	0.4
A§	7.0	16.2	7.9	15.3	100.0	100.0	0.1	81.5	99.9	8.2	0.0	10.3
B§	5.5	15.6	(8.5)	(12.6)	0.0	0.0	3.1	99.6	96.9	0.3	0.0	0.1
S65G												
A+B§	6.8	14.8	8.1	13.5	88.6	9.9	0.2	89.8	99.8	5.7	0.0	4.5
A§	8.4	15.9	8.1	16.2	100.0	100.0	0.0	18.2	100.0	45.6	0.0	36.2
B§	5.9	14.9	(7.9)	(12.9)	0.0	0.0	1.2	98.2	98.7	0.8	0.0	1.0
T203V												
w19a¶	(7.7)	(13.5)	5.8	15.4	—	—	0.0	1.2	98.4	0.6	1.6	98.2
w19b¶	7.4	13.8	6.3	14.9	—	—	0.0	7.2	99.5	1.8	0.5	91.0
T203V/S65G												
w3‖	7.7	15.4	7.7	15.4	—	—	0.0	40.0	100.0	20.0	0.0	40.0
w2‖	7.5	15.7	8.1	15.1	—	—	0.0	63.8	100.0	20.2	0.0	16.0
w0‖	7.7	17.1	9.4	15.4	—	—	0.0	65.7	100.0	33.0	0.0	1.3

Given are the values for wild-type GFP and hypothetical mutants (for further details see text and Fig. 4). pK_{a1} and pK_{a4} refer to acid-base transitions for Y66 in the presence of neutral and anionic E222, respectively. pK_{a2} and pK_{a3} denote protonation changes of E222 in the presence of anionic and neutral Y66, respectively. Values in brackets indicate that the corresponding equilibrium is not realized.

*Relative occupancy of side chain orientation A for residue T203.

#Relative populations of neutral and anionic forms are calculated from the pK_a (Eq. 3, a–c).

§Entries A and B refer to simulations with fixed orientation for the T203 side chain; in simulations A+B both conformations were sampled.

¶For simulation T203V, water W19 was either fixed in the orientation shown in Fig. 1 (w19a) or removed (w19b).

‖Simulation T203V/S65G was carried out with varying polarity in the binding site including either three (w3: W12, W19, W27), two (w2: W12, W27), or none (w0) of the buried water molecules located above the chromophore plane.

$10^{(pK_{a3}-pK_{a1})}$ (compare Eqs. 3a,c) in the pH region extending from min(pK_{a1}, pK_{a3}) to min(pK_{a2}, pK_{a4}). This is the case for wild-type GFP if T203 is free to stabilize chromophore ionization (T203B). The fixation of T203 in an orientation without favorable interaction to anionic Y66 (T203A) or hydrophobic substitution (T203V) shift both protonation equilibria of the chromophore (pK_{a1} and pK_{a4}) in favor of the neutral form. Therefore, the only residue titrating in the neutral pH range is E222 (pK_{a3}) and the protonation of Y66 is fixed for pH $< pK_{a4}$. On the other side, the upshift of the two equilibria for E222 (pK_{a2}, pK_{a3}) by S65 substitution (S56T or S65G) forces protonation of E222 for pH $< pK_{a2}$ and reduces measurable titration to Y66 (with pK_{a1}). Additionally, depending on which residue is ionized first with increasing pH, the other will have its pK_a shifted up by the charge on the other site, screened by the aligning surrounding dipoles. Hydrophobic substitution near Y66 as well as near E222 (T203V/S65G) shift the equilibria in a comparable way and restore the proton-sharing relationship in neutral and basic pH range. Decreasing amount of buried water above the chromophore influences mainly the pK_a of E222.

The calculated similarity of pK_{a1}-values for wild type and S65T substitution is in accordance with experiment (reported values for mutants carrying the S65T motif are 5.98 [Kneen et al., 1998] and 5.8 [Haupts et al., 1998]). The analysis shows that, even without shifts in pK_{a1}, the relative population of neutral and anionic fluorophore at physiologic pH experiences drastic changes. As a consequence of the internal proton exchange with E222, the ratio $f_{YH}/f_{Y-} = 10^{-x_1}(1 + 10^{-x_3})$ (compare Eqs. 4a,b) deviates from the expected dependency for a one-step titration ($f_{YH}/f_{Y-} = 10^{-x_1}$).

Analysis of conformational substates

The electrostatic calculations yield results at two levels: the macroscopic result is the average population of protonation and conformational states, the microscopic result is the individual state vector $\vec{x}$ describing protonation and conformation of the residues included. Each state vector is related to a particular free energy $\Delta G(\vec{x})$ (see Methods, Eq. 1). The description of the available states becomes relevant in connection to single molecule spectroscopy. To study protein heterogeneities in the presence of protons on several acceptor sites, we analyze the dipole orientations making up the lowest configurations. This analysis also provides information about conformational dependencies and coupled movements. The permissible orientations sampled for the three ionization states YH/EH, YH/E- and Y-/EH are summarized in Table 5 (wild-type GFP simulation, free energy values

TABLE 5 Lowest energy conformations for wild-type GFP

Conformation	Y66	T203	W22	S205	E222	S65	W12	GS*	ES*
A1	—	bh1	b	h2	h4	h1	a	1.2	−21.3
A2	—	bh1	c	h2	h4	h1	a	1.6	−23.2
A3	—	bh1	b	h2	h3	h1	a	4.3	−18.2
A4	—	ah3#	b	h2	h4	h1	a	4.5	−21.4
A5	—	bh1	c	h2	h3	h1	a	4.7	−20.0
A6	—	bh1	d	h2	h4	h1	a	5.3	−21.6
A7	—	bh3	b	h2	h4	h1	a	5.3	−18.9
N1	h2	ah3#	a	h1	—	h1	a	0.0	−0.5
N2	h2	bh1	a	h1	—	h1	a	1.9	1.4
N3	h2	ah3#	d	h1	—	h1	a	2.1	1.6
N4	h2	bh3	a	h1	—	h1	a	3.6	3.0
N5	h2	bh1	d	h1	—	h1	a	3.8	3.2
N6	h2	bh2	a	h1	—	h1	a	4.1	3.5
N7	h2	bh2	c′	h1	—	h1	a	5.9	5.3
N8	h2	bh1	c′	h1	—	h1	a	7.0	6.5
N9	h1	bh2	e	h1	—	h1	a	4.8	4.2
N10	h1	bh2	a	h1	—	h1	a	5.8	5.2
NN1	h2	ah3#	d	h2	h4	h1	a	3.3	2.6
NN2	h2	ah3#	d	h2	h3	h1	a	5.4	4.7
NN3	h1	bh2	c	h2	h4	h1	a	5.6	4.9

Given are the orientations associated with the ionization states of Y66 and E222 for the residues in the hydrogen-bond net (for details see Figs. 1 and 2). The shaded area represents photoactive states for the RH_{eq} manifold (see Table 6 for details concerning proton transfer).

*GS, state energies for the electronic ground state; ES, excited state. Energies are for pH = 8 relative to state N1 in units of k_BT (=2.48 kJ/mol).

#T203 samples orientations ah3 and ah2 with approximately the same probability, corresponding to state energy differences $<0.3\ k_BT$.

are for pH = 8 relative to the lowest energy state). States in an energy range of $\approx 5\ k_BT$ have been included, which is approximately twice the numerical error inherent in the determination of state energies (Gilson et al., 1987). In the nomenclature introduced by Lossau et al. (1996), the manifold of states {A1, . . . , A7} characterizes the equilibrium protein state R_{eq}^{-}, the states {N1, . . . , NN3} characterize the protein state RH_{eq}.

The sampled states for anionic Y66 (states A1, . . . , A7) can be grouped according to the fluctuating hydrogen-bond pattern between the phenolate oxygen, T203 and water W22. With T203-bh1 and W22-h1 (in b-orientation), state A1 provides two short hydrogen bonds to Y66. Only one hydrogen bond develops for states A2 (from T203-bh1) and A7 (from W22-h1)—as compared to A1, the second partner is involved in T203B–W22 interaction. In all conformations, the oxygen of W22 is tied up by the hydroxyl group of S205. The energy difference between *syn*- and *anti*-configuration of the hydroxyl group of E222 is $\approx 3\ k_BT$, indicating the modulation of the intrinsic preference (4.6 k_BT [Alexov and Gunner, 1997]) by electrostatic interaction with S205.

The interaction pattern of the neutral fluorophore exhibits large heterogeneity: additional to the two orientations of its phenolic hydroxyl (Y66-h1, -h2), the environment differs in both, protonation of E222 and orientation of dipolar groups. In the presence of negative E222, S205 favors the hydrogen bond to OE2, at the same time allowing the reorientation of water W22, which accepts a strong hydrogen bond from Y66-h2 and flips between S205 and the main chain oxygens of residues 146 and 203 (orientations a, d). Consistent with the x-ray structure (Brejc et al., 1997) T203 adopts A-orientation in the lowest energy configuration (state N1). Although the presence of a proton on Y66 disfavors T203-bh1/bh3 orientations (e.g., N2), the dipole reorientation to T203-bh2 becomes energetically comparable to A-orientations (N6), allows a 90°-flip of water W22 (c′) and facilitates hydrogen bonding for Y66-h1 (states N9, N10). After reorientation of S205, W22 and the T203 hydroxyl to almost exclusive A-orientation, the protein accommodates both, protonated Y66 and protonated E222 (states NN1, . . . , NN3). In accordance with crystallographic data for buried water molecules (Denisov et al., 1997), W22 is engaged in three to four hydrogen bonds with protein atoms. At least two short hydrogen bonds characterize the lowest energy states. The population of T203A-orientation together with the anionic chromophore (state A4) and B-orientation together with neutral Y66 (e.g., N2, N4, N5, N6, NN3) shows that this residue can be disordered even for fixed protonation of the fluorophore.

As an immediate consequence, the different dipole orientations will lead to a distribution of electrostatic potential around the chromophore. Figure 5 shows the individual contributions to the electrostatic potential at the atomic positions of the fluorophore. Among the polar side chains with fixed charge and orientation, the positive arginine R96 near the imidazolinone ring dominates and initiates a potential gradient along the chromophore. All dipole orientations (states A1, . . . , A7 in Table 5) near the phenolic end of ionized Y66 compensate this potential difference approximately. Fluctuation among these states introduces no drastic difference, because the movement of one dipole away

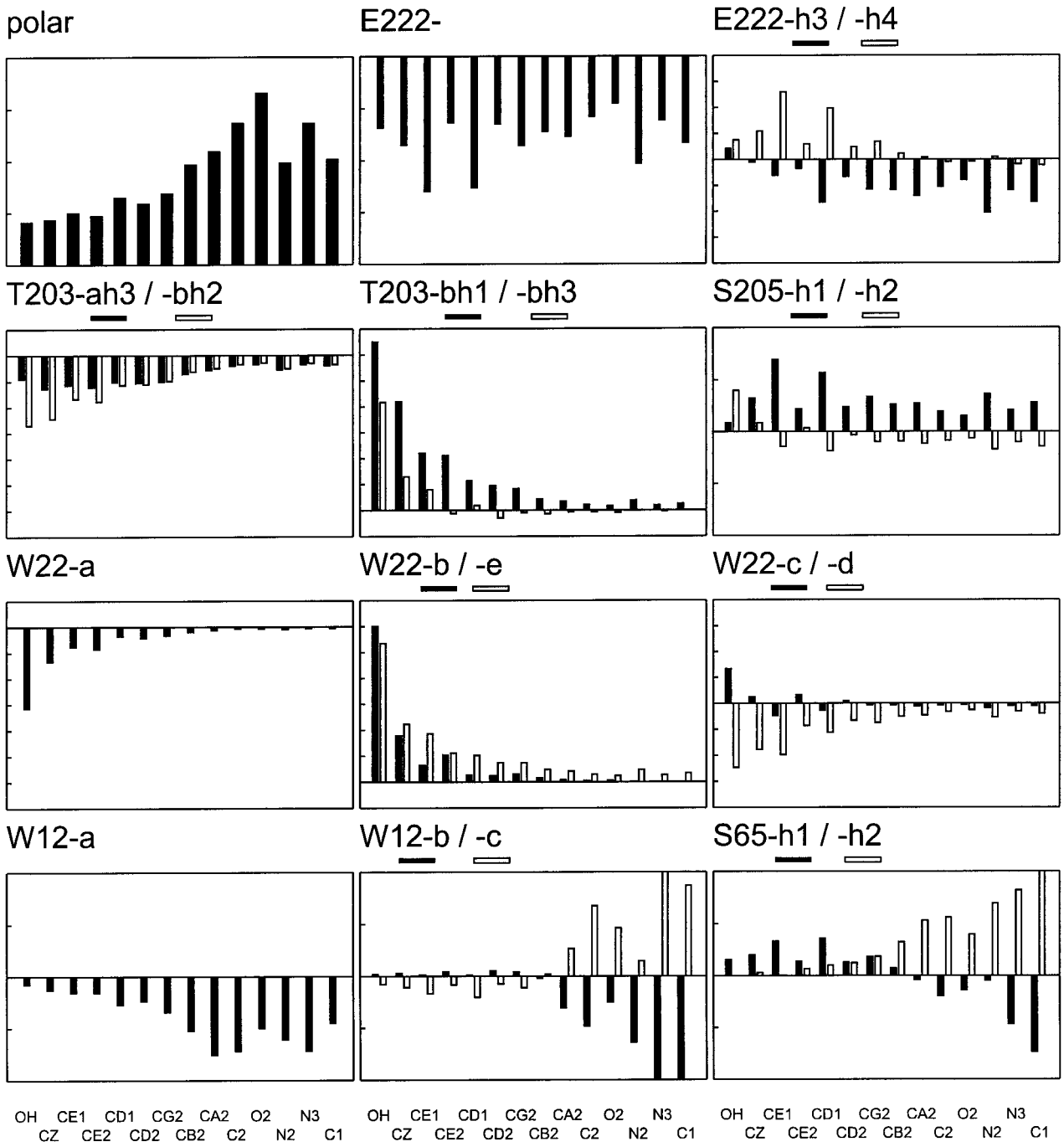

FIGURE 5 Electrostatic potentials from surrounding residues at the atomic positions of the fluorophore in wild-type GFP. Potentials were calculated from protein partial charges (compare Methods) using continuum electrostatic methods. Bars pointing up (down) indicate positive (negative) potential. One section on the *y*-axis corresponds to 10 k_BT/e for the contribution from polar side chains and negatively charged E222, and to 2 k_BT/e for the other conformations (1 $k_BT/e \equiv 25.6$ mV).

from the phenolate is compensated by the reorientation of the others. For the neutral fluorophore states in the presence of negative E222 (N1, . . . , N6) the potential gradient intensifies. The sum values are sensitive to dipole conformations of T203, but not to fluctuations of water W22 between a- and d-orientation. T203-ah3/bh2 dipoles are connected with stabilizing, and the other T203-orientations with destabilizing potential for the phenolic proton.

Response to instantaneous dipole change induced by light absorption

In accordance with other theoretical investigations (Vojtyuk et al., 1998) our quantum chemical calculations (see Methods) reveal that excitation of the chromophore is connected with transfer of electronic charge density from the phenolic part to the imidazolinone ring. The effect is more pro-

nounced for the anionic chromophore, whereas the charge displacement for the neutral fluorophore is more localized in the imidazolinone ring (compare Table 1). This difference is also mirrored by the calculated change in dipole moment of $\Delta\mu = 13$ D as compared to $\Delta\mu = 1$ D for the neutral chromophore. Within the error limits of the semiempirical approach, the calculated values fit to $\Delta\mu = 6.8 \pm 0.3$ D (Chattoraj et al., 1996) for the anionic fluorophore and $\Delta\mu = 2.5$ D (Bublitz et al., 1998) determined for the chemically modified neutral Y66H-fluorophore from Stark effect spectroscopy. Our calculations are also able to reproduce the experimental observation of a nearly parallel orientation of $\Delta\mu$ and the transition moment for the anion (calculated angle: 3°) in contrast to the almost perpendicular orientation (calculated angle: 87°) for the neutral fluorophore.

Generally, the various distributions of dipoles around the chromophore will respond in a distinct way to the changed charge distribution, developing toward the new equilibrium in a variety of time scales. For our analysis, we neglect further rearrangements in the chromophore and in the protein surrounding. The last column in Table 5 shows the response of the state energies $\Delta G(\vec{x})$ (Eq. 1) to the changed charge distribution. The reduction of electronic density at the phenolate oxygen weakens the hydrogen-bond strength from ionized Y66 to T203 and W22 (compare also the site–site interactions in Table 3), thereby facilitating rearrangements in the immediate environment reflected by the close energies of states A2, A4, A6, A1. These conformations are representative for the protein state R_{eq}^{-}*, if the hydrogen bonds are able to rearrange themselves to the lower energy conformations in the lifetime of the excited fluorophore. In contrast, the smaller and more localized charge displacement associated with excitation of the neutral chromophore conserves the energetic order of the ground-state dipole distribution, e.g., excited-state equilibrium protein states RH_{eq}^{*} correspond to the ground state manifold RH_{eq}.

Energetics of proton transfer

The photophysical pattern of GFP has been rationalized by ESPT from the fluorophore hydroxyl group to the hydrogen-bonded network, leading to the deprotonated chromophore from which the actual emission occurs. Kinetic results reveal that the excited-state deprotonation happens within the picosecond time scale, is slightly activated, and involves nonequilibrium protein states (Chattoraj et al., 1996; Lossau et al., 1996). A variety of quantitative evaluations of proton transfer dynamics in proteins extending from semiclassic trajectory studies (Warshel, 1982; Warshel, 1991) to quantum dynamic simulations (Laria et al., 1992; Bala et al., 1996; Berendsen and Mavri, 1996) reveal that adiabatic proton transfer is driven by local polarity changes at the donor and acceptor sites. This provides a general rule for the discussion of the kinetics in terms of the underlying molecular structure and energetics (Warshel, 1986). For hydrogen-bonded donor-acceptor pairs (distance <3.5 Å), the activation barrier $\Delta G^{\#}$ is correlated with the free energy difference ΔG_{PT} between initial and final state,

$$\Delta G^{\#} \approx \Delta G_{PT} \quad \text{for} \quad \Delta G_{PT} > 0$$

and

$$\Delta G^{\#} \approx 0 \quad \text{for} \quad \Delta G_{PT} < 0.$$

Depending on the reaction time, as compared to the relaxation of the environment, two limiting cases have to be distinguished: the transport can occur in a constant, nonequilibrated or in a relaxed environment. A prerequisite for fast transport along an extended chain is the rigidity of the conducting structure in the time window of the transport.

We model proton translocation in GFP as a series of thermally activated hopping processes with localized proton jumps between neighboring sites along the preformatted hydrogen-bond network extending from Y66 to E222. The large isotope effect for ESPT favors the 3-step model with E222 as final acceptor over, e.g., translocation to H148 (Palm et al., 1997). Varying electrostatic fields enter in terms of fluctuating hydroxyl dipoles and lead to widespread driving forces and barriers. A particular state of the system is identified by the location of the proton and the state of the surrounding hydrogen-bond net (compare Table 6). The first step ($1 > 2$) involves transport of the phenolic hydrogen h1 to water W22 forming hydronium ions W22-a+, W22-d+, and W22-c′+. In the next step ($2 > 3$) S205 accepts proton h2 from W22, forming protonated S205+. The final step ($3 > 4$) produces neutral E222-h3 (*anti*-orientation) after transfers of proton h1 from S205+ to E222. The short hydrogen bonds between S205 to W22 and E222 (compare Table 2) have been treated as gating bonds (Nagel et al., 1980) favoring ionic transport, simultaneously slowing down the formation of a turning defect, which requires at least 10 $k_B T$ for the breaking of the hydrogen bond.

If the reaction time is shorter than the relaxation time of the surrounding, nonequilibrium effects arise. In this case, the relevant potential energy surface for proton transfer is the one obtained for the current configuration of the environment. States N1, . . . , N8 (a subset of RH_{eq}, characterized in Table 5) represent possible photoactive arrangements of the hydrogen-bond net (states N9 and N10 require additional reorientation of the Y66-OH group and are not discussed). Table 6 summarizes the calculated free energy gaps along the transport chain for ground and excited states.

Due to the increased acidity of the hydroxyaromatic compound (compare Methods), electronically excited Y66 is a strong proton injector. On the other end of the transport chain, E222 is a strong proton acceptor, characterized by the large negative free energy difference for the last step. The intermediate steps depend on both the orientation of the aligning dipoles and the orientation of the groups participating in the translocation. The negative potential at W22,

TABLE 6 Free energy gaps ΔG_{PT} for forward transport of a proton along the preformatted hydrogen-bond net extending from Y66 to E222

	1 Y66-h3/E222-				2 Y66-/W22+/E222-				3 Y66-/S205+/E222-				4 Y66-/E222-h3	
			ΔG_{PT} (1 > 2)				ΔG_{PT} (2 > 3)				ΔG_{PT} (3 > 4)		ΔG	
State *j** Conformation	T203	W22	GS	ES	T203	W22	GS	ES	T203	W22	GS	ES	GS	ES
N1	ah3	a	21.8	2.4	ah3	a+	4.0	−0.9	ah3	b	−18.1	−20.0	7.7	−19.0
N2	bh1	a	19.4	4.4	bh1	a+	1.1	−3.9	bh1	b	−18.1	−20.0	4.3	−18.2
N3	ah3	d	19.6	0.5	ah3	d+								
N4	bh3	a	23.7	7.2	bh3	a+	−0.4	−5.4	bh3	b	−18.4	−20.3	8.5	−15.5
N5	bh1	d	17.0	2.4	bh1	d+								
N6	bh2	a	21.2	1.0	bh2	a+	5.2	0.2	bh2	b	−17.9	−19.7	12.6	−15.0
N7	bh2	c′	13.4	−4.9	bh2	c′+	11.3	4.4[#]						
N8	bh1	c′	11.4	−1.7	bh1	c′+	4.0[#]	−2.9						

Calculations are carried out for the lowest energy photoactive states of the RH_{eq} manifold (states N1 to N8 in Table 5). The driving force for transfer in the opposite direction is connected to forward transfer by $\Delta G_{PT}(j > i) = -\Delta G_{PT}(i > j)$. Energies are given in k_BT (=2.48 kJ/mol).
*A state j in the transport chain is identified by the electronic state of the fluorophore (GS, ground state; ES, excited state), the location of the proton (underlined) and the orientation of the surrounding dipoles. ΔG is the energy of the final state ($j = 4$) in unrelaxed dipolar surrounding.
[#]The relaxed state in the presence of the proton on water W22 is marked for the corresponding electronic state.

provided by dipole orientations T203-ah3/bh2 (states N3, N6), stabilizes the hydronium ion; thereby enhances the first step and disfavors simultaneously the second step. Consistently, the rotation of the dipole orientation (T203-bh1/bh3, states N2 and N4) reverses the situation. Water W22 flips between orientations W22-a which facilitates complete hydrogen-bonding between Y66 and E222, and orientation W22-d, which requires rotation before the second step. Experimental rotational relaxation times for water molecules range from 3 ps (the lifespan reported for a hydrogen bond in a percolating water network [Gutman and Nachliel, 1990]) to the subnanoseconds regime for the torsional vibration around an orthogonal axis (Denisov et al., 1997). A waiting period is also related to the probability of reaching W22 orientation c′ (states N7 and N8). This analysis shows how the translocation rate depends on details of the proton solvation on the W22 site. The lowest energy configurations, N1 and N2, allow activated transport in an ordered chain without reorientation. However, the presence of the ion on W22 will affect the positions of the protons on neighboring sites, especially T203 participates in the solvation process. Dipole orientations characterized by state N7 provide the greatest stabilization for hydronium W22+. If the reaction time is slower than solvation, the hydrogen-bond net will become polarized in a way against transport. Assuming that solvation processes in the excited state occur on a longer time scale, ESPT from state N1 will end up in a nonequilibrium state $N1^-$*, an example for protein states R^-_{neq}*, differing from the lowest energy dipole arrangement for excited anionic fluorophore (state A2* from manifold R^-_{eq}*, compare Table 5) with respect to the orientation of T203 (ah3 versus bh1), of water W22 (b versus c), and the isomerization state of the hydroxyl on E222 (h3 versus h4). Energy contributions to these relaxation processes (intrinsic barriers refer to the force field values and do not include specific interactions in the protein, compare Methods) involve: break of the hydrogen bond from E222-h3 to S205 (5 k_BT), *anti-/syn* isomerization to E222-h4 (intrinsic barrier: 2.5 k_BT), reorientation of T203 from A to B (intrinsic barrier: 5.4 k_BT), rotation of the T203 hydroxyl (intrinsic barrier: 2.2 k_BT). During the short lifetime of the excited anionic state (3.3 ns [Chattoraj et al., 1996; Lossau et al., 1996]) the full equilibrium conformation will not be achieved, especially reorientation of T203 can occur on a longer time scale. The calculations reveal further a small activation energy in the order of $\approx k_BT$, consistent with the slowing down of ESPT by a factor of at least 5 between room temperature and 80 K (Chattoraj et al., 1996; Lossau et al., 1996). The estimate for the free energy gap between initial RH^*_{eq} and final (nonequilibrated) state R^-_{neq}* for excited state deprotonation is $-19\ k_BT$ (compare Tables 5 and 6). The large isotope effect that increases with the driving force is indicative of an energy gap of that scale (Lossau et al., 1996).

At any state in the ESPT process, deexcitation to the ground state may occur. The system is locked in disequilibrium in relation to proton distribution and dipole orientation. Depending on the state and driven by the competition of their acidities, the proton will be pumped back to Y66 or forward to E222 on different time scale. Because the lifetime of the state formed via ESPT (e.g., $N1^-$*) is presumably too short to allow full relaxation toward the equilibrium state (A2* in Table 5), deexcitation in a protein configuration resembling still N-character (especially in respect to orientation of T203) provides a surrounding that favors fast repopulation of the starting configuration (e.g., N1).

In contrast to the excited state, the ground-state free energy profile for the $RH_{eq} \Leftrightarrow R^-_{eq}$ reaction (Table 5) has only a small gap ($\approx 2\ k_BT$) between initial and final state (as mirrored also by the ground-state equilibrium population, compare Table 4). Ionic entry from both sides of the translocation network requires approximately 20 k_BT, consistent with the experimentally determined (Haupts et al., 1998)

340 μs time for the internal ground-state proton exchange and accompanying hydrogen bond rearrangements. In a manner similar to the excited state, the actual injection energy from the phenolic fluorophore is altered by dipolar fluctuations of T203 and water W22. The alternative transport sequence involving transport of the negative charge of E222 in the opposite direction to proton transfer is energetically less favorable because the positive charge on W22 is not sufficient to stabilize both, the negative charge on Y66 and the negative charge on S205.

DISCUSSION

X-ray crystallography, ultrafast optical spectroscopy, and site-specific mutagenesis data suggest that the protonation state of the tyrosylhydroxyl group of the chromophore is responsible for the pH sensitivity of GFP. Taking into consideration the experimentally shown interdependence of protonation and rearrangements of the protein framework (e.g., Ormö et al., 1996; Brejc et al., 1997; Palm et al., 1997; Dickson et al., 1997; Jung et al., 1998; Haupts et al., 1998; Elsliger et al., 1999), theoretical methods that provide a sound basis for the calculation of pH-dependent properties of the GFP protein have to include, therefore, coupling between the ionization states and protein flexibility (You and Bashford, 1995; Scharnagl and Fischer, 1996; Beroza and Case, 1996; Ripoll et al., 1996; Alexov and Gunner, 1997). For our analysis, we use the atomic model provided by the low-pH x-ray structure for monomeric wild-type GFP (Brejc et al., 1997). Guided by the results of crystallographic analysis (Ormö et al., 1996; Brejc et al., 1997; Palm et al., 1997; Elsliger et al., 1999), circular dichroism (Kneen et al., 1998) and FTIR spectroscopy (Van Thor et al., 1998) conformational sampling was restricted to reorientations of hydroxyl dipoles and buried water in the immediate surrounding of the fluorophore.

The calculations correctly reproduce experimentally established behavior of key residues: two-step titration of the fluorophore and individual acid-base equilibrium constants; pK_a-shifts due to chemical modification of aligning residues T203 and S65; coupling of the orientation of T203 with the protonation state of the fluorophore; response of the hydrogen-bond net to ionization changes; and increased acidity of the fluorophore in the excited state. In addition, the investigation extends existing knowledge. First, a molecular model for the two-step titration is developed. Second, the dependence of the fluorophore protonation on protein and buried water structural elements is quantified. Third, molecular models for the dipolar microheterogeneity around the fluorophore are described and characterized in relation to spectroscopic properties and relaxation pathways and time scales after light absorption and ESPT. Fourth, the heterogeneous driving forces for proton transfer (in ground and excited state) are discussed in terms of the underlying molecular structure and electrostatic properties at the donor and acceptor sites. The use of only one structure and a restricted number of conformational degrees of freedom impose, however, some limitations.

Molecular basis for the two-step titration

For pH $>$ 4, the response of GFP fluorescence to pH changes occurs in $<$1 ms and is reversible (Ward, 1981; Kneen et al., 1998). The pK_a for the phenol–phenolate transition of the fluorophore in wild-type protein was reported as $\approx$5, . . . , 6 (Bokman and Ward, 1981; Kneen et al.; 1998), but the titration curve of the protein has a second pK_a near 12 (Bokman and Ward, 1981). Between pH 8 and 11, neutral and charged forms are present in a constant ratio. The calculated pH-dependent occupancies of the ionized residues (Fig. 4) reveal that this titration behavior is the consequence of the competition between the acidities of the phenolic group of Y66 and the carboxylate of E222. The coupling of their chemical equilibria has been quantified by four pK_a values (compare Fig. 4) describing acid–base transitions of one residue dependent on the charge state of the other. Consistent with results from fluorescence correlation spectroscopy (Haupts et al., 1998), we find the pH-dependent population of three protonation states for pH $<$ 12, the fourth state Y-/E- develops for pH $>$ 12. At low pH, the state YH/EH dominates. The two states YH/E- and Y-/EH, coupled by internal proton transfer and reorientation of the internal hydrogen-bond net, are present in a constant ratio in the neutral and basic pH region. Their relative population is driven by the difference of pK_{a1}, the protonation equilibrium of Y66 in the presence of neutral E222, and pK_{a3}, the protonation equilibrium of E222 in the presence of neutral fluorophore. This finding implies that the fractional occupancies f_{Y-}, as determined, e.g., by fluorescence monitoring at the emission wavelength of the anionic fluorophore, depend not on a single protonation equilibrium (Eq. 4a). Coupled ionization equilibria of residues are also the key factors for extracellular proton release in bacteriorhodopsin (Balashov et al., 1995, 1996). In this proton-translocating membrane protein, light-induced protonation of the primary acceptor, an aspartate residue (D85) near the retinal chromophore binding site is communicated to the distant extracellular proton release group (involving E204), presumably mediated by the reorientation of the positively charged arginine (R82) located in between (Scharnagl and Fischer, 1996) and the rearrangement of internal water molecules.

Correlation of the neutral-to-anionic ratio with structural elements

Reorientation of the T203 hydroxyl dipole or substitution with a hydrophobic side chain shifts pK_{a1}, similar modifications of S65 modulate pK_{a3}. Depending on which residue is ionized first with increasing pH, the other will have its pK_a shifted up by the charge on the other site, but screened by the dipoles aligning in the new electrostatic field. Our

results pointed out that the internal proton-sharing relationship will lead to drastic changes in the relative population of neutral and anionic fluorophore in the physiologic pH-range, despite only small changes in pK_{a1}. The approach of the two pK_a within 0.9 units for the small fraction of T203A orientations in the simulation of the hypothetical S65T mutant (compare Table 4) leads to an equilibrium constant $K' = f_{Y\text{-}/EH}/f_{YH/E\text{-}} = 10^{0.9} \approx 8$ corresponding to the experimentally reported (Haupts et al., 1998) ratio (7.55 ± 0.60) for a mutant carrying the S65T substitution. In addition, the calculated barrier for the ground-state protonation equilibrium between these two states (compare Table 6) in the order of ≈20 k_BT gives the right scale for the associated relaxation time (340 μs, Haupts et al., 1998).

Mutual influence of chemical and conformational substates

This interdependence as revealed by the calculations becomes evident in experiments monitoring the fluorescence of anionic chromophores for a small ensemble of molecules (Dickson et al., 1997; Jung et al., 1998; Haupts et al., 1998). The dependence on the time scale of the individual conformational transitions can be illustrated by the following examples. It is realistic to assume that equilibration between the two orientations of T203 is faster than proton exchange with external solvent (≈300 μs at neutral pH for an S65T mutant [Haupts et al., 1998]), whereas the reorientation of the S65 hydroxyl is—due to its cooperative nature involving additional water and side-chain orientations—a slower conformational change. This protein rearrangement shifts the ionization equilibrium for E222 (=pK_{a3}) from 6.3 to 8.9, thus increasing the population of the Y-/EH state by a factor of 4 on expense of the state YH/E- (values from Table 4 for pH = 8). Our analysis shows that T203 relaxation involves formation and breaking of hydrogen bonds to neighboring main chain oxygen atoms. The A- to B-reorientation will, therefore, be coupled to the larger scale cooperative movements of the β-sheet backbone (Haupts et al., 1998) on a slower time scale. Starting with configuration {Y-/EH; T203B; S65-h2}, a possible sequence of conformational changes can proceed via states {Y-/EH; T203B; S65-h1} and {Y-/EH; T203A; S65-h1}, while the concentration of the anionic fluorophore changes from 97 to 56% and, finally, to 1%. The time-resolved disappearance of fluorescence will scan the time scales for the individual transitions. Because the internal proton transfer is coupled to the presence of a buried water near the fluorophore, translational movement or exchange of W22 will be additional determinants for the on/off behavior of GFP.

Nature and relationship of dipole fluctuations around the fluorophore

The most detailed picture of the protein microheterogeneity within the limitations of the predefined conformational changes is provided by the analysis of the properties of the accessible substates sampled by each single molecule even in an equilibrium state. The results are linked to experimental properties like absorption wavelengths, reorganization of the hydrogen-bond net after charge movement upon excitation, and pathways for equilibration after proton transfer. Compared with the anionic fluorophore, the neutral one experiences a greater deal of electrostatic potential fluctuation associated with different orientation of its tyrosine hydroxyl as well as the charge and dipole distribution in the surrounding (compare Table 5). At the same time, due to the greater amount of internal charge transfer upon excitation of the ionized chromophore, protein response is more pronounced for this charge state, resulting in the stabilization of dipoles with flipped orientations in the region of the phenolate oxygen. The experimental fluorescence decay kinetics (Chattoraj et al., 1996; Lossau et al., 1996) revealed monoexponential decay for the anionic chromophore with a decay time of ≈3 ns, indicating equilibration in the lifetime of the excited state. Our results suggest that possible solvation processes involve mainly reorientation of a water molecule (W22). The experimental time scale for water rotations around a perpendicular axis in the subnanosecond regime (Denisov et al., 1997) supports the idea that the hydrogen-bond net is able to rearrange itself to relax to the lower energy conformations upon excitation of the anionic chromophore. An explanation for the nonexponential decay kinetics (even without ESPT) of the neutral fluorophore is provided by our findings that the ground state dipole distribution is conserved upon excitation. Under the assumption that the lifetime of the neutral fluorophore without ESPT is 0.9 ns (a value obtained for the nonproton-transferring Y66H mutant, Lossau et al., 1996), the nonexponential decay kinetics mirrors the ground-state heterogeneity rather than dielectric relaxation. Along with this argumentation, one has to keep in mind that the discussed structural changes include only a small set of relaxations and conserved chromophore geometry.

Energetics of proton transfer

The heterogeneous environment of the neutral chromophore will also influence the driving forces for proton transfer. Therefore, dispersive kinetics for ESPT has to be expected—in accordance with experiments (Lossau et al., 1996). The excited-state proton transfer reaction rapidly (shortest deprotonation time 3 ps [Chattoraj et al., 1996; Lossau et al., 1996]) converts the excited neutral fluorophore, forming an intermediate protein state (denoted by I* [Chattoraj et al., 1996] or R^-_{neq}* [Lossau et al., 1996]) with deprotonated excited chromophore in unrelaxed surrounding, giving rise to the green fluorescence. Ground-state conversion of the intermediate state is rapid (<1 ns [Lossau et al., 1996]) to the starting equilibrium configuration RH_{eq} or slow to the fully relaxed deprotonated state R^-_{eq}.

Although transport of protons involves breaking and formation of covalent bonds, modeling in a classic way is

possible under the assumption that the proton charge density can instantaneously adapt to the classical particle positions. As a consequence, proton transfer will be driven by the dynamics of the dielectric surrounding (Warshel, 1991; Laria et al., 1992; Bala et al., 1996). This provides an estimate for the driving forces and barriers derived from the electrostatic potentials at the relevant sites. The total translocation can be described as a hopping of the proton across a series of activation barriers along a hydrogen-bonded chain (Nagel et al., 1980; Marrink et al., 1996; Pomès and Roux, 1998) assuming that the conductive structure is stable enough not to loose rigidity in the time window of the transport. The changed proton affinity of the hydroxyphenolic compound in the excited state enters via a simple thermodynamic approach (Förster cycle) and altered charged distribution. We analyzed the driving forces for proton transfer along the preformatted hydrogen-bonding net extending from the phenolic part of the chromophore to E222. This three-step model is consistent with conclusions about the number of involved intermediates as judged from deuterium effects on the transport (Palm et al., 1997). The identity of the internal proton acceptor in ground and excited state is corroborated by fluorescence correlation spectroscopy results (Haupts et al., 1998) indicating a photoassisted shift of the ground state population. Our analysis is, however, limited to experimental conditions where the structural motif as revealed by the used x-ray structure (Brejc et al., 1997) prevails (see below).

If the reaction time is shorter than the relaxation time of the environment, nonequilibrium effects arise, and the relevant potential energy surface for ESPT will be the current configuration of the environment. The analysis, therefore, has been carried out for the photoactive subset of ground state populations RH_{eq}. Consistent with experiment (Chattoraj et al., 1996; Lossau et al., 1996), we found that ESPT is not a totally activationless process. The barrier in the order of $\approx k_BT$ is connected to the injection of the proton from the excited phenolic chromophore into the hydrogen-bonded net and unfavorable stabilization at the primary acceptor site, water W22. Subsequent transport steps are enhanced due to the restricted motion of the hydrogen-coupled pair S205–E222. Whereas the lowest energy dipole configurations are capable of forming complete hydrogen bonding among the groups participating in the transfer, higher energy states include additional waiting periods in the order of water reorientation times. If solvation processes are faster than the reaction, the response of dipoles to the presence of the charge on water W22 polarizes the surrounding in a way against transport. It is beyond the aim of the present investigation to monitor the dynamics of these relaxation processes. The overall ESPT is characterized by a large free energy gap ($\approx -19\ k_BT$) as compared to $\approx 2\ k_BT$ for back transfer in the equilibrated ground-state environment. From the absorption and emission data follows that the total free energy for deprotonation in the excited and reprotonation in the ground state cannot exceed 5000 cm^{-1} or 24 k_BT. This correspondence indicates that the calculations allow a correct description of the energetics of initial and final states in the proton transfer process.

Molecular nature of the relaxation processes leading to the equilibration of the deprotonated intermediate R^-_{neq}*

Immediately after ESPT to E222, the surrounding of the deprotonation site still resembles dipole orientations corresponding to the starting state RH_{eq}. The lowest energy R^-_{neq}* configuration will, therefore, be {T203-ah3, W22-b, E222-h3} in contrast to the environment of the relaxed dipole distribution R^-_{eq}* {T203-bh1, W22-c, E222-h4} around the excited anionic fluorophore. The participation of T203 in the nonequilibrium solvation reveals a new interaction pattern as compared to the previously (Brejc et al., 1997) discussed *syn*-/*anti*-isomerization of E222. The different nature of the surrounding becomes evident from low-temperature absorption and emission spectra (Lossau et al., 1996). At 150 K, emission from R^-_{neq}* occurs at 505 nm, emission from R^-_{eq}* at 486 nm. The room temperature absorption spectrum of R^-_{eq} peaks at 477 nm; for 150 K, absorption at 446 nm has been reported. To quantify the influence of the surrounding on the absorption spectra, we used semiempirical quantum chemical methods and calculated the absorption spectrum for the anionic chromophore in the presence of the relevant dipole orientations. The analysis revealed that the dipole orientation corresponding to R^-_{neq}* (T203 hydroxyl pointing away from the phenolate oxygen) shifts the longest wavelength absorption only a few nanometers as compared to the result for the isolated chromophore (472 nm). The strong hydrogen bond between the T203-bh1 hydroxyl and the phenolate oxygen leads to a complex mixture of the local $\pi \rightarrow \pi^*$ excitation with charge transfer transitions involving T203 as well as arginine R96, which is in hydrogen-bonding distance to the imidazolinone oxygen. This mixing splits the absorption in two bands around 470 nm and 440 nm. The intensity ratio is modulated by the dipole orientation of W22: W22-b and W22-c orientations (the lowest energy configurations A1 and A2 for protein state R^-_{eq}, Table 5) show an intensified 440 nm band, for W22-d (state A6) the two absorptions have comparable oscillator strengths. The blue-shifted absorption and emission of R^-_{eq} upon lowering the temperature can, therefore, be an indicator for the strong hydrogen bond to T203 frozen in its B-orientation as compared to the A-orientation in R^-_{neq}. These calculations show also that the absorption and fluorescence characteristics, referred to as blinking in single molecule spectroscopy and attributed to protonation changes of the anionic chromophore, may have additional biophysical determinants, e.g., larger scale rearrangements in the protein framework leading to hydrogen-bonding changes near the phenolate.

Limitations and extensions

The established sensitivity to details in the hydrogen-bond pattern indicates that the use of only one x-ray structure for the analysis imposes limitations to the range of validity of the developed models. The competition between the acidities of the phenolic part of the fluorophore and E222, linked via an internal hydrogen-bond net, was found to be responsible for the two-step titration, and E222 to be the final acceptor in ground- and excited-state proton transfer. However, a recent FTIR study (Van Thor et al., 1998) does not provide indications for a net change of the protonation state of a glutamate side chain upon change in ionization of the fluorophore. This behavior cannot yet be understood on the basis of the current investigation. The high concentration of GFP used in this FTIR study leads presumably to protein/buried water conformational changes, as mirrored by deviations from known GFP properties (e.g., increase of the neutral to anionic ratio from 6 to 13 at pH 8, the absence of signals indicative for the rearrangement of T203 and the isotope independence of the photoconversion rate [Van Thor et al., 1998]). The use of the dimeric structure (Yang et al., 1996) will provide a more adequate analysis of proton pathways under these experimental conditions. Particularly, the additional hydrogen bond between the phenolic end of the fluorophore and the imidazole group of H148 opens the option for an alternative proton acceptor. Beyond this, the altered structure and orientation of internal hydrogen-bonded chains of water molecules has to be taken into account. However, a recent x-ray structural study for the S65T mutant (Elsliger et al., 1999) seems to confirm the conservation of the neutral charge state of H148 even for the unprotonated fluorophore. But this study also shows that the orientation of H148 responds to fluorophore deprotonation at high pH and indicates that additional degrees of freedom for this residue should be included. Like oligomerization, side-chain substitutions will induce structural and hydrogen-bonding rearrangements not included in our discussion of the hypothetical mutants.

For our analysis, we assume the heterocyclic ring of the fluorophore to be unprotonated. This is consistent with experimental findings (Wachter et al., 1997; Niwa et al., 1996; Terry et al., 1995; Patterson et al., 1997; Elsliger et al., 1999) but at variance to quantum chemical calculations (Vojtyuk et al., 1998), which use modeled geometries for the chromophore. Using the same level of quantum chemical approach (INDO/S SCF-CI), but the x-ray-determined geometry of the fluorophoric core, the calculated absorption characteristics for phenol and phenolate form reproduce the experimental absorption wavelength as well as intensity ratios and dipole moment changes without invoking the additional protonation site. An extension of our analysis will include all four different protonation forms of the fluorophore. Experimental pH-titration of difunctional compounds in solution (or denatured protein) reveals only two apparent acid dissociation constants (Stewart, 1985), but not the relevant microscopic pK_a for the four single deprotonation steps. Their determination requires electronic structure calculations and the use of statistical thermodynamic concepts for the chromophore in solution and the protein. Additionally, the sensitivity of the results to the charge model used for ground and excited state of the fluorophore requires the determination of a set of partial charges optimized for use with electrostatic calculations (Sigfridsson and Ryde, 1998).

CONCLUDING REMARKS

The present work provides a first theoretical analysis of the coupling between chemical and conformational equilibria and a determination of the nature and relationship of protein microheterogeneities in GFP. The investigation is carried out on the basis of the recently solved x-ray (Brejc et al., 1997) structure for monomeric wild-type protein. Because the multiconformation calculations require an initial choice of possible conformations (Alexov and Gunner, 1997), a broader set of reorientational flexibility (e.g., including disorder in E222 and H148) and the inclusion of other internal protonation sites will improve the analysis. Guided by the atomic model, E222 has been suggested as proton acceptor in ground and excited state. Varying experimental conditions (e.g., the E222Q mutant [Jung et al., 1998] or aggregated proteins [Yang et al., 1996; Van Thor et al., 1998]) indicate the presence of alternative proton acceptors like H148 or a chain of water molecules. Applying this extended method to a variety of solved tertiary structures will, quite likely, provide additional insight into the biophysical basis of GFP. Our results provide essential information about classic degrees of freedom involved in the proton transfer process and are a prerequisite for evaluations of the exact barriers, taking the quantum mechanical character of the transferred particle into account.

We thank Dr. M. Thompson for making the quantum chemical program ARGUS available to us. The work was supported by the Deutsche Forschungsgemeinschaft (projects SFB 377/C2 and SFB 533/C2).

REFERENCES

Alexov, E. G., and M. R. Gunner. 1997. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* 74:2075–2093.

Antosiewicz, J., and J. A. McCammon. 1996. The determinants of pK_a's in proteins. *Biochemistry.* 35:7819–7833.

Baca, M., G. E. O. Borgstahl, M. Boissinot, P. M. Burke, R. W. DeWright, K. A. Slater, and E. D. Getzoff. 1994. Complete chemical structure of photoactive yellow protein: novel thioester-linked 4-hydroxycinnamyl chromophore and photocycle chemistry. *Biochemistry*. 33: 14369–14377.

Bala, P., P. Grochowski, B. Lesyng, and J. A. McCammon. 1996. Quantum-classical molecular dynamics. Models and applications. *In* Quantum Mechanical Simulation Methods for Studying Biological Systems. D. Bicout and M. Field, editors. Springer, Berlin. 119–156.

Balashov, S. P., R. Govindjee, E. S. Imasheva, S. Misra, T. E. Ebrey, Y. Feng, R. K. Crouch, and D. R. Menick. 1995. The two pK_a's of aspartate-85 and control of thermal isomerization and proton release in

the arginine-82 to lysine mutant of bacteriorhodopsin. *Biochemistry.* 34:8820–8834.

Balashov, S. F., E. S. Imasheva, R. Govindjee, and T. G. Ebrey. 1996. Titration of aspartate-85 in bacteriorhodopsin: what is says about chromophore isomerization and proton release. *Biophys. J.* 70:473–481.

Berendsen, H. J. C., and J. Mavri. 1996. Quantum dynamics simulation of a small quantum system embedded in a classical environment. *In* Quantum Mechanical Simulation Methods for Studying Biological Systems. D. Bicout and M. Field, editors. Springer, Berlin. 157–179.

Beroza, P., and D. A. Case. 1996. Including side chain flexibility in continuum electrostatic calculations of protein titration. *J. Phys. Chem.* 100:20156–20163.

Bokman, S. H., and W. W. Ward. 1981. Renaturation of *Aequorea* green-fluorescent protein. *Biochem. Biophys. Res. Comm.* 101:1372–1380.

Brejc, K., T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, and S. J. Remington. 1997. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. USA.* 94:2306–2311.

Brooks, C. L., III, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Bublitz, G., B. A. King, and S. G. Boxer. 1998. Electronic structure of the chromophore in green fluorescent protein (GFP). *J. Am. Chem. Soc.* 120:9370–9371.

Chattoraj, M., B. A. King, G. U. Bublitz, and S. G. Boxer. 1996. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl. Acad. Sci. USA.* 93:8362–8367.

Cometta-Morini, C., C. Scharnagl, and S. F. Fischer. 1993. Proton transfer to ubiquinone Q_B in the photosynthetic reaction center of *Rps. viridis*: the role of electrostatic interactions. *Int. J. Quant. Chem. Quant. Biol. Symp.* 20:89–106.

Denisov, V. P., K. Venu, J. Petersen, H. D. Hörlein, and B. Halle. 1997. Orientational disorder and entropy of water in protein cavities. *J. Phys. Chem. B.* 101:9380–9389.

Dickson, R. M., A. B. Cubitt, R. Y. Tsien, and W. E. Moerner. 1997. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature.* 388:355–358.

Ehrig, T., D. J. O'Kane, and F. G. Prendergast. 1995. Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett.* 367:163–166.

Elsliger, M.-A., R. M. Wachter, G. T. Hanson, K. Kallio, and S. J. Remington. 1999. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry.* 38:5296–5301.

Gilson, M. K., K. A. Sharp, and B. Honig. 1987. Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comp. Chem.* 9:327–335.

Gutman, M., and E. Nachliel. 1990. The dynamic aspects of proton transfer processes. *Biochim. Biophys. Acta.* 1015:391–414.

Haupts, U., S. Maiti, P. Schwille, and W. W. Webb. 1998. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proc. Natl. Acad. Sci. USA.* 95: 13573–13578.

Heim, R., A. B. Cubitt, and R. Y. Tsien. 1995. Improved green fluorescence. *Nature.* 373:663–664.

Heim, R., and R. Y. Tsien. 1996. Engineering green fluorescent protein for improved brightness, longer wavelength and fluorescence resonance energy transfer. *Curr. Biol.* 6:178–182.

Honig, B., and A. Nicholls. 1995. Classical electrostatic in biology and chemistry. *Science.* 268:1144–1149.

Jung, G., J. Wiehler, W. Göhde, J. Tittel, T. Basché, B. Steipe, and C. Bräuchle. 1998. Confocal microscopy of single molecules of the green fluorescent protein. *Bioimaging* 6:54–61.

Karplus, M., and G. A. Petsko. 1990. Molecular dynamics simulations in biology. *Nature.* 347:631–639.

Kneen, M., J. Farinas, Y. Li, and A. S. Verkman. 1998. Green fluorescent protein as a noninvasive intracellular pH indicator. *Biophys. J.* 74: 1591–1599.

Lanyi, J. 1997. Mechanism of ion transport across membranes: bacteriorhodopsin as a prototype for proton pumps. *J. Biol. Chem.* 272: 31209–31212.

Laria, D., G. Ciccotti, M. Ferrario, and R. Kapral. 1992. Molecular-dynamics study of adiabatic proton-transfer reactions in solution. *J. Chem. Phys.* 97:378–388.

Lossau, H., A. Kummer, R. Heinecke, F. Pöllinger-Dammer, C. Kompa, G. Bieser, T. Jonsson, C. M. Silva, M. M. Yang, D. C. Youvan, M. E. Michel-Beyerle. 1996. Time-resolved spectroscopy of wild-type and mutant green fluorescent proteins revealed excited state deprotonation consistent with fluorophore-protein interactions. *Chem. Phys.* 213:1–16.

Marrink, S. J., F. Jähnig, and H. J. C. Berendsen. 1996. Proton transport across transient single-file water pores in a lipid membrane studied by molecular dynamics simulations. *Biophys. J.* 71:632–647.

Misteli, T., and D. L. Spector. 1997. Applications of the green fluorescent protein in cell biology and biotechnology. *Nature Biotechnol.* 15: 961–964.

Nagle, J. F., M. Mille, and J. J. Moritz. 1980. Theory of hydrogen bonded chains in bioenergetics. *J. Phys. Chem.* 72:3959–3971.

Niwa, H., S. Inouye, T. Hirano, T. Matsuno, S. Kojima, M. Kubota, M. Ohashi, and T. I. Tsuji. 1996. Chemical nature of the light emitter of the *Aequorea* green fluorescent protein. *Proc. Natl. Acad. USA.* 93: 13617–13622.

Ormö, M., A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, S. J. Remington. 1996. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science.* 273:1392–1395.

Palm, G. J., A. Zdanov, G. A. Gaitanaris, R. Stauber, G. N. Pavlakis, and A. Wlodawer. 1997. The structural basis for spectral variations in green fluorescent protein. *Nature Struct. Biol.* 4:361–365.

Patterson, G. H., S. M. Knobel, W. D. Sharif, S. R. Kain, and D. W. Piston. 1997. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys. J.* 73:2782–2790.

Pomès, R., and B. Roux. 1998. Free energy profiles for H^+ conduction along hydrogen-bonded chains of water molecules. *Biophys. J.* 75: 33–40.

Ripoll, D. R., Y. N. Vorobjev, A. Liwo, J. A. Vila, and H. A. Scheraga. 1996. Coupling between folding and ionization equilibria: effects of pH on the conformational preferences of polypeptides. *J. Mol. Biol.* 264: 770–778.

Robey, R. B., O. Ruiz, A. V. P. Santoz, J. Ma, F. Kear, L.-L. Wang, C.-J. Li, A. A. Bernardo, and J. A. L. Arruda. 1998. pH-dependent fluorescence of a heterologously expressed *Aequorea* green fluorescent protein mutant: in situ spectral characteristics and application to intracellular pH estimation. *Biochemistry.* 37:9894–9901.

Scharnagl, C., and S. F. Fischer. 1996. Conformational flexibility of arginine-82 as source for the heterogeneous and pH-dependent kinetics of the primary proton transfer step in the bacteriorhodopsin photocycle: an electrostatic model. *Chem. Phys.* 212:231–246.

Sharp, K. A., and B. Honig. 1990. Electrostatic interactions in macromolecules: theory and applications. *Ann. Rev. Biophys. Biophys. Chem.* 19:301–332.

Sigfridsson, E., and U. Ryde. 1998. Comparison of methods for deriving atomic partial charges from the electrostatic potential and moments. *J. Comp. Chem.* 19:377–395.

Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.

Stewart, R. 1985. The Proton: Application to Organic Chemistry. Academic Press, New York.

Stryer, L. 1988. Biochemistry. W. H. Freeman and Company, New York.

Terry, B. R., E. K. Matthews, and J. Haseloof. 1995. Molecular characterisation of recombinant green fluorescent protein by fluorescence correlation microscopy. *Biochem. Biophys. Res. Comm.* 217:21–27.

Thompson, M. A., and M. Zerner. 1991. A theoretical examination of the electronic structure and spectroscopy of the photosynthetic reaction center from *Rhodopseudomonas viridis*. *J. Am. Chem. Soc.* 113: 8210–8215.

Tsien, R. Y. 1998. The green fluorescent protein. *Annu. Rev. Biochem.* 67:509–544.

Van Thor, J. J., A. J. Pierik, I. Nugteren-Roodzant, A. Xie, and K. J. Hellingwerf. 1998. Characterization of the photoconversion of green fluorescent protein with FTIR spectroscopy. *Biochemistry*. 37: 16915–16921.

Van Vlijmen, H. W. T., M. Schäfer, and M. Karplus. 1998. Improving the accuracy of protein pK_a calculations: conformational averaging versus the average structure. *Proteins: Struct. Funct. Genet.* 33:145–158.

Voityuk, A. A., M.-E. Michel-Beyerle, and N. Rösch. 1998. Quantum chemical modeling of structure and absorption spectra of the chromophore in green fluorescent proteins. *Chem. Phys.* 231:13–27.

Wachter, R. M., B. A. King, R. Heim, K. Kallio, R. Y. Tsien, S. G. Boxer, and S. J. Remington. 1997. Crystal structure and photodynamic behavior of the blue emission variant Y66H/Y145F of green fluorescent protein. *Biochemistry*. 36:9759–9765.

Wachter, R. M., M.-A. Elsliger, K. Kallio, G. T. Hanson, and S. J. Remington. 1998. Structural basis of spectral shifts in the yellow-emission variant of green fluorescent protein. *Structure.* 6:1267–1277.

Ward, W. W. 1981. Properties of the coelenterate green-fluorescent protein. *In* Bioluminescence and Chemiluminescence. M. A. DeLuca and W. D. McElroy. Academic Press, New York. 235–242.

Ward, W. W., H. J. Prentice, A. F. Roth, C. W. Cody, and S. C. Reeves. 1982. Spectral perturbation of the *Aequorea* green-fluorescent protein. *Photochem. Photobiol.* 35:803–808.

Ward, W. W., and S. H. Bokman. 1982. Reversible denaturation of *Aequorea* green-fluorescent protein: physical separation and characterization of the renatured protein. *Biochemistry.* 21:4535–4540.

Warshel, A. 1982. Dynamics of reactions in polar solvent. Semiclassical trajectory studies of electron-transfer and proton-transfer reactions. *J. Phys. Chem.* 86:2218–2224.

Warshel, A. 1986. Correlation between the structure and efficiency of light-induced proton pumps. *Methods Enzymol.* 127:578–587.

Warshel, A. 1991. Computer Modeling of Chemical Reactions in Enzymes and Solutions. Wiley-Interscience, New York.

Yang, F., L. G. Moss, and G. N. Phillips, Jr. 1996. The molecular structure of green fluorescent protein. *Nature Biotechnol.* 14:1246–1251.

Yang, T. T., P. Sinai, G. Green, P. A. Kitts, Y. T. Cheng, L. Lybarger, R. Chervenak, G. H. Patterson, D. W. Piston, and S. R. Kain. 1998. Improved fluorescence and dual color detection with enhanced blue and green variants of the green fluorescing protein. *J. Biol. Chem.* 273:8212–8216.

You, T. J., and D. Bashford. 1995. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys. J.* 69:1721–1733.