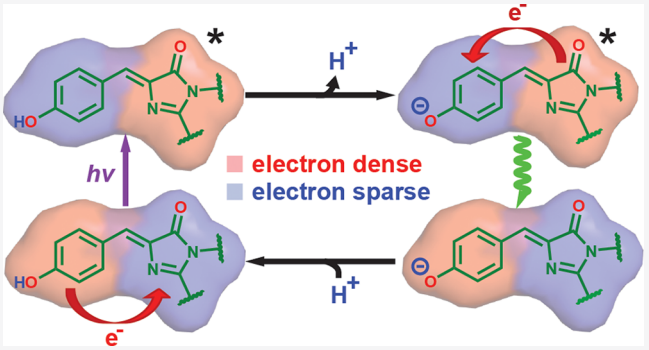
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Mechanism of Color and Photoacidity Tuning for the Protonated Green Fluorescent Protein Chromophore

Chi-Yun Lin* and Steven G. Boxer*

Cite This: J. Am. Chem. Soc. 2020, 142, 11032−11041
Read Online

ABSTRACT: The neutral or A state of the green fluorescent protein (GFP) chromophore is a remarkable example of a photoacid naturally embedded in the protein environment and accounts for the large Stokes shift of GFP in response to near UV excitation. Its color tuning mechanism has been largely overlooked, as it is less preferred for imaging applications than the redder anionic or B state. Past studies, based on site-directed mutagenesis or solvatochromism of the isolated chromophore, have concluded that its color tuning range is much narrower than its anionic counterpart. However, as we performed extensive investigation on more GFP mutants, we found that the color of the neutral chromophore can be more sensitive to protein electrostatics than can the anionic counterpart. Electronic Stark spectroscopy reveals a fundamentally different electrostatic color tuning mechanism for the neutral state of the chromophore that demands a three-form model as compared to that of the anionic state, which requires only two forms (*J. Am. Chem. Soc.* **2019**, *141*, 15250−15265). Specifically, an underlying zwitterionic charge-transfer state is required to explain its sensitivity to electrostatics. As the Stokes shift is tightly linked to excited-state proton transfer (ESPT) of the protonated chromophore, we infer design principles of the GFP chromophore as a photoacid through the color tuning mechanisms of both protonation states. The three-form model could also be applied to similar biological and nonbiological dyes and complements the failure of the two-form model for donor−acceptor systems with localized ground-state electronic distributions.



1. INTRODUCTION

The jellyfish *Aequorea victoria* green fluorescent protein (avGFP) exhibits two visible absorption bands: a major band (A) at 395 nm and a minor band (B) at 475 nm.[1] The A and B bands correspond to the protonated neutral and the deprotonated anionic chromophores, respectively (Figure 1). The hallmark of the neutral state is its large Stokes shift (~5600 cm^{-1}), which is attributed to its ability to generate the deprotonated intermediate (I) state via excited-state proton transfer (ESPT) upon photoexcitation (Figure 1).[2] The chromophore's photoacidity and the associated proton transfer process(es) have prompted many subsequent studies. Within those studies, GFP serves either as a unique model system with a well-defined hydrogen-bonding network that accepts a proton following photoexcitation[3−7] or as a naturally occurring photoacid from which inspiration is drawn for designing small-molecule analogues.[8−10] Because of phototoxicity from near-ultraviolet irradiation, the lower cross section, and incompatibility with the standard fluorescein filter sets for practical applications,[11] much less attention has been paid to the A state as compared to the B state, and the latter is the main optimization target for imaging.[12] The Stark tuning rate (electronic redistribution upon excitation) is a critical quantity for understanding both the color tuning behavior[13,14] and the photoacidity[15] of the neutral chromophore in response to environmental mutations; however, there has been only limited investigation of the Stark tuning rate for the neutral chromophore.[16]

Previously, we have elucidated how the photophysical and electro-optical properties of the anionic B-state GFP chromophore are modulated by its environment through electrostatics.[13] We applied the Marcus−Hush theory, which is based on electronic coupling between the intuitive resonance structures with explicit consideration of vibronic coupling to the bond length alternation (BLA) mode, to explain the strong and monotonic correlations among properties including electronic excitation energy, Stokes shift, and Stark tuning rate. We then identified the relative energetics of the resonance forms, or the driving force, to be the dominant factor that is tuned by environmental electrostatics. The additivity of the

Received: March 11, 2020
Published: May 26, 2020

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

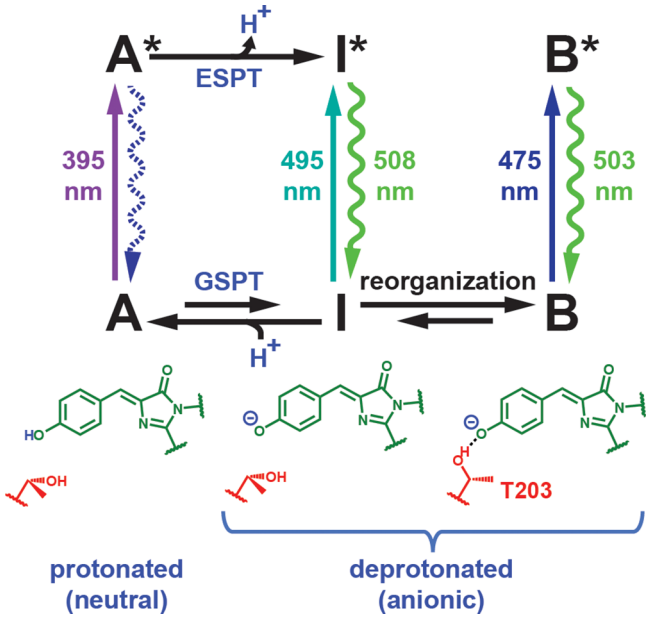


Figure 1. Canonical ESPT scheme for avGFP with spectroscopically characterized protonation states (A, B, and I states) and their structures (with T203 shown in red). The wavelengths associated with the excitation and emission (straight and curved vertical arrows, respectively) are measured at room temperature,[1,2] except for the absorption maximum of the I state, which can only be identified as a small band at low temperature.[1,17] Ground-state processes, including proton transfer (GSPT) and environmental reorganization, are required to reach an equilibrium among the states. The A* emission (shown with a dashed curved arrow) is largely suppressed at steady state due to efficient ESPT.[2] T203 is not the only residue that undergoes structural change following ESPT; see Figure 4 in ref 18.

driving force from combinations of mutations allowed us to demonstrate the predictive power of the model and infer design principles for the anionic state. In this study, we continue this effort but instead focus on the neutral state of the chromophore. We find that the Stark tuning rates for the A state, which are obtained via electronic Stark spectroscopy,[19] can sometimes exceed those of the B state, and protonation completely alters the qualitative trend between the absorption maximum and the Stark tuning rate previously documented for the B state. The simple resonance picture, which worked well for the anionic chromophore, breaks down, but, as shown in the following, inclusion of a third form is sufficient to quantitively capture the new correlation while inheriting the minimalistic spirit of the Marcus−Hush model. Combined with our understanding of the anionic chromophore,[13] we can track the electron flow within the chromophore before and after ESPT and reveal design principles for the excited state pK_a, or pK_a^*, in the protein environment. In the subsequent discussion, to clarify the term "states" in different contexts, "states" are reserved for electronic states (e.g., S_0 and S_1) and protonation states (e.g., A and B states) that can be experimentally observed and are energy eigenstates (adiabatic states), while "(resonance) forms" refer to resonance structures, which are convenient diabatic states for chemical intuition and whose linear combinations can well approximate the electronic eigenstates S_0 and S_1. As these forms are thought to be consistent in different environments,[13] they provide an ideal set of bases for understanding modulation of electronic distribution caused by environmental effects.

2. RESULTS AND DISCUSSION

2.1. Mutant and Variant Design. In this work, we take advantage of a subset of GFP mutants from our B-state color-tuning work[13] that also populate the A state at nondenaturing pH's (Tables 1 and S1; see Figure 2 for the critical residues in the protein environment). The mutations involving R96 and noncanonical amino acids were enabled by the semisynthetic reconstitution strategy of split GFP as described previously,[13,20] in combination with amber suppression. Because T203, H148, and R96 in the GFP environment can stabilize the anionic chromophore via hydrogen bonding, any mutations at these positions that remove the hydrogen bond(s) disfavor chromophore deprotonation, which can also be achieved by negatively supercharging the β-barrel (supercharged −30, where surface residues are extensively replaced with negatively charged amino acids).[13,21] The model chromophore 4-hydroxybenzylidene-1,2-dimethylimidazolinone (HBDI) in ethanol was also examined to complement the protein study.

Table 1. GFP Mutation Sites Chosen for This Study to Sample the A State in a Wide Range of Protein Environments[a]

mutation sites	amino acids
	Environmental Mutants
R96	R, M
T203	T, V, H, F, Y, 4-F_1F, F_5F, 4-NH_2F, 3-OCH_3Y
	Chromophore Variants
S65	S, T

[a]Noncanonical amino acids are abbreviated as follows: 4-fluorophenylalanine is 4-F_1F, 2,3,4,5,6-pentafluorophenylalanine is F_5F, 3-methoxytyrosine is 3-OCH_3Y (3-OMeY), and 4-aminophenylalanine is 4-NH_2F (see also Table S1).

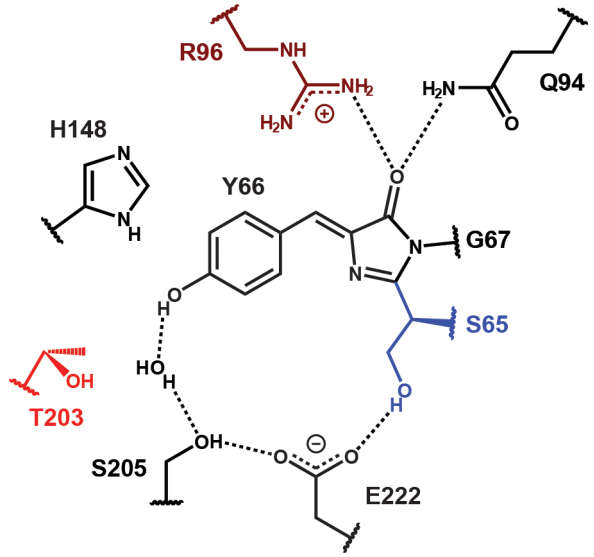


Figure 2. Surrounding residues of the GFP chromophore in the protonated state (A state) for avGFP, whose residue 65 is a serine.[18] The ESPT chain is explicitly shown here. The colored residues represent those mutated in this work. Note that E222 mutants cannot be used for investigating the A state as the mutation suppresses chromophore protonation.[13]

2.2. 77 K Absorption Spectra and Color Tuning. We first focus on the absorption spectra at 77 K and the color tuning behavior of A states. Even though the A bands are rather

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

featureless at room temperature, a clear vibronic progression can be seen at 77 K (Figure 3A–C). Substructure can be resolved by

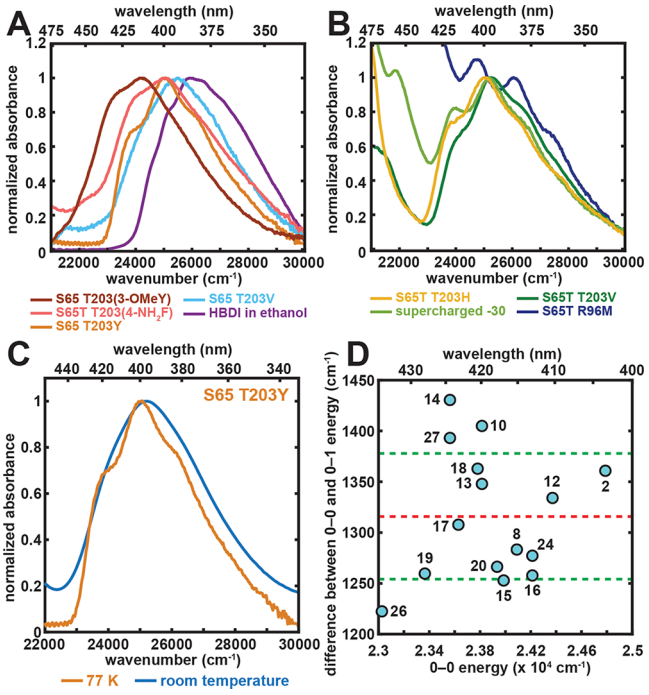


Figure 3. Spectral properties and color tuning of the A state. (A and B) Representative 77 K absorption spectra of GFP mutants and the model chromophore HBDI in ethanol. The spectra are all normalized at their A-state peak maxima. The absorption maximum, later used in Figure 4, corresponds to the wavenumber at which the largest absorbance is observed for each mutant and the model chromophore, except for the R96M mutant, for which spectral deconvolution is done through Stark spectroscopy to extract the A-state absorption maximum (Figure S1). For GFP mutants, the samples of S65 T203(3-OMeY) and S65T T203(4-NH_2F) were at pH 5.0 and 8.0, respectively. The rest were prepared at pH 10.0. Panel B includes constructs that contain the longer wavelength B band, corresponding to the anionic chromophore, in the spectral region of interest.[13] (C) Representative room-temperature and 77 K absorption spectra for the A state of S65 T203Y. (D) Difference between the 0–0 and 0–1 energy plotted against 0–0 electronic excitation energy for GFP mutants obtained from a second derivative analysis (Figure S2). The x–y coordinates and the numerical labels for the species are listed in Table S2. The values are within a quite narrow range with an average of 1320 ± 60 cm^{-1}. The red and green dashed lines represent the mean value and ±1 σ, respectively.

taking the second derivative of the absorption, and the results resemble those from B states,[13] where the energy spacing (1320 ± 60 cm^{-1}) between the presumed 0–0, 0–1, and 0–2 electronic transitions agrees well with the vibrational frequency assigned to the BLA mode (Figures 3D and S2; see also Figure S23 in ref 13; an exception is the 610 cm^{-1} spacing for neutral HBDI in ethanol; see Figure S2). Because the vibronic peaks are rather broad, the assignment to the BLA mode is not definitive and could be mixed with other normal modes with similar frequencies. On the basis of the relative intensity of these vibronic peaks, we can assert that the Huang–Rhys factors for the A states consistently fall between 1 and 2, showing spectral features similar to those of the anionic HBDI model chromophores in water/glycerol.[13]

The qualitative trend of A state color tuning follows that of the B state (Table S3): any mutation that eliminates the positive charge of R96[22−24] results in a blue-shift (see also ref 25), while mutating T203 to an aromatic residue that π–π stacks with the chromophore's phenol moiety red-shifts the absorption maximum. This suggests that the same direction of electron flow from phenol(ate) to imidazolinone during excitation can be assigned to both protonation states.[13,26] Within the π–π stacking series, as we modulate the π-system's electron density at position 203 from electron-deficient to electron-rich by using electron-withdrawing and -donating groups, respectively, a clear red-shift is observed. The corresponding color tuning range is appreciably larger for the A state than for the B state (1550 cm^{-1} vs 650 cm^{-1}, respectively, Table S3).[13] In contrast, for neutral and anionic HBDI in various organic solvents, the color tuning range is much smaller for the former as compared to the latter (962 cm^{-1} vs 2755 cm^{-1}, respectively).[27] As such, we see a drastic change in color tuning behaviors for the chromophore in its two protonation states, which hints at a fundamental difference in the magnitude of charge redistribution upon excitation between the A and B states. Consequently, we measured the Stark tuning rates using electronic Stark spectroscopy to uncover the A-state color tuning mechanism.

2.3. Stark Tuning Rate and Color Tuning Mechanism.

Stark spectra of the A states are dominated by the second derivative of their corresponding absorption spectra (Figure S1 and Table S4), which indicates a significant contribution from the linear Stark effect characterized by a Stark tuning rate $f\Delta\mu$[19] (where f is the local field factor, which is necessary due to the larger field experienced by the chromophore as compared to the externally applied field based on the unavoidable polarization effect of the chromophore environment; see section S6 in ref 13). The largest Stark tuning rate for the A states (24 D) from various mutants is twice as large as the B-state counterpart (12 D),[13] suggesting the A-state excitation is strongly associated with charge-transfer character, and its electronic excitation energy can be more sensitive to electrostatic color tuning. The large Stark tuning rates $f\Delta\mu$ of the A state for some protein mutants may seem surprising at first. However, the calculated Stark tuning rates $\Delta\mu$ as large as 10 D have been suggested by Filippi et al. when accounting for nearby residues in GFP using quantum mechanics/molecular mechanics (QM/MM) methods,[28] in good agreement with our experimental results because f is likely to be close to 2 (see section S6 in ref 13). As gas-phase calculations yield Stark tuning rates between 2 and 2.5 D,[28,29] the protein environment is indeed a crucial factor that increases the Stark tuning rate.

From the prior investigation conducted by Drobizhev et al.[14] and us,[13] it is informative to plot the absorption maximum against the Stark tuning rate for the A state to reveal how electron delocalization within the neutral chromophore is modulated by electrostatics (Figure 4A). Here, we follow our previous B-state study and plot the S65 and S65T mutants together to illustrate that the presence or absence of the methyl group does not modify the intrinsic response of the π system to the protein environment (Figure 4A and B).[23,30] It can be readily seen that the A-state correlation plot indicates larger Stark tuning rates for redder species as compared to the B-state counterpart (Figure 4B).[13]

To understand this trend, as a first approximation it might seem reasonable to treat the A-state chromophore as an anionic chromophore experiencing the electric field exerted by the proton and assume that the naïve picture of a two-form resonance (P and I forms, Figure 4D), that is, interaction between two distinct resonance forms, still holds. In this picture,

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

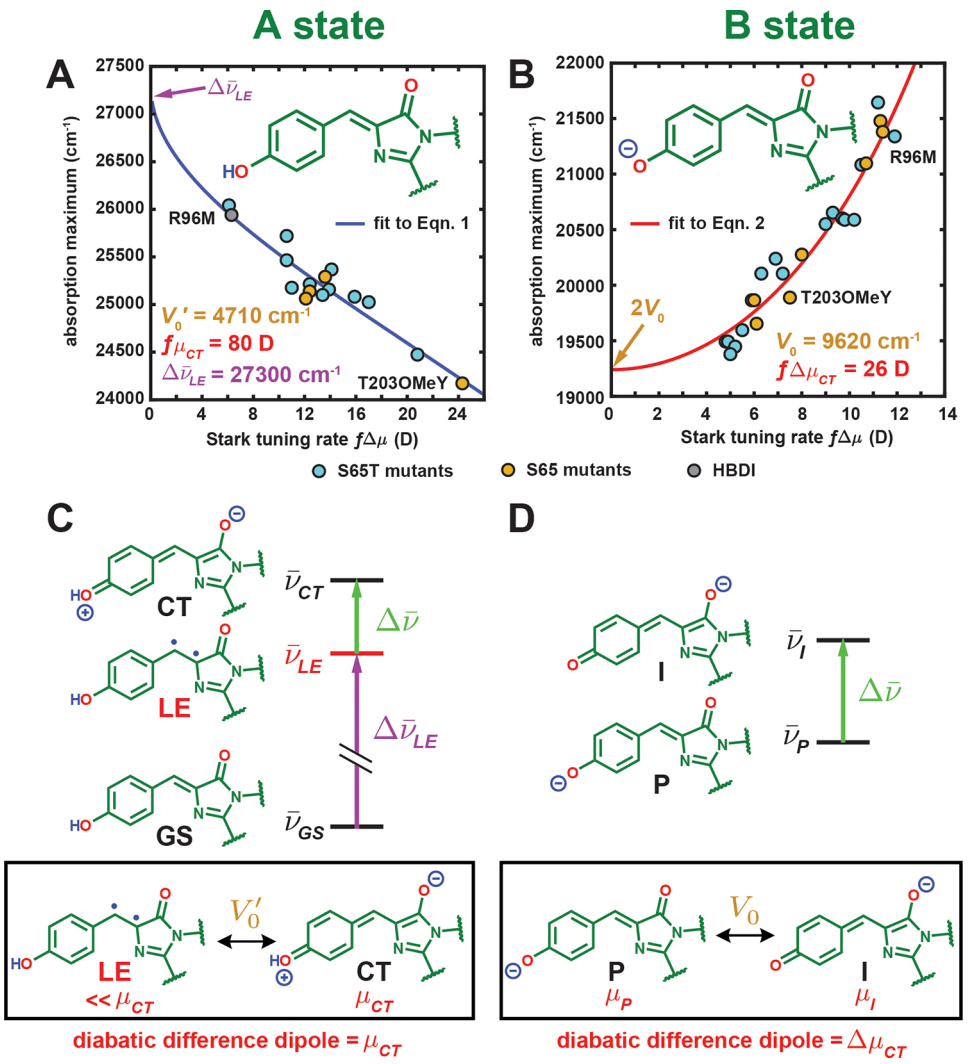


Figure 4. Correlation between the absorption maximum and Stark tuning rate for GFP mutants and HBDI in the (A) A and (B) B states at 77 K, respectively. Panel B is reproduced from Figure 9 in ref 13 including only the S65T and S65 mutants' data. The curves for panels A and B come from fitting all data (S65T for the latter) to eqs 1 and 2, respectively. See Figure S3 for an identical figure reproduced with numerical labels to identify the species. (C) The three-form model for the A state and (D) the two-form model for the B state. The two diabatic forms with substantial mixing are shown in the black box below each model with the corresponding electronic coupling and diabatic difference dipole. For both models, only $\Delta\bar{\nu}$ is assumed to be tunable by the environment among all parameters. Note that the LE form has a nontrivial dipole moment, but it is much smaller than μ_{CT} (5 D[29] vs 40 D for the CT form if $f = 2$), so we can approximately assign μ_{CT} to the dipole moment of the CT form. The diradicaloid structure illustrated for the LE form should not be treated as exact, but rather it is meant to represent the locally excited and nonzwitterionic nature of this third form (see section 2.3 for further discussion).

the proton biases the driving force $\Delta\bar{\nu}$ through electrostatics by strongly stabilizing the P form (now ground-state, GS, form after proton attachment) over the I form (now charge-transfer (CT) form, Figure 4C). The direct logical consequence is that the data points of the A-state absorption maximum versus Stark tuning rate (Figure 4A) should simply fall on the corresponding correlation line obtained from the B state in Figure 4B, contrary to what is observed for the A state. This is not at all surprising because the phenol oxygen is in fact covalently linked to the proton.

On the basis of Olsen's multiconfigurational calculation on a series of modified neutral GFP chromophores,[31] there exists a third form close in energy with the CT form (Figure 4C). This form is locally excited (LE) and mixes with the CT form to yield two adiabatic states, the lower of which is the S_1 state. Because of the large energy gap between the CT and GS forms caused by proton stabilization, the latter stays mostly intact upon coupling and becomes the S_0 state. We can accordingly set up a minimal three-form model, in which the GS and LE forms are neutral with their relative energies largely unaltered across different environments, while the CT form's energy is strongly affected by the environment due to its dipolar nature (Figure 4C). In the context of Figure 4C, the energy gap between LE and GS ($\Delta\bar{\nu}_{LE} \equiv \bar{\nu}_{LE} - \bar{\nu}_{GS}$) is assumed to be constant, while that between the CT and LE ($\Delta\bar{\nu} \equiv \bar{\nu}_{CT} - \bar{\nu}_{LE}$) is the only tunable parameter by electrostatics from the environment. We expect the latter to be positive because the charge-separated CT state should be higher in energy than the neutral LE state, which we will verify later. An electronic coupling V_0' is required to account for the mixing between LE and CT, and the notation is primed to distinguish it from the electronic coupling V_0 between the P and I forms of the B state (Figure 4D). For simplicity, we neglect the coupling between GS and the other two forms due to the large energy differences. To calculate the Stark tuning rate associated with the

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

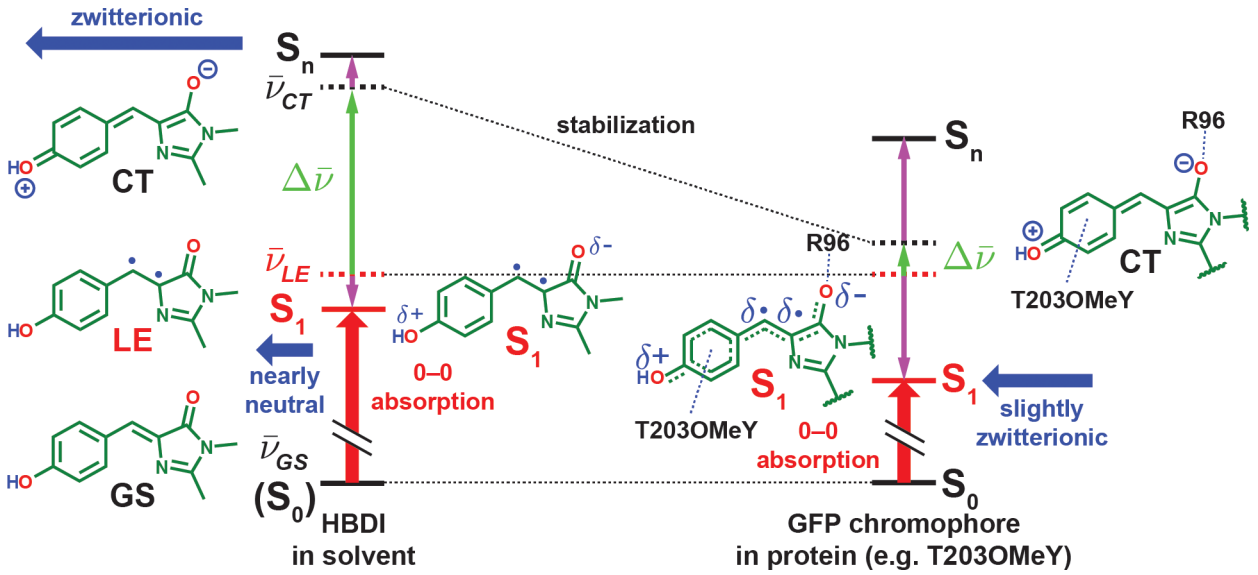


Figure 5. Electrostatic color tuning mechanism of the protonated GFP chromophore, HBDI in solvent, and the covalently linked chromophore in proteins, illustrated qualitatively in different environments (energies not drawn to scale). The energy levels of the diabatic forms and adiabatic states (electronic eigenstates) are represented with dashed and solid lines, respectively. Because the GS form stays intact upon mixing and becomes the S_0 state (denoted with parentheses on the left), we represent its energy levels with solid lines. The energy changes before and after mixing are shown with vertical purple arrows. The dipole moments of the CT form and the resulting S_1 states are depicted with horizontal blue thick arrows, while the 0−0 electronic excitation energy is shown as a vertical red thick arrow in each scenario. The protein environment is more capable of stabilizing the zwitterionic CT form with R96 and the residue at position T203 than solvents, thereby bringing the CT form closer in energy to the LE form and leading to stronger mixing, which subsequently lowers the S_1 energy and imparts stronger charger-transfer character to the S_1 state and thus a larger Stark tuning rate. The resulting CT character of the S_1 state ranges from 7−31% for the mutants we sampled (Table S4). The diradicaloid character for the LE form and the S_1 state should not be taken literally (see section 2.3 for further discussion).

excitation, we also assign a large dipole moment $\vec{\mu}_{CT}$ with a magnitude of μ_{CT} ($\equiv|\vec{\mu}_{CT}|$) for the CT form relative to the other two forms, which possess similar yet relatively negligible electronic dipoles (vide infra). Therefore, only four essential parameters, $\Delta\bar{\nu}$, $\Delta\bar{\nu}_{LE}$, V_0', and μ_{CT}, are required, and only $\Delta\bar{\nu}$ is influenced by the environment. Again, to avoid potential confusion, we refer to diabatic states as "forms" and the adiabatic states as "states".

As we will elaborate later, the S_1 state for the chromophore in vacuo or in solvents is almost exclusively LE-like, and thus inferring the nature of LE form is possible from excited-state calculations. The LE form has been proposed as a diradicaloid structure at the methine bridge according to Olsen's computational study.[31] It has likewise been introduced in the context of retinal and other polyene chromophores as the $2A_g$ state (as opposed to the $1B_u$ state with a dominant charge-transfer character).[32−40] While for the neutral all-*trans* polyene, the diradicaloid state is indeed lower in energy than the charge-separated state, the former tends to have a small oscillator strength when excited from the ground state ($1A_g$ for all-*trans* polyenes) due to the unfavorable parity.[37] The assignment of the diradicaloid behavior seems to be in contradiction to the large extinction coefficients for the protonated GFP chromophore observed from experiments and calculations,[28,29,41−43] where the S_0−S_1 transition is consistently designated to be of π−π* character. Some of these studies offer results for us to re-examine the exact nature of the LE form.

First, consistent with our model, there is nearly no net charge transfer across the neutral GFP chromophore in vacuo upon excitation,[28,29,41] so the corresponding LE form does not have a strong CT character and its name is thereby justified. In other words, the GS and LE forms share similar dipole moments (~5 D[29]). Our model is flexible enough to accommodate these nontrivial dipole moments through a constant offset in dipole moments for all three forms (and three resulting states after mixing), because our observable is the difference dipole moment (i.e., Stark tuning rate) between S_0 and S_1 states. The same flexibility also analogously applies to a constant energy shift in all energy levels as we only measure the difference energy (i.e., electronic excitation energy) between S_0 and S_1. Moreover, the calculated difference electron density between S_0 and S_1 states suggests localized π electron redistribution at the rings and/or the methine bridge,[41,43] which is still "locally excited" but slightly more delocalized than the diradicaloid structure would imply. Given that the GFP chromophore is two rings connected by a short methine bridge, whose symmetry is grossly different from C_{2h} for long all-*trans* polyenes, it is conceivable that the LE form is essentially excited at the rings (excitonic state) with a large oscillator strength, which is more similar to the heterodimeric special pair of the bacterial reaction centers[44] (section S5) and other donor−acceptor systems with short bridge(s).[45,46] This assignment with localized excitation at the rings and/or bridge rather than just at the bridge for the LE form is consistent with a large π−π* character, a high oscillator strength, and still retains its small dipole moment. However, because the corresponding resonance form is not exactly known, we will still illustrate the LE form with a diradicaloid structure in the figures to represent its locally excited and nonzwitterionic nature with this caveat in mind. As Olsen's calculation is largely inspired by the allylic model[31] that could imply localized radicals, the diradicaloid structure may actually represent a subset of neutral valence bond structures that allow for radical delocalization within the rings and the conundrum could likely be resolved as such. Because we aim to present a physically reasonable model that captures the global charge distribution (i.e., dipolar or not) rather than the exact wave function for each electronic state of the

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

chromophore, this ambiguity is not relevant to our subsequent analysis.

By including the LE form as a third form, this model already qualitatively predicts the opposite trend from that inferred from the two-form Marcus−Hush theory that was used to describe the deprotonated B state as in ref 13: if the environment stabilizes the CT form and brings it closer in energy to the LE form, the LE form mixes more strongly with the CT form and leads to a decrease in electronic excitation energy and a larger dipolar character for S_1 (Figure 5). The quantitative correlation of the Stark tuning rate $f\Delta\mu$ and absorption maximum $\Delta\bar{\nu}_{abs}$ can be deduced as (section S4):

$$\bar{\nu}_{abs} = \Delta\bar{\nu}_{LE} - V_0' \frac{2\frac{f\Delta\mu}{f\mu_{CT}}}{\sqrt{1-\left(2\frac{f\Delta\mu}{f\mu_{CT}}-1\right)^2}} \quad (1)$$

The vibronic shift between 0−0 electronic excitation energy and absorption maximum (0−1 transition) is absorbed by the $\Delta\bar{\nu}_{LE}$, so the use of either observable for $\Delta\bar{\nu}_{abs}$ would not affect the parameters in the second term on the right of eq 1 because the shift is roughly constant (Figures 3D and S2). The absorption maximum is chosen to facilitate comparison between 77 K and room-temperature spectra, for which the vibronic features are poorly resolved (Figure 3C). Using this equation, we obtain an excellent fit to the data in Figure 4A, and we can thus determine $\Delta\bar{\nu}_{LE}$, V_0', and $f\mu_{CT}$ to be 27 300 ± 100 cm^{-1}, 4710 ± 50 cm^{-1}, and 80 ± 4 D, respectively. For reference, the corresponding correlation for the anionic chromophore derived from the two-form model (Figure 4D) is

$$\bar{\nu}_{abs} = \frac{2V_0}{\sqrt{1-\left(\frac{f\Delta\mu}{f\Delta\mu_{CT}}\right)^2}} \quad (2)$$

which is exactly eq 2 in ref 13 and to which the data in Figure 4B are fit.

From eq 1, $\Delta\bar{\nu}_{LE}$, corresponding to 366 nm, is the predicted bluest possible absorption maximum through electrostatic color tuning of the neutral chromophore. The bluest absorption maximum among our protein mutants is 384 nm (from R96M); however, studies of neutral HBDI absorption in various solvents from Tolbert and colleagues show a tight range of absorption maxima between 360 and 373 nm,[27] which is comparable with our prediction. V_0' is about one-half of V_0 between the P and I forms of the B state (9620 cm^{-1}, Figure 4B), while $f\mu_{CT}$ is roughly 3 times as large as the intrinsic dipole moment difference $f\Delta\mu_{CT}$ between the P and I forms of the B state (26 D).[13] If f is approximately 2 as suggested by comparing previous experiments from us and Drobizhev et al. (see section S6 in ref 13), μ_{CT} corresponds to two opposite elementary charges with a separation distance of 8.3 Å for the CT form, which closely matches the actual O−O distance of the chromophore (8.7 Å). This suggests that the CT form is indeed fully charge-separated, as opposed to the smaller $\Delta\mu_{CT}$ from the B state, which is due to strong electron delocalization (section S9 in ref 13). The corresponding $\Delta\bar{\nu}$ for each environment is also estimated via eq S5 (Table S4) and shows that the CT form is always higher in energy than is the LE form, which justifies our aforementioned claim.

It is interesting that the solvatochromism of the neutral HBDI never recapitulates the color range observed from the protein mutants, unlike the anionic counterpart.[27] Because the absorption maxima in solvents are all close to $\Delta\bar{\nu}_{LE}$, our model suggests that solvents are much less effective at stabilizing the dipolar CT form than is the GFP environment. This results in a consistently large energy gap $\Delta\bar{\nu}$ between CT and LE in the solvents (>$2V_0'$), which leads to minimal CT and LE mixing, so no appreciable solvatochromic shift is observed (Figure 5, left). This is not surprising because the organization of solvent molecules maximally stabilizes the GS form and is unable to simultaneously solvate the CT form with a significantly larger dipole moment. On the other hand, the GFP environments are preorganized and can better stabilize the CT form via R96 and electron-donating residues at the 203 position, which brings the CT form close to the LE form and leads to substantial mixing that results in a much wider range of absorption maxima (Figure 5, right). This reconciliation suggests that one should be cautious when extrapolating the color tuning behavior from solvatochromic studies on the absorption of dyes to proteins.

While the neutral HBDI has a weak solvatochromism for absorption, if the CT form is indeed present and closely coupled to the LE form, the ability for polar solvents to reorganize and stabilize the CT form should impart HBDI with strong solvatochromism for emission (solvatofluorochromism). This is in fact not that easy to test, because HBDI in solutions are fraught with side photoreactions such as photoisomerization and, to a much lesser extent, ESPT.[8] However, for neutral HBDI analogues that suppress these side reactions, solvatofluorochromism ranges as wide as 95 nm (~3700 cm^{-1}) have been observed.[47−49] In contrast, absorption and emission maxima for GFP mutants that allow for significant A* emission by disrupting the ESPT chain (e.g., S205V,[50] T203V/S205A,[51] and deGFPs[52]) are all fairly close to 400 and 460 nm, respectively, which could be due to the limited sampling of protein environments and/or the relatively poor ability for protein environments to reorganize as compared to simple solvents.

As discussed in detail in ref 13, the color tuning of the B state is achieved by modulating the driving force between the P and I forms (Figure 4D), while we now see that color tuning of the A state is done through changing the relative energies between the CT and LE forms (Figure 4C). Given the same direction of the difference dipoles $\vec{\mu}_{CT}$ and $\Delta\vec{\mu}_{CT}$ between the corresponding underlying diabatic states for the A and B states, respectively, we can explain why both states show the same qualitative color tuning trends in various protein environments (Figure 4C and D), as noted in section 2.2. However, due to the dissimilar mechanisms between a two- and three-form model for B and A states, while the B state shows the reddest possible absorption maximum ($2V_0$) for electrostatic color tuning, the A state possesses the bluest possible absorption maximum ($\Delta\bar{\nu}_{LE}$) instead (Figure 4A and B). Therefore, as opposed to the common belief generalized from solvent studies, we conclude that the A state's color can be more easily tunable in the protein environment than can the B state, as was illustrated previously by the $\pi-\pi$ stacking series in section 2.2. A correlation plot between the A and B state maxima from protein mutants can also attest this trend (Figure S4), in which the only violations can be found when comparing T203 and T203Y mutants. The T203Y mutation is a frequently invoked example for demonstrating the insensitivity of the A-state color to environmental changes, a common notion derived from solvatochromic studies,[27] because T203Y red-shifts the B state by 1400 cm^{-1} while leaving the A state relatively unaffected (red-shift by 375 cm^{-1}).[53] However, there is an important rotameric difference in T203 when interacting with the neutral and anionic chromophores (Figure

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

1). Specifically, while there is a drastic change in the environment from hydrogen bonding to $\pi-\pi$ stacking for the B-state chromophore when T203 is replaced with tyrosine, the corresponding environmental change for the A-state chromophore is much less significant because the hydrogen bond is no longer present to begin with (as in the T203V mutant). Therefore, the situations are not entirely comparable, and we cannot conclude the lack of color tunability for the A state on the basis of this example. In other words, inferring the intrinsic color tunability between the two protonation states from color shifts is only valid when both states are embedded in the same exact environments (Figure S5). Of course, this is also not possible in solvents as solvent molecules reorganize to accommodate different protonation states.

From our explanations of A and B state color tuning, we can now discern the (de)stabilizing characteristics of specific interactions to link structure to energetics. Specifically, because T203 stabilizes the P form and destabilizes the CT form, while R96 prefers both I and CT forms, we can conclude that hydrogen bonds tend to stabilize anions but destabilize cations;[54] in contrast, $\pi-\pi$ stacking on the phenol(ate) moiety selectively destabilizes the P form over the I form and stabilizes the CT form, which shows a tendency of the π system to interact more favorably with cations. This observation casts doubt on the early assertion that the redder color for yellow fluorescent protein from wild-type GFP through the mutation of T203Y is achieved by the polarizability of the π system,[55,56] because a polarizable electronic system should be able to accommodate both cations and anions via redistribution of its electrons accordingly. Instead, it is the electron-rich nature of the π system that results in the overall stabilization of cations, which rationalizes the prevalence of cation$-\pi$ over anion$-\pi$ interactions in protein structures,[57,58] also partially due to the lack of naturally occurring amino acids bearing electron-deficient aromatic side chains such as F_5F.[59] It is also satisfying to see that we are able to treat these interactions on an equal footing through models of electrostatic color tuning demonstrated in Figure 4A and B.[60−62]

2.4. Photoacidity. The color tuning behaviors and Stark tuning rates of neutral and anionic chromophores have profound implications on the photoacidity of the A state chromophore. According to the Förster cycle,[63] the difference in pK_a between the excited and ground states of the protonated chromophore ($\Delta pK_a = pK_a^* - pK_a$) can be readily estimated from the Stokes shift after ESPT from A state excitation, if we assume the entropy change is the same for ground- and excited-state protonation of the corresponding anionic species (I and I*, Figure 6).

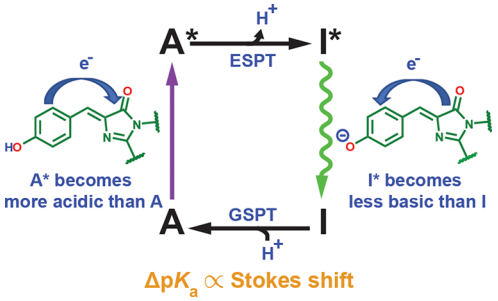


Figure 6. Scheme for proton and electron flow in the Förster cycle of GFP in the A state (Figure 1). Electron redistribution serves as a driving force for ESPT and photoacidity.

Therefore, any color tuning of the deprotonated and protonated states caused by the environment directly determines how the environment modulates ΔpK_a. In other words, both electron redistribution during excitation and emission before and after ESPT, respectively, can be driving forces for promoting excited-state deprotonation.[15] Because we know from section 2.2 that the excitation of both protonation states leads to an electron flow from the phenol(ate) to the imidazolinone moiety, electron redistribution upon excitation of both states works in the same direction to aid the photoacidity (Figure 6). Specifically, inferred from the electron flows, the excited-state protonated chromophore (A*) is a stronger acid than the ground-state counterpart (A), and the excited-state deprotonated chromophore (I*) is a weaker base than the ground-state counterpart (I). Note that we approximate the emission of the I* state as the reverse of its absorption process, because the anionic chromophore is known to exhibit a small Stokes shift.[13] On the basis of the measured magnitudes of Stark tuning rates, the intramolecular electron flow for the protonated state in the redder mutants (e.g., T203OMeY) is nearly the only contributor to the photoacidity due to its much larger Stark tuning rate than the I-state counterpart, while for bluer mutants such as R96M, electron flow for the I state is dominant (Figure 4A and B). Because the protein environment is more capable of red-shifting the color of the protonated than is the deprotonated chromophore (section 2.3), the Stokes shift is decreased and the ΔpK_a becomes less negative for redder mutants. Consequently, to render the protonated GFP chromophore a better photoacid, one should follow the same strategy as blue-shifting the chromophore, by designing hydrogen-bonding groups that interact with the phenol oxygen or attaching electron-withdrawing groups to the phenol moiety, the latter of which echoes Fang and Solntsev's recent work on ESPT of GFP model chromophores.[10] However, the ground-state pK_a itself is not correlated with color when modulated by electrostatics as we previously argued (see sections S7 and S8 in ref 13). Specifically, as the chromophore deprotonation involves a net charge loss, while color tuning only relates to the electronic redistribution across the chromophore with an intact overall charge, the pK_a and the color of the chromophore are governed by electrostatic potentials and projected electric fields (gradients of potentials) exerted by the environment, respectively. These two electrostatic quantities are not necessarily correlated with each other. However, because ΔpK_a entails the difference in energetics of the two deprotonation processes, the net charge changes associated with them are balanced out, and the resulting response to the environment can be understood simply through color tuning, as the Förster cycle indicates. Also note that, in addition to the strong photoacidity, the chromophore has to be somewhat buried and interacts with a terminal proton acceptor through a well-positioned hydrogen-bond network for ESPT to occur in proteins.

3. CONCLUSIONS

Through extensive mutational studies and electronic Stark spectroscopy, we have elucidated the electrostatic color tuning behavior of the protonated GFP chromophore using a simple three-form model, which is surprisingly different from yet not much more sophisticated than the two-form model for the deprotonated chromophore. In particular, after the anionic chromophore is protonated, while the electron flow direction from phenol(ate) to imidazolinone upon excitation is retained, the ease of color tuning in response to environmental

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

electrostatics is drastically changed due to the difference in relative energetics of the diabatic forms. This difference not only allows us to better understand the solvatochromism of the GFP model chromophores, but also offers a strategy for increasing the GFP chromophore's photoacidity through electrostatic modulation. This work showcases the correlation plot between Stark tuning rates and absorption energies as a powerful tool for characterizing excited-state electronic structures and the underlying color tuning mechanisms for dyes, and also the electrostatic role of the environment in which they are embedded.

The three-form model is not only limited to the protonated GFP chromophore but can likely be also applied to donor−acceptor dyes with strongly asymmetric characters and narrow solvatochromic ranges for absorption,[64−68] which would be two contradictory properties if the two-form model were invoked.[69,70] In those cases, the corresponding CT forms are energetically penalized and comparable in energy with the rather unexpected LE form. Because of the moderate electronic couplings between the donor and acceptor moieties, electrons are effectively delocalized in the excited state for such dyes, as opposed to systems with weaker couplings such that thermally activated excited-state electron transfer can occur instead.[71,72] Intriguingly, these dyes also tend to exhibit strong solvatofluorochromism as expected from the model. As another notable biological example, heterodimeric special pairs of bacterial photosynthetic reaction centers, in which one of the bacteriochlorophylls has been replaced with bacteriopheophytin,[44] also obey the same color tuning mechanism (section S5). It could be informative to re-examine the validity of the two-form model with donor−acceptor systems with relatively localized ground-state electronic distributions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c02796.

Detailed experimental procedures, including sample preparation and spectroscopic methods; further technical discussion on related topics, such as elaboration on theoretical models, electrostatic color tuning, and Stark analysis; additional experimental data in figures and tables (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Chi-Yun Lin − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0001-6555-8767; Email: chiyunl@stanford.edu

Steven G. Boxer − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0001-9167-4286; Email: sboxer@stanford.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c02796

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Professor Robert Stanley at Temple University for the Stark spectra fitting software and Professor Pakorn Tony Kanchanawong at the National University of Singapore for his suggestions on the Stark spectroscopy setup. We greatly appreciate Professor Sharon Hammes-Schiffer's illuminating feedback on various related topics. We also thank Matt Romei for reading the manuscript and Tom Carver in Stanford Nano Shared Facilities for depositing nickel on Stark windows. C.-Y.L. was supported by a Kenneth and Nina Tai Stanford Graduate Fellowship and the Taiwanese Ministry of Education. This work was supported, in part, by NIH Grant GM118044 (to S.G.B.) and NSF CCI Phase I: Center for First-Principles Design of Quantum Processes (CHE-1740645).

■ REFERENCES

(1) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.

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

(3) van Thor, J. J. Photoreactions and dynamics of the green fluorescent protein. *Chem. Soc. Rev.* **2009**, *38*, 2935−2950.

(4) Tonge, P. J.; Meech, S. R. Excited state dynamics in the green fluorescent protein. *J. Photochem. Photobiol., A* **2009**, *205*, 1−11.

(5) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **2009**, *462*, 200−204.

(6) Remington, S. J. Green fluorescent protein: a perspective. *Protein Sci.* **2011**, *20*, 1509−1519.

(7) Salna, B.; Benabbas, A.; Sage, J. T.; van Thor, J.; Champion, P. M. Wide-dynamic-range kinetic investigations of deep proton tunnelling in proteins. *Nat. Chem.* **2016**, *8*, 874−880.

(8) Baranov, M. S.; Lukyanov, K. A.; Borissova, A. O.; Shamir, J.; Kosenkov, D.; Slipchenko, L. V.; Tolbert, L. M.; Yampolsky, I. V.; Solntsev, K. M. Conformationally locked chromophores as models of excited-state proton transfer in fluorescent proteins. *J. Am. Chem. Soc.* **2012**, *134*, 6025−6032.

(9) Chen, C.; Liu, W.; Baranov, M. S.; Baleeva, N. S.; Yampolsky, I. V.; Zhu, L.; Wang, Y.; Shamir, A.; Solntsev, K. M.; Fang, C. Unveiling structural motions of a highly fluorescent superphotoacid by locking and fluorinating the GFP chromophore in solution. *J. Phys. Chem. Lett.* **2017**, *8*, 5921−5928.

(10) Chen, C.; Zhu, L.; Baranov, M. S.; Tang, L.; Baleeva, N. S.; Smirnov, A. Y.; Yampolsky, I. V.; Solntsev, K. M.; Fang, C. Photoinduced proton transfer of GFP-inspired fluorescent superphotoacids: principles and design. *J. Phys. Chem. B* **2019**, *123*, 3804−3821.

(11) Heim, R.; Cubitt, A. B.; Tsien, R. Y. Improved green fluorescence. *Nature* **1995**, *373*, 663−664.

(12) Day, R. N.; Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* **2009**, *38*, 2887−2921.

(13) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified model for photophysical and electro-optical properties of green fluorescent proteins. *J. Am. Chem. Soc.* **2019**, *141*, 15250−15265.

(14) Drobizhev, M.; Callis, P. R.; Nifosì, R.; Wicks, G.; Stoltzfus, C.; Barnett, L.; Hughes, T. E.; Sullivan, P.; Rebane, A. Long- and short-range electrostatic fields in GFP mutants: implications for spectral tuning. *Sci. Rep.* **2015**, *5*, 13223.

(15) Silverman, L. N.; Spry, D. B.; Boxer, S. G.; Fayer, M. D. Charge transfer in photoacids observed by Stark spectroscopy. *J. Phys. Chem. A* **2008**, *112*, 10244−10249.

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

(17) Creemers, T. M. H.; Lock, A. J.; Subramaniam, V.; Jovin, T. M.; Völker, S. Three photoconvertible forms of green fluorescent protein identified by spectral hole-burning. *Nat. Struct. Biol.* **1999**, *6*, 557−560.

(18) Craggs, T. D. Green fluorescent protein: structure, folding and chromophore maturation. *Chem. Soc. Rev.* **2009**, *38*, 2865−2875.

(19) Bublitz, G. U.; Boxer, S. G. Stark spectroscopy: applications in chemistry, biology, and materials science. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213−242.
(20) Kent, K. P.; Oltrogge, L. M.; Boxer, S. G. Synthetic control of green fluorescent protein. *J. Am. Chem. Soc.* **2009**, *131*, 15988−15989.
(21) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging proteins can impart unusual resilience. *J. Am. Chem. Soc.* **2007**, *129*, 10110−10112.
(22) Shinobu, A.; Palm, G. J.; Schierbeek, A. J.; Agmon, N. Visualizing proton antenna in a high-resolution green fluorescent protein structure. *J. Am. Chem. Soc.* **2010**, *132*, 11093−11102.
(23) Takaba, K.; Tai, Y.; Eki, H.; Dao, H.-A.; Hanazono, Y.; Hasegawa, K.; Miki, K.; Takeda, K. Subatomic resolution X-ray structures of green fluorescent protein. *IUCrJ* **2019**, *6*, 387−400.
(24) Shibazaki, C.; Shimizu, R.; Kagotani, Y.; Ostermann, A.; Schrader, T. E.; Adachi, M. Direct observation of the protonation states in the mutant green fluorescent protein. *J. Phys. Chem. Lett.* **2020**, *11*, 492−496.
(25) Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Rocha-Rinza, T.; Christiansen, O.; Bravaya, K. B.; Erokhin, A. V.; Bochenkova, A. V.; Solntsev, K. M.; Dong, J.; Kowalik, J.; Tolbert, L. M.; Petersen, M. Å.; Nielsen, M. B. Photoabsorption studies of neutral green fluorescent protein model chromophores *in vacuo*. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9996−10002.
(26) Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science* **2020**, *367*, 76−79.
(27) Dong, J.; Solntsev, K. M.; Tolbert, L. M. Solvatochromism of the green fluorescent protein chromophore and its derivatives. *J. Am. Chem. Soc.* **2006**, *128*, 12038−12039.
(28) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. Bathochromic shift in green fluorescent protein: a puzzle for QM/MM approaches. *J. Chem. Theory Comput.* **2012**, *8*, 112−124.
(29) Das, A. K.; Hasegawa, J.-Y.; Miyahara, T.; Ehara, M.; Nakatsuji, H. Electronic excitations of the green fluorescent protein chromophore in its protonation states: SAC/SAC-CI study. *J. Comput. Chem.* **2003**, *24*, 1421−1431.
(30) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 2306−2311.
(31) Olsen, S. Locally-excited (LE) versus charge-transfer (CT) excited state competition in a series of para-substituted neutral green fluorescent protein (GFP) chromophore models. *J. Phys. Chem. B* **2015**, *119*, 2566−2575.
(32) Salem, L.; Stohrer, W.-D. A double-well potential for olefin isomerization in polar solvents. *J. Chem. Soc., Chem. Commun.* **1975**, 140−142.
(33) Bonačić-Koutecký, V.; Koutecký, J.; Michl, J. Neutral and charged biradicals, zwitterions, funnels in S_1, and proton translocation: their role in photochemistry, photophysics, and vision. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 170−189.
(34) Turro, N. J.; Ramamurthy, V.; Scaiano, J. C. *Modern Molecular Photochemistry of Organic Molecules*, 1st ed.; University Science Books: Sausalito, CA, 2010; pp 341−351.
(35) Laricheva, E. N.; Gozem, S.; Rinaldi, S.; Melaccio, F.; Valentini, A.; Olivucci, M. Origin of fluorescence in 11-*cis* locked bovine rhodopsin. *J. Chem. Theory Comput.* **2012**, *8*, 2559−2563.
(36) Huntress, M. M.; Gozem, S.; Malley, K. R.; Jailaubekov, A. E.; Vasileiou, C.; Vengris, M.; Geiger, J. H.; Borhan, B.; Shapiro, I.; Larsen, D. S.; Olivucci, M. Toward an understanding of the retinal chromophore in rhodopsin mimics. *J. Phys. Chem. B* **2013**, *117*, 10053−10070.
(37) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and simulation of the ultrafast double-bond isomerization of biological chromophores. *Chem. Rev.* **2017**, *117*, 13502−13565.
(38) Manathunga, M.; Yang, X.; Orozco-Gonzalez, Y.; Olivucci, M. Impact of electronic state mixing on the photoisomerization time scale of the retinal chromophore. *J. Phys. Chem. Lett.* **2017**, *8*, 5222−5227.
(39) Manathunga, M.; Yang, X.; Olivucci, M. Electronic state mixing controls the photoreactivity of a rhodopsin with all-*trans* chromophore analogues. *J. Phys. Chem. Lett.* **2018**, *9*, 6350−6355.
(40) Gromov, E. V.; Domratcheva, T. Four resonance structures elucidate double-bond isomerisation of a biological chromophore. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8535−8544.
(41) Bravaya, K. B.; Bochenkova, A. V.; Granovsky, A. A.; Savitsky, A. P.; Nemukhin, A. V. Modeling photoabsorption of the asFP595 chromophore. *J. Phys. Chem. A* **2008**, *112*, 8804−8810.
(42) Topol, I.; Collins, J.; Polyakov, I.; Grigorenko, B.; Nemukhin, A. On photoabsorption of the neutral form of the green fluorescent protein chromophore. *Biophys. Chem.* **2009**, *145*, 1−6.
(43) Georgieva, I.; Aquino, A. J. A.; Trendafilova, N.; Lischka, H. High-level *ab initio* absorption spectra simulations of neutral, anionic, and neutral+ chromophore of green fluorescent protein chromophore models in gas phase and solution. *Photochem. Photobiol.* **2017**, *93*, 1356−1367.
(44) Zhou, H.; Boxer, S. G. Charge resonance effects on electronic absorption line shapes: application to the heterodimer absorption of bacterial photosynthetic reaction centers. *J. Phys. Chem. B* **1997**, *101*, 5759−5766.
(45) Patrizi, B.; Cozza, C.; Pietropaolo, A.; Foggi, P.; Siciliani de Cumis, M. Synergistic approach of ultrafast spectroscopy and molecular simulations in the characterization of intramolecular charge transfer in push-pull molecules. *Molecules* **2020**, *25*, 430.
(46) Chen, C.; Fang, C. Devising efficient red-shifting strategies for bioimaging: a generalizable donor-acceptor fluorophore prototype. *Chem. - Asian J.* **2020**, *15*, 1514−1523.
(47) Chatterjee, T.; Mandal, M.; Mandal, P. K. Solvent H-bond accepting ability induced conformational change and its influence towards fluorescence enhancement and dual fluorescent of hydroxy *meta*-GFP chromophore analogue. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24332−24342.
(48) Ermakova, Y. G.; Sen, T.; Bogdanova, Y. A.; Smirnov, A. Y.; Baleeva, N. S.; Krylov, A. I.; Baranov, M. S. Pyridinium analogues of green fluorescent protein chromophore: fluorogenic dyes with large solvent-dependent Stokes shift. *J. Phys. Chem. Lett.* **2018**, *9*, 1958−1963.
(49) Deng, H.; Yu, C.; Yan, D.; Zhu, X. Dual-self-restricted GFP chromophore analogues with significantly enhanced emission. *J. Phys. Chem. B* **2020**, *124*, 871−880.
(50) Shu, X.; Leiderman, P.; Gepshtein, R.; Smith, N. R.; Kallio, K.; Huppert, D.; Remington, S. J. An alternative excited-state proton transfer pathway in green fluorescent protein variant S205V. *Protein Sci.* **2007**, *16*, 2703−2710.
(51) Wineman-Fisher, V.; Simkovitch, R.; Shomer, S.; Gepshtein, R.; Huppert, D.; Saif, M.; Kallio, K.; Remington, S. J.; Miller, Y. Insight into the structure and the mechanism of the slow proton transfer in the GFP double mutant T203V/S205A. *Phys. Chem. Chem. Phys.* **2014**, *16*, 11196−11208.
(52) Hanson, G. T.; McAnaney, T. B.; Park, E. S.; Rendell, M. E. P.; Yarbrough, D. K.; Chu, S.; Xi, L.; Boxer, S. G.; Montrose, M. H.; Remington, S. J. Green fluorescent protein variants as ratiometric dual emission pH sensors. 1. Structural characterization and preliminary application. *Biochemistry* **2002**, *41*, 15477−15488.
(53) Jung, G.; Wiehler, J.; Zumbusch, A. The photophysics of green fluorescent protein: influence of the key amino acids at positions 65, 203, and 222. *Biophys. J.* **2005**, *88*, 1932−1947.
(54) Zhang, X.; Liu, G.; Ciborowski, S.; Bowen, K. Stabilizing otherwise unstable anions with halogen bonding. *Angew. Chem., Int. Ed.* **2017**, *56*, 9897−9900.
(55) Wachter, R. M.; Elsliger, M.-A.; Kallio, K.; Hanson, G. T.; Remington, S. J. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure* **1998**, *6*, 1267−1277.
(56) Chica, R. A.; Moore, M. M.; Allen, B. D.; Mayo, S. L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 20257−20262.

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

(57) Ma, J. C.; Dougherty, D. A. The cation−π interaction. *Chem. Rev.* **1997**, *97*, 1303−1324.
(58) Lucas, X.; Bauzá, A.; Frontera, A.; Quiñonero, D. A thorough anion−π interaction study in biomolecules: on the importance of cooperativity effects. *Chem. Sci.* **2016**, *7*, 1038−1050.
(59) Smith, M. S.; Lawrence, E. E. K.; Billings, W. M.; Larsen, K. S.; Bécar, N. A.; Price, J. L. An anion-π interaction strongly stabilizes the β-sheet protein WW. *ACS Chem. Biol.* **2017**, *12*, 2535−2537.
(60) Fried, S. D.; Boxer, S. G. Measuring electric fields and noncovalent interactions using the vibrational Stark effect. *Acc. Chem. Res.* **2015**, *48*, 998−1006.
(61) Hunter, C. A.; Sanders, J. K. M. The nature of π−π interactions. *J. Am. Chem. Soc.* **1990**, *112*, 5525−5534.
(62) Mecozzi, S.; West, A. P.; Dougherty, D. A. Cation−π interactions in simple aromatics: electrostatics provide a predictive tool. *J. Am. Chem. Soc.* **1996**, *118*, 2307−2308.
(63) Tolbert, L. M.; Solntsev, K. M. Excited-state proton transfer: from constrained systems to "super" photoacids to superfast proton transfer. *Acc. Chem. Res.* **2002**, *35*, 19−27.
(64) Webber, N. M.; Meech, S. R. Electronic spectroscopy and solvatochromism in the chromophore of GFP and the Y66F mutant. *Photochem. Photobiol. Sci.* **2007**, *6*, 976−981.
(65) Voliani, V.; Bizzarri, R.; Nifosì, R.; Abbruzzetti, S.; Grandi, E.; Viappiani, C.; Beltram, F. *Cis-trans* photoisomerization of fluorescent-protein chromophores. *J. Phys. Chem. B* **2008**, *112*, 10714−10722.
(66) Sun, W.; Zhou, C.; Xu, C.-H.; Zhang, Y.-Q.; Li, Z.-X.; Fang, C.-J.; Sun, L.-D.; Yan, C.-H. Intramolecular charge transfer in 5-methoxy-2-(2-pyridyl)thiazole-derived fluorescent molecules with different acceptor or donor substituents. *J. Phys. Chem. A* **2009**, *113*, 8635−8646.
(67) Greiner, R.; Schlücker, T.; Zgela, D.; Langhals, H. Fluorescent aryl naphthalene dicarboximides with large Stokes shifts and strong solvatochromism controlled by dynamics and molecular geometry. *J. Mater. Chem. C* **2016**, *4*, 11244−11252.
(68) Zhang, Y.; Liang, C.; Jiang, S. A solvatochromic cyanostilbene derivative as an intensity and wavelength-based fluorescent sensor for water in organic solvents. *New J. Chem.* **2017**, *41*, 8644−8649.
(69) Marder, S. R.; Gorman, C. B.; Meyers, F.; Perry, J. W.; Bourhill, G.; Brédas, J.-L.; Pierce, B. M. A unified description of linear and nonlinear polarization in organic polymethine dyes. *Science* **1994**, *265*, 632−635.
(70) Bublitz, G. U.; Ortiz, R.; Marder, S. R.; Boxer, S. G. Stark spectroscopy of donor/acceptor substituted polyenes. *J. Am. Chem. Soc.* **1997**, *119*, 3365−3376.
(71) Treynor, T. P.; Yoshina-Ishii, C.; Boxer, S. G. Probing excited-state electron transfer by resonance Stark spectroscopy: 4. Mutants near B_L in photosynthetic reaction centers perturb multiple factors that affect $B_L^* \rightarrow B_L^+H_L^-$. *J. Phys. Chem. B* **2004**, *108*, 13523−13535.
(72) Chen, K.-Y.; Hsieh, C.-C.; Cheng, Y.-M.; Lai, C.-H.; Chou, P.-T.; Chow, T. J. Tuning excited-state electron transfer from an adiabatic to nonadiabatic type in donor−bridge−acceptor systems and the associated energy-transfer process. *J. Phys. Chem. A* **2006**, *110*, 12136−12144.

https://dx.doi.org/10.1021/jacs.0c02796
J. Am. Chem. Soc. 2020, 142, 11032−11041

Supporting Information for

"Mechanism of Color and Photoacidity Tuning for the Protonated Green Fluorescent Protein Chromophore"

Chi-Yun Lin[1]* and Steven G. Boxer[1]*

[1]Department of Chemistry, Stanford University, Stanford, CA 94305, USA.

*Correspondence to: chiyunl@stanford.edu; sboxer@stanford.edu

Table of Contents

S1 Sample Preparation

Plasmid Construction

The logic of GFP plasmid design followed our previous works on Superfolder GFPs [1][2]. Point mutations were made using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's protocol. The residue numbering scheme follows GFPs without circular permutation. Supercharged GFP -30 gene in pET-29 was generously provided by the David Liu Lab at Harvard University [3] and used without further modification.

GFP Constructs in This Study

We adopted the nomenclature devised for split GFP circular permutants in our previous works [4]. Labels describe elements (separated by colons) of GFP progressing from the N terminus to the C terminus when read from left to right. Specific β-strands in the GFP β-barrel are denoted sX, where X is the number of the strand of interest, while the internal helix is denoted ih. Loop refers to a sacrificial loop with proteolytic cleavage sites. GFP refers to the remainder of the protein. A strike through an element indicates that the element has been removed. Synthetic elements are underlined. A dot is used to indicate a noncovalent interaction. For example, s10(203F) · ~~s10:loop~~:GFP denotes a synthetic β-s10 carrying the mutation T203F noncovalently bound to circularly permuted GFP with its original N-terminal s10 and loop removed.

Table S1. GFP constructs in this study, forming a subset of those characterized in our previous work [2]. The following entries were colored based on their parent circular permutants. The parent proteins for the colors orange, pink, and green are s10:loop:GFP, ih:GFP, and ih:loop:GFP, respectively. Red letters denote non-wild-type amino acids, and superscript "mat" indicates an internal helix with a matured chromophore. To facilitate readability, the mutation carried by the synthetic strand is enclosed by parentheses rather than superscripted as in our previous publications.

GFP Constructs	ih		s4	s7	s10	s11
	65	66	96	148	203	222
s10:loop:GFP	S	Y	R	H	T	E
s10:loop:GFP T203V	S	Y	R	H	V	E
s10:loop:GFP T203Y	S	Y	R	H	Y	E

ih:GFP T203(3-OMeY)	S	Y	R	H	3-OCH_3Y	E
s10:loop:GFP S65T	T	Y	R	H	T	E
ih:GFP S65T	T	Y	R	H	T	E
supercharged -30	T	Y	R	H	T	E
ih^{mat}(65T) · ~~ih:loop~~:GFP R96M	T	Y	M	H	T	E
s10:loop:GFP S65T T203V	T	Y	R	H	V	E
s10:loop:GFP S65T T203H	T	Y	R	H	H	E
s10(203F) · ~~s10:loop~~:GFP S65T	T	Y	R	H	F	E
s10(203(4-F_1F)) · ~~s10:loop~~:GFP S65T	T	Y	R	H	4-F_1F	E
s10(203(F_5F)) · ~~s10:loop~~:GFP S65T	T	Y	R	H	F_5F	E
s10(203(4-NH_2F)) · ~~s10:loop~~:GFP S65T	T	Y	R	H	4-NH_2F	E
ih:GFP S65T T203(3-OMeY)	T	Y	R	H	3-OCH_3Y	E
s10:loop:GFP S65T T203Y	T	Y	R	H	Y	E

DNA and Amino Acid Sequences

The sequences have been described in our previous work [2] in detail.

Synthetic Peptide Design

Peptides were designed to match native s10 of s10:loop:GFP and were synthesized by Elim Biopharmaceuticals.

s10(203F): LPDNHYLS**F**QTVLSKDPNE

s10(203(4-F_1F)): LPDNHYLS **(4-F_1F)** QTVLSKDPNE

s10(203(F_5F)): LPDNHYLS **(F_5F)** QTVLSKDPNE

s10(203(4-NH_2F)): LPDNHYLS **(4-NH_2F)** QTVLSKDPNE

Semi-synthetic Method for Split GFPs

The protocol has been described in our previous work [2] in detail, including the subsequent purification and characterization.

Sample Preparation for 77 K Absorption and Electronic Stark Spectroscopy

Glass forming solvents, such as ethanol or a 1:1 mixture of glycerol and aqueous buffer, are required for low-temperature electronic Stark spectroscopy experiments. The concentrated samples were mixed with an equal volume of glycerol (Fisher, CAS 56-81-

5) right before Stark measurements. The final sample concentrations for Stark spectroscopy were checked with a NanoDrop spectrometer (ND-1000 Spectrometer; NanoDrop) to ensure a maximum absorbance of 0.2 – 0.9 for a 25 μm path length, the optimum OD for good signal-to-noise ratio in low temperature absorption.

S2 Spectroscopic Methods

UV–Vis Absorption Measurements

UV–Vis absorption spectra at room temperature were all measured with a PerkinElmer Lambda 25 UV–Vis spectrometer and a 1 mL quartz cuvette. Data acquisition was performed every 1.0 nm at a maximum scan rate of 480 nm/min.

Low-Temperature (77 K) Absorption Measurements and Electronic Stark Spectroscopy

The detailed method has been reviewed in our previous works [2][5]. The cell consisted of a pair of 12.7 mm diameter by 1 mm thick fused silica windows (FOCtek Photonics, Inc.) deposited with 45 Å of nickel on the surfaces facing the sample. The windows were separated by a pair of 27-micron thick Teflon spacers and held in place with a metal clamp and four adjustable screws. The interference fringes were optimized under a fluorescent lamp, and the path length was determined by the undulations in UV –Vis absorption from 500 – 1100 nm. The path length was then used to calculate the electric field strength applied during the measurement knowing the applied voltage. The Stark cell was mounted onto a home-built rod with electrical wires and alligator clips attached to the nickel electrodes. The whole apparatus was insulated with electrical tape, and a sample (at most 10 µL) was loaded into the cell by capillary uptake. The whole rod was then rapidly plunged into an immersion cryostat [6] pre-filled with liquid nitrogen to allow the sample to form a transparent glass upon flash freezing. Protein samples with glycerol were centrifuged at 17000 rcf for at least 40 min prior to sample loading.

The custom-built spectrometer could be switched between Stark spectroscopy and absorption modes with the latter dual-beamed. For Stark spectroscopy, the sinusoidal high voltage signal was generated from the sample channel of a lock-in amplifier (SR830; Stanford Research) with a frequency of 203 Hz, amplified 1000-fold via a high-voltage power supply (TREK 10/10; TREK), and the voltage was applied through the rod onto the sample. The root-mean-square voltage (V_{rms}) applied before dielectric breakdown can range from 0.6 – 3.0 kV, which amounts to a peak external field strength F_{ext} of 0.3 – 1.6

MV/cm given the sample thickness. The *X* and *Y* components of Stark signal Δ*I* were detected at the second harmonic of the applied field. The direct output voltage *I* was also simultaneously recorded. The Stark spectra were then obtained from the ratio: $\Delta A = \frac{2\sqrt{2}}{\ln 10}\frac{\Delta I}{I}$ as a function of the scanning wavelengths [5]. A wavelength scan rate of 0.3 nm/s and a time constant of 300 ms were chosen. The polarized probe light was set to be horizontal, and depolarization along the beam path was carefully checked. χ angles between the applied electric field and the polarization of the probe light at 90°, 70°, and 50° were sampled at each applied field strength with an increment of 0.3 kV in V_{rms} to ensure a complete data set.

For absorption spectroscopy, the sample channel was reconfigured by replacing the polarizer with another beam chopper, and the reference channel was employed. The magnitudes of output signals were detected at the first harmonic of the chopper modulation frequency (3029 Hz). The scanning rate and time constant were set to match those of the Stark measurements. The blank sample was prepared by carefully blowing the Stark sample out of the cell with air and then loading the cell with a 1:1 buffer and glycerol mixture (or ethanol). The absorbance *A* was determined at normal incidence with an absolute uncertainty around ±0.01. The final absorbance was obtained by averaging over three to four scans for each sample. LabView programs were used to facilitate data collection in both modes. Undulation can be occasionally seen in the baseline at the red-edge of the absorption spectra due to light interference between two windows of the sample cell.

Stark Spectroscopy Data Analysis

All Stark spectra Δ*A* are shown with their corresponding absorbance *A* normalized. Δ*A* are also scaled to 1 MV/cm with χ = 90° according to their proportionality to $(F_{ext})^2$, where χ is the angle between probe light polarization and field direction. The Stark spectra (as functions of wavenumbers $\bar{\nu}$) were analyzed as linear combinations of wavenumber-weighted zeroth, first, and second derivatives of the absorbance spectra with coefficients A_χ, B_χ, and C_χ as functions of χ, respectively, to extract the apparent Stark tuning rates

$\Delta\mu_{app}$ (= $|\Delta\vec{\mu}_{app}|$) and the measured angles ζ between difference dipoles and transition dipoles [5]:

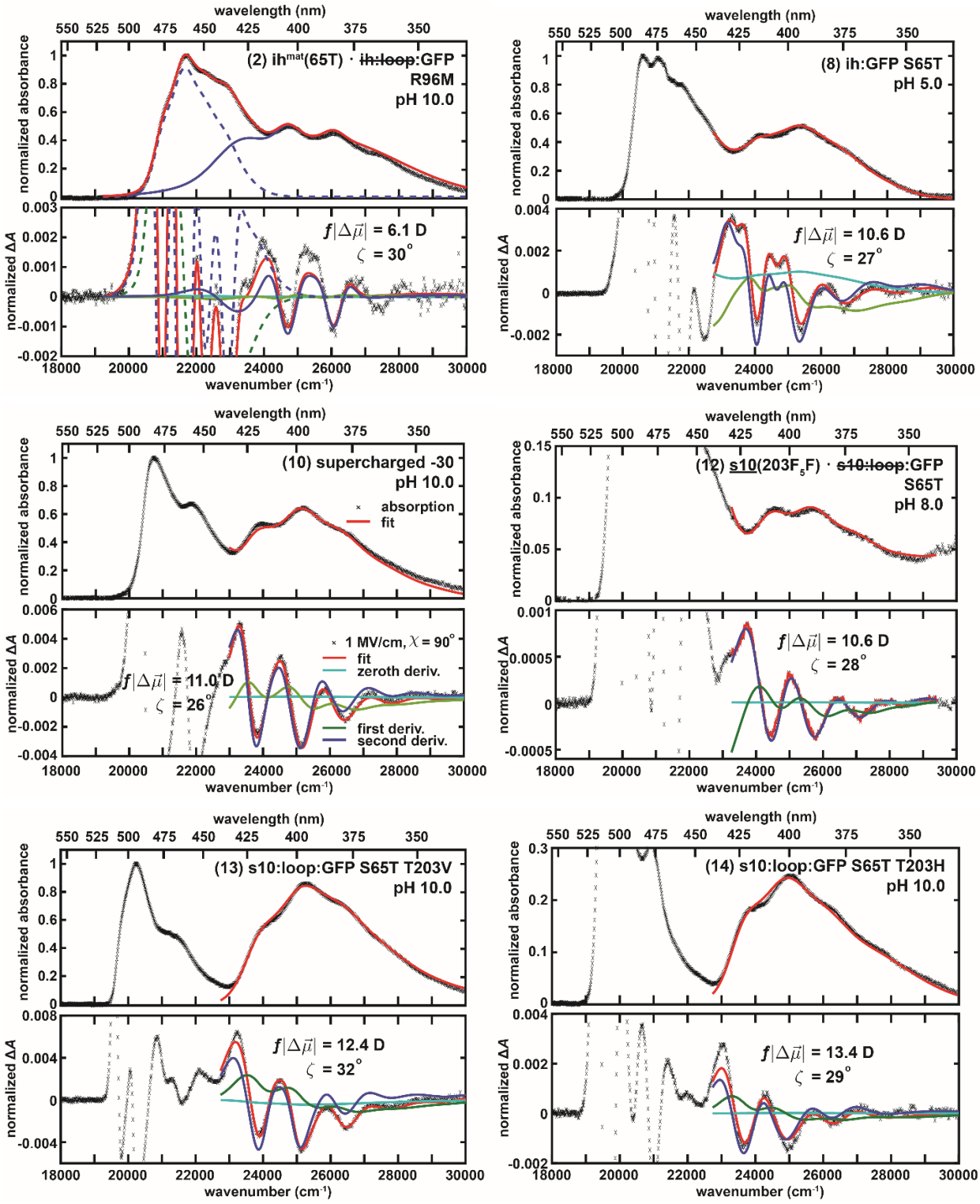
$$\begin{aligned}\Delta A(\bar{\nu}, F_{ext}) &= A(\bar{\nu}, F_{ext}) - A(\bar{\nu}, F_{ext} = 0) \\ &= (F_{ext})^2 \left[A_\chi A(\bar{\nu}) + \frac{B_\chi}{15hc}\bar{\nu}\frac{d}{d\bar{\nu}}\left(\frac{A(\bar{\nu})}{\bar{\nu}}\right) + \frac{C_\chi}{30h^2c^2}\bar{\nu}\frac{d^2}{d\bar{\nu}^2}\left(\frac{A(\bar{\nu})}{\bar{\nu}}\right)\right] \end{aligned} \quad \text{(S1)}$$

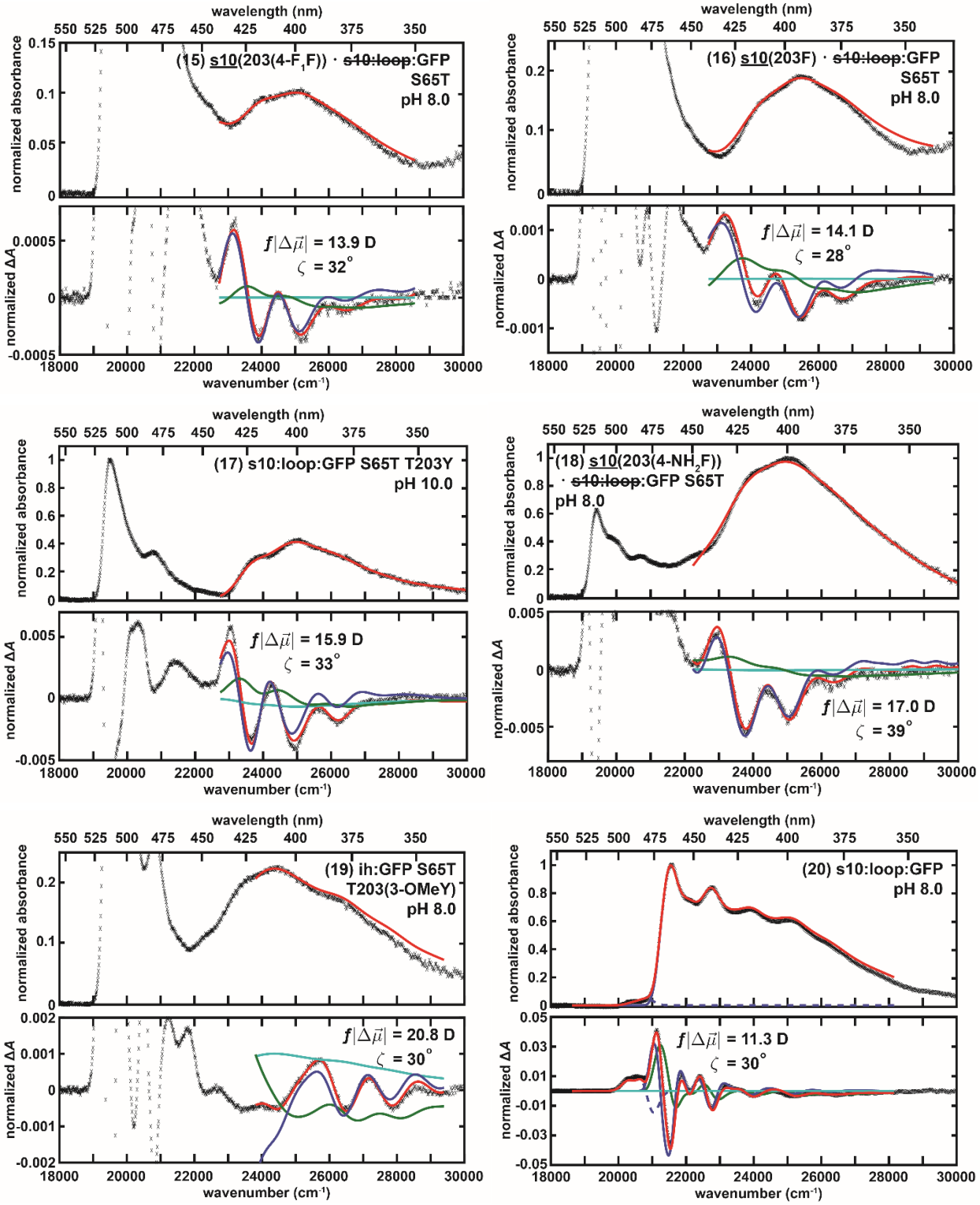
and

$$C_\chi = \left(\Delta\mu_{app}\right)^2[5 + (3\cos^2\chi - 1)(3\cos^2\zeta - 1)] \quad \text{(S2)}$$

where F_{ext} is the strength of the externally applied field through the parallel-plate capacitor. The magnitude of a vector quantity is denoted by dropping the vector notation. The data was processed by the MATLAB code kindly provided by Professor Robert Stanley at Temple University [7]. Simultaneous fitting of Δ*A* and *A* at χ = 90°, 70°, and 50° were performed with a minimal number of Gaussian components and their analytical derivatives to model the vibronic progression and effectively smooth the absorbance spectra. No real physical meaning is associated with the individual peak positions of these fit Gaussians, and any attempt to do so should be treated with great caution. One set of electro-optical parameters ($\Delta\mu_{app}$, ζ, A_χ, and B_χ) was first assigned to recapitulate the transition with the dominant Stark effect. More bands were employed only if the result from the one-band fit was unsatisfactory (Section S5). Due to the dominant contribution from $\Delta\mu_{app}$, no attempt was done to isolate the difference polarizabilities $\Delta\alpha$ from B_χ. The uncertainties in $\Delta\mu_{app}$ from both fitting and duplicates amounted to ±5%, while those in ζ were ±5°, unless the bands were too small (*A* < 0.1) to be properly analyzed. Throughout this study, $\Delta\mu_{app}$ was treated as the product of the true difference dipole moment of the chromophore $\Delta\mu$ and the local field factor f, with the latter assumed to be a constant scalar across different environments. The necessity of including f reflects our lack of certainty over the magnitude of the local field sensed by the chromophore [5] (see also Section S6 in [2]).

S3 Stark Spectra and Fitting of GFP Mutants and HBDI




wavelength (nm)
550 525 500 475 450 425 400 375 350
normalized absorbance
(15) s10(203(4-F1F)) · s10:loop:GFP S65T pH 8.0
(16) s10(203F) · s10:loop:GFP S65T pH 8.0
(17) s10:loop:GFP S65T T203Y pH 10.0
(18) s10(203(4-NH2F)) · s10:loop:GFP S65T pH 8.0
(19) ih:GFP S65T T203(3-OMeY) pH 8.0
(20) s10:loop:GFP pH 8.0
normalized ΔA
f|Δμ| = 13.9 D
ζ = 32°
f|Δμ| = 14.1 D
ζ = 28°
f|Δμ| = 15.9 D
ζ = 33°
f|Δμ| = 17.0 D
ζ = 39°
f|Δμ| = 20.8 D
ζ = 30°
f|Δμ| = 11.3 D
ζ = 30°
wavenumber (cm-1)
18000 20000 22000 24000 26000 28000 30000

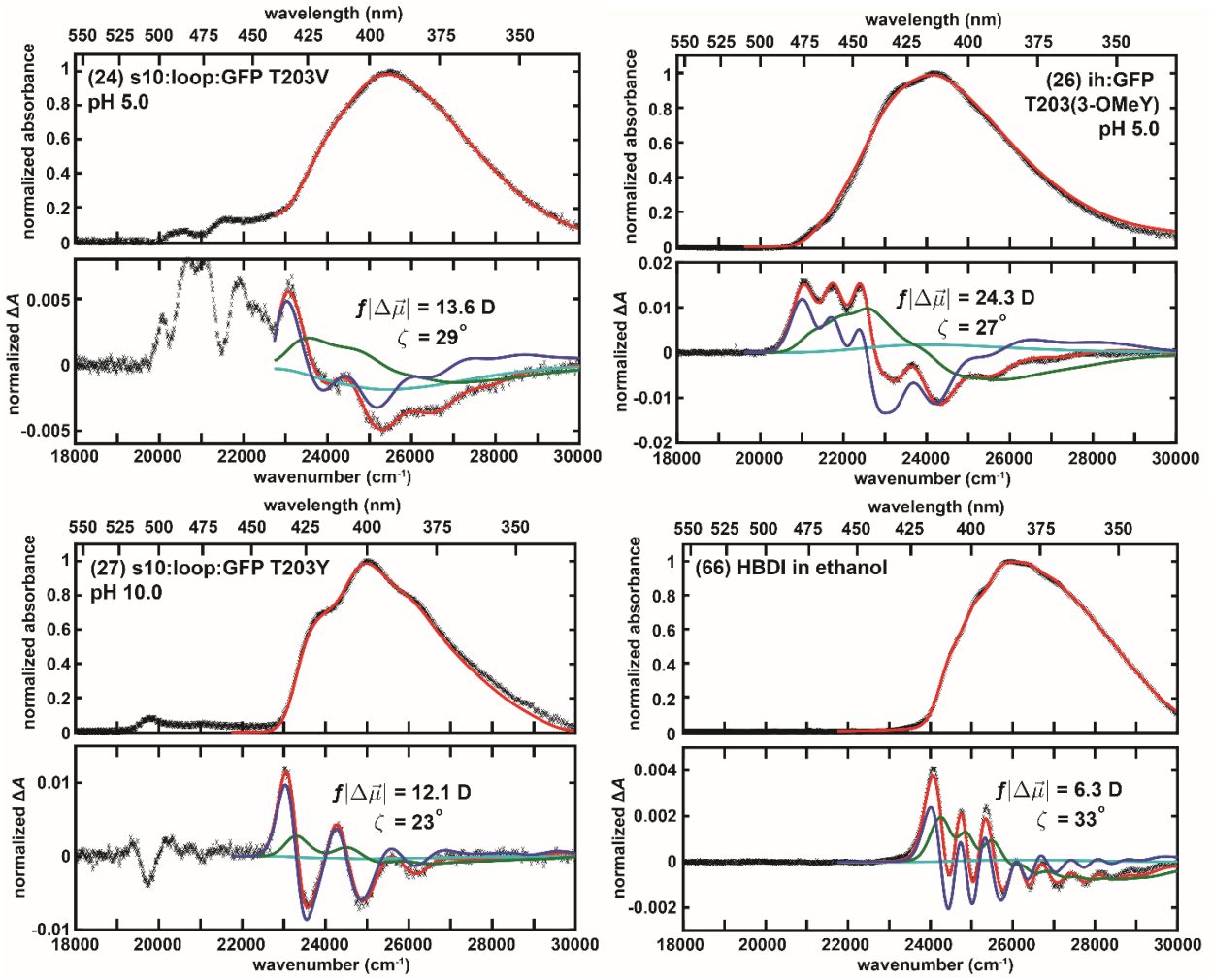


Figure S1. The classical sum-of-derivative analysis for 77 K UV–Vis absorption and Stark spectra, which are dominated by second-derivative lineshapes (Figure S2). The corresponding B-state analysis can be found in Figure S10 of [2]. The panels are numbered and listed according to Table S4. The absorption spectra are normalized to 1 at the maximum absorbance in 18000 – 30000 cm^{-1} and are magnified if the corresponding normalized absorbance for A state is less than 0.3. The Stark spectra are measured at χ = 90° and scaled to 1 MV/cm to facilitate comparison. The color scheme of fit lines and data, as shown in panel 10, is consistent throughout the figure. Solid lines represent the band of interest (the neutral state), for which the measured Stark tuning rate (±5%) and the ζ angle (±5°) is noted, while dashed lines in panel 2 represent other absorption bands that require simultaneous fitting to extract out electro-optic parameters from the higher energy band. In most cases, in which either only one dominant band or sufficient spectral separation between bands is observed, assigning one set of electro-optic parameters with occasionally a limited spectral range is preferred, even when two bands with distinct origins (such as A and B states) share similar Stark parameters (panel 20). An extra set of electro-optic parameters is only considered when the fit demands such a scenario due to significant overlap of two bands with different Stark tuning rates (panel 2). Due to less sharp vibronic features of A band compared to B band, the A-state

Stark signals are much less significant than the B-state counterparts and tend to be buried when both bands overlap.

S4 The Three-Form Coupling Model for Protonated GFP Chromophore

In this section, we present a detailed derivation for the relation between Stark tuning rate and absorption maximum for the neutral GFP chromophore from the three-form model (Figure 4C), which is based on Olsen's multi-configurational calculation [8]. Even though the model contains three forms, it is effectively a two-form model since the GS form is decoupled from the other two forms by assumption due to the relatively large energy gap. By setting the energy difference between the LE and GS forms to be $\Delta\bar{\nu}_{LE}$ ($\equiv \bar{\nu}_{LE} - \bar{\nu}_{GS}$) and that between the CT and LE forms to be $\Delta\bar{\nu}$ ($\equiv \bar{\nu}_{CT} - \bar{\nu}_{LE}$) (see Figure 4C in main text), we can write down the potential energy matrix to describe the coupling between the LE and CT forms:

$$V = \begin{pmatrix} \Delta\bar{\nu}_{LE} & {V_0}' \\ {V_0}' & \Delta\bar{\nu}_{LE} + \Delta\bar{\nu} \end{pmatrix} \quad \text{(S3)}$$

where V_0' is the associated electronic coupling and the diagonal element are the energies for the LE and CT forms from the GS form. We do not include vibrational degrees of freedom in contrast to our previous B-state color tuning model [2]; this is justified below. The relative energies for the resulting two excited adiabatic states with respect to the GS form can be solved by diagonalizing Equation S3 and are:

$$\Delta\bar{\nu}_{1,2} = \Delta\bar{\nu}_{LE} + \frac{\Delta\bar{\nu}}{2} \pm \frac{\sqrt{\Delta\bar{\nu}^2 + 4{V_0}'^2}}{2} \quad \text{(S4)}$$

the lower of which is the absorption energy from the S_0 to S_1 state:

$$\bar{\nu}_{abs} = \Delta\bar{\nu}_{LE} + \frac{\Delta\bar{\nu}}{2} - \frac{\sqrt{\Delta\bar{\nu}^2 + 4{V_0}'^2}}{2} \quad \text{(S5)}$$

Therefore, when $\Delta\bar{\nu}$ is much larger than $2V_0$', no appreciable mixing between the two states is expected, and $\bar{\nu}_{abs}$ approaches $\Delta\bar{\nu}_{LE}$ as the bluest possible absorption for the neutral GFP chromophore. As $\Delta\bar{\nu}$ becomes smaller or more negative, the coupling lowers the LE form's energy and red shifts the absorption (Figure 5). It does not matter whether we use the 0–0 electronic excitation energy or the absorption maximum (which is technically 0–1 excitation energy) as $\bar{\nu}_{abs}$, since it should only differ by a constant frequency (Figure 3D), which is absorbed by $\Delta\bar{\nu}_{LE}$ and justifies why we neglect the energy offset. After mixing between the CT and LE forms, the S_1 state inherits some dipolar character from the CT form, which carries a dipole moment of μ_{CT}. Since the GS form has

a much smaller dipole moment than the CT form, the Stark tuning rate $\Delta\mu$ associated with the absorption becomes:

$$\Delta\mu = \left(\frac{1}{2} - \frac{\Delta\bar{\nu}}{2\sqrt{\Delta\bar{\nu}^2 + 4{V_0}'^2}}\right)\mu_{CT} \tag{S6}$$

(cf. Equation 2 in the main text). When $\Delta\bar{\nu}$ is much larger than $2V_0'$, the Stark tuning rate becomes zero because the S_1 state exists purely as the neutral LE form (Figure 5). The fact that zero Stark tuning rate corresponds to the bluest possible absorption for the neutral chromophore is not a mere coincidence, as discussed below. Since $\Delta\bar{\nu}$ is not an observable and assumed to be the only quantity that can be modulated by the chromophore's environment through electrostatic interactions, as developed in detail for the driving force in the Marcus–Hush model for the B state [2], we can combine Equations S5 and S6 by eliminating $\Delta\bar{\nu}$ to obtain the correlation between $\bar{\nu}_{abs}$ and $\Delta\mu$:

$$\boxed{\bar{\nu}_{abs} = \Delta\bar{\nu}_{LE} - {V_0}' \frac{2\frac{\Delta\mu}{\mu_{CT}}}{\sqrt{1-\left(2\frac{\Delta\mu}{\mu_{CT}}-1\right)^2}}} \tag{S7}$$

Note that we include the local field factor f in the corresponding equation in the main text (Equation 1) to emphasize the fact that all experimentally determined dipole moments are associated with f due to environmental polarization in response to the externally applied field.

In addition to allowing us to extract parameters for a specific model from the correlation plot of absorption maxima and Stark tuning rates for various mutants, we advocate that the correlation plot is a useful strategy for understanding a chromophore's electrostatic color tuning behavior even if the underlying model is unknown. The Stark tuning rate is the linear response of the chromophore's color to the effective electric field $\vec{F}$ experienced by the chromophore:

$$\Delta\mu = \left|\frac{\partial\bar{\nu}_{abs}}{\partial\vec{F}}\right| \tag{S8}$$

which is a *derivative* of $\bar{\nu}_{abs}$ with respect to $\vec{F}$. By plotting $\bar{\nu}_{abs}$ against $\Delta\mu$ for one-dimensional systems, we can in principle capture the behavior of the function $\bar{\nu}_{abs}(\vec{F})$ without knowing the function itself explicitly. This is in fact a widely used strategy in different contexts. For instance, in classical Hamiltonian mechanics, it is useful to plot the

velocity (which is the time *derivative* of the position) of a particle against its position to illustrate the phase space trajectory and analyze the dynamical information without directly solving the equation of motion [9]. Similarly, when encountering a nonlinear ordinary differential equation (ODE) of a function *y*(*t*) that is not analytically solvable, a plot with $\frac{dy}{dt}$ against *y* itself, known as the phase portrait, is an invaluable tool to understand the dynamical behavior of the ODE, especially in terms of fixed points, flows, and limit cycles [10]. In the case of electrostatic color tuning, *y* is $\bar{\nu}_{abs}$, while the external field plays the role of time. Therefore, the correlation curve between $\bar{\nu}_{abs}$ and $\Delta\mu$ can serve as a calibration curve for electrostatic color tuning for a specific chromophore.

Furthermore, fixed points, defined as *y* at which $\frac{dy}{dt} = 0$, correspond to the extremal values of *y*(*t*). Analogously, even if the function $\bar{\nu}_{abs}(\vec{F})$ that encodes all information of electrostatic color tuning of a given chromophore is unknown, we can still figure out where the reddest or bluest possible absorption maximum is by knowing where $\Delta\mu = 0$ occurs, which is in fact the fixed point(s). Whether it is the reddest or bluest can be readily determined from the correlation plot itself: if $\Delta\mu$ is a decreasing or increasing function of $\bar{\nu}_{abs}$, the fixed point corresponds to the bluest or reddest possible absorption maximum as for the neutral and anionic GFP chromophore [2], respectively.

The aforementioned analysis is only strictly applicable to one-dimensional systems (and two-form systems from the perspective of the minimal Hilbert space for diabatic states) since the direction of $\Delta\vec{\mu}$ stays the same (or at most 180° flipped) across mutants. The direct consequence of being one-dimensional is that the correlation plots are monotonic and environmental effects from any combinations of mutations can be completely captured by the relative energy between the two underlying forms (i.e. driving forces $\Delta\bar{\nu}$). Therefore, the *magnitude* of $\Delta\vec{\mu}$ from each mutant is sufficient to encompass all electro-optic properties and can be readily measured through electronic Stark spectroscopy. The correlation plot for red fluorescent proteins, on the other hand, fails to show the simple monotonic trend [11], suggesting the necessity of incorporating at least one more coupled form likely due to the additional acylimine tail from the GFP chromophore [12][13]. Since at least three energetically close forms with noncolinear charge distributions are involved, the *direction* of $\Delta\vec{\mu}$ likely changes across mutants and

is sensitive to the direction of the electric field exerted by the protein environment. In other words, the effects of mutations can no longer be encoded by only one driving force, so presumably one has to determine both *x* and *y* components of $\Delta\vec{\mu}$ for each mutant (assuming the *z* axis is normal to the chromophore plane) in order to conduct the same analysis. That is to say, a correlation plot with data points of ($\Delta\mu_x$, $\Delta\mu_y$, $\bar{\nu}_{abs}$) is required to evaluate multiple driving forces and electronic couplings between the forms [14] and understand the electrostatic color tuning behavior for non-one-dimensional systems, such as porphyrins, chlorins, and bacteriochlorins [15][16].

S5 Analogy with Special Pair in Bacterial Photosynthetic Reaction Center

As a side note of interest in the spectroscopy of photobiological systems, it is intriguing to draw an analogy between the GFP chromophore and special pair P (i.e. bacteriochlorophyll dimer $BChl_2$) in the *Rhodobacter sphaeroides* photosynthetic reaction center with regards to color tuning and charge transfer. In particular, the B-state GFP chromophore resembles the oxidized P (P^+) and both chromophores can both be treated as mixed-valence systems since the electron and hole are strongly delocalized within the chromophores, respectively [17]. On the other hand, owing to the large energetic asymmetry between the two interacting moieties within the chromophore, the A state is more similar to a heterodimer D, which is formed by selectively replacing one of the BChls with a bacteriopheophytin (BPhe): both have a CT form ($phenol^+imidazolinone^-$ or $BChl^+BPhe^-$) close in energy to an LE-type form (or an exciton state) that can be excited from the ground state [18]. An obvious difference, however, between the special pair (P^+ and D) and the GFP chromophore (A and B states) is that the electronic coupling between the two rings of the latter is an order of magnitude larger than that for the former due to the direct π conjugation. For the charged states P^+ and B, V_0 is approximately 1000 cm^{-1} [17] and 9500 cm^{-1} [2], respectively; for the neutral states D and A, V_0' is approximately 500 cm^{-1} [18] and 5000 cm^{-1}, respectively. Following this analogy, it might be possible to understand why the redder A bands tend to be broader through the Fano theory [19], in which a broadened absorption band is caused by a decrease in excited lifetime of the LE form by virtue of coupling to the CT form (homogeneous broadening) [18]. It could also be explained by more significant inhomogeneous broadening from the environmental electric field distribution experienced by the chromophore with a larger Stark tuning rate, as explicated by Drobizhev *et al.* [20] and our previous work [2].

S6 Supplementary Figures

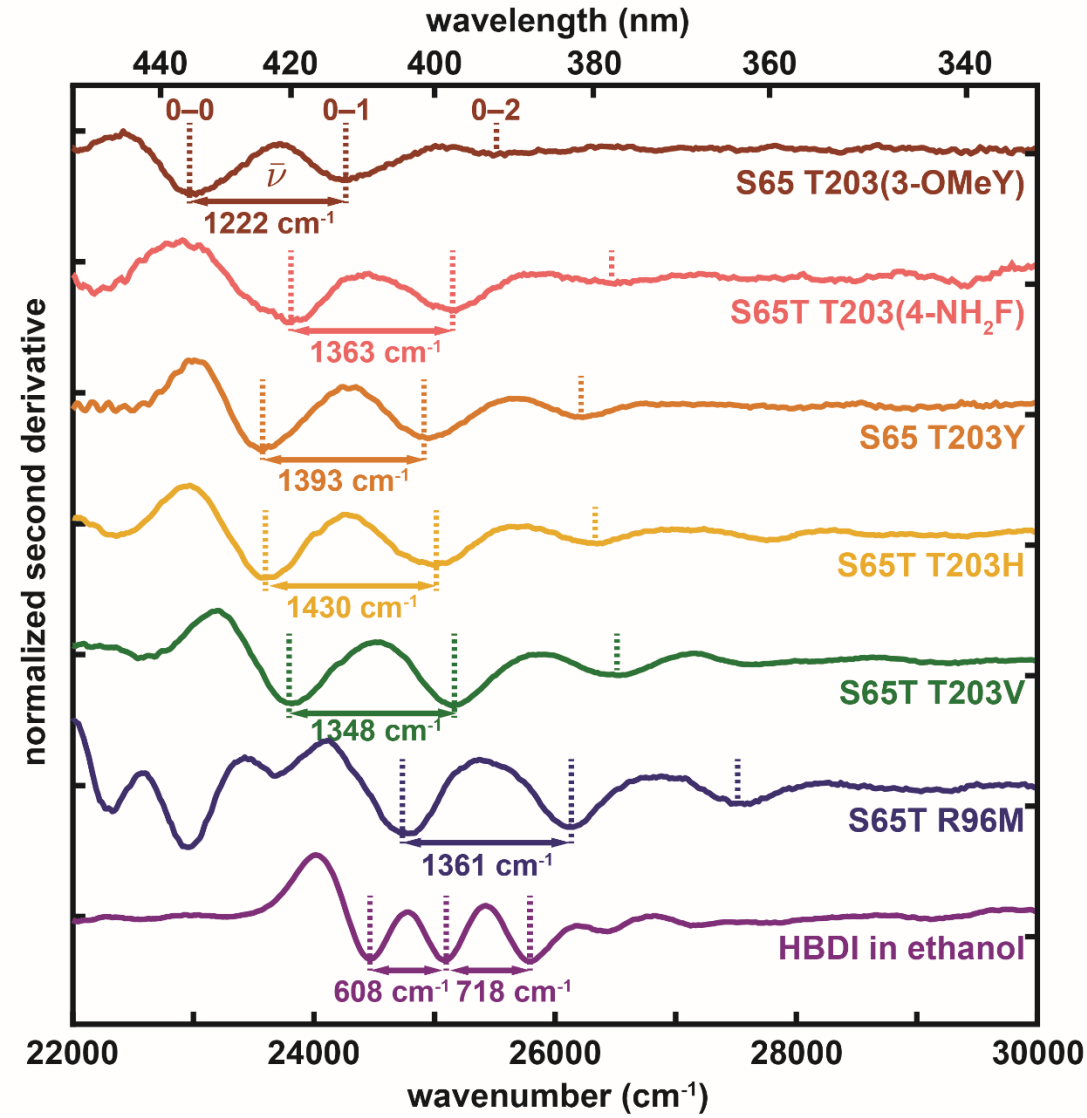


Figure S2. Method for estimating the difference between 0–0 and 0–1 transitions using the second derivative of 77 K absorption spectra (Figures 3A and 3B) of protonated GFP mutants and HBDI. 0–0 and 0–1 energies are assigned from the negative peak positions; this difference corresponds to the BLA and other possible normal modes with similar frequencies for GFP mutants (Figure 3D). The color coding follows that of Figures 3A and 3B. For HBDI in ethanol, the frequency spacings are about half of those from GFP mutants, suggesting that the 0–2 feature assigned in the figure could also correspond to a 0–1 transition as in protein mutants, while the 0–1 feature results from another vibrational mode that is more vibronically coupled in ethanol than in the GFP environment. Interestingly, this additional feature is not seen for the anionic HBDI in ethanol (Figure S24 in [2]). The redder features from S65T R96M GFP correspond to the B and I bands [2], which are also observed in the absorption spectrum (Figure 3B).

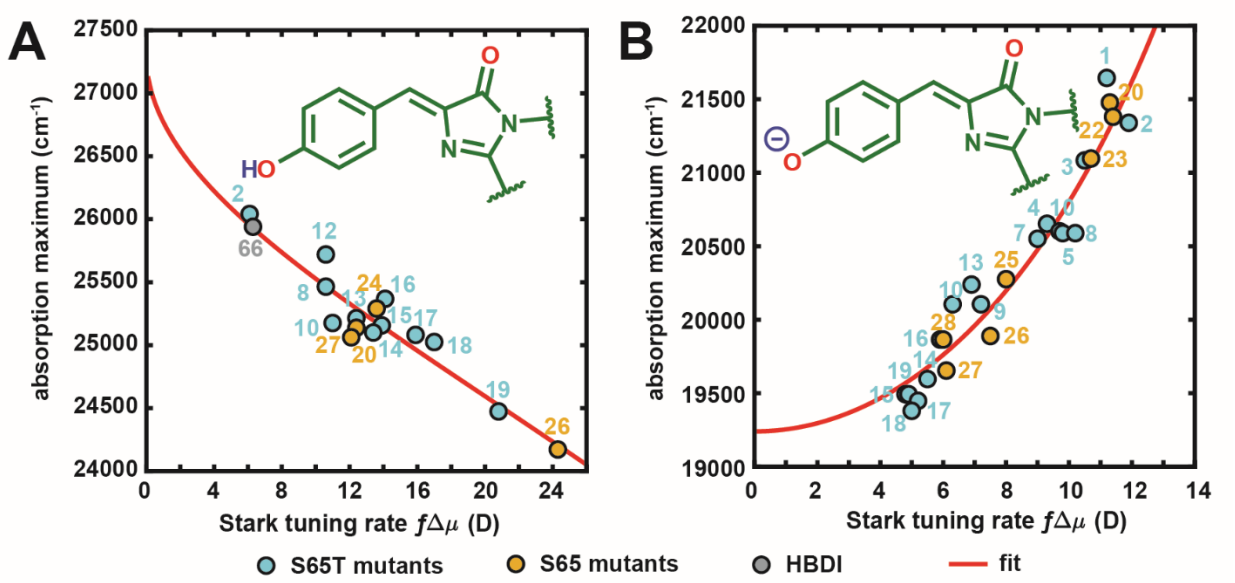


Figure S3. Correlation between the absorption maximum and Stark tuning rate for GFP mutants and HBDI in the (A) A and (B) B states at 77 K. This figure is reproduced from Figure 4 to include numerical labels defined in Table S4.

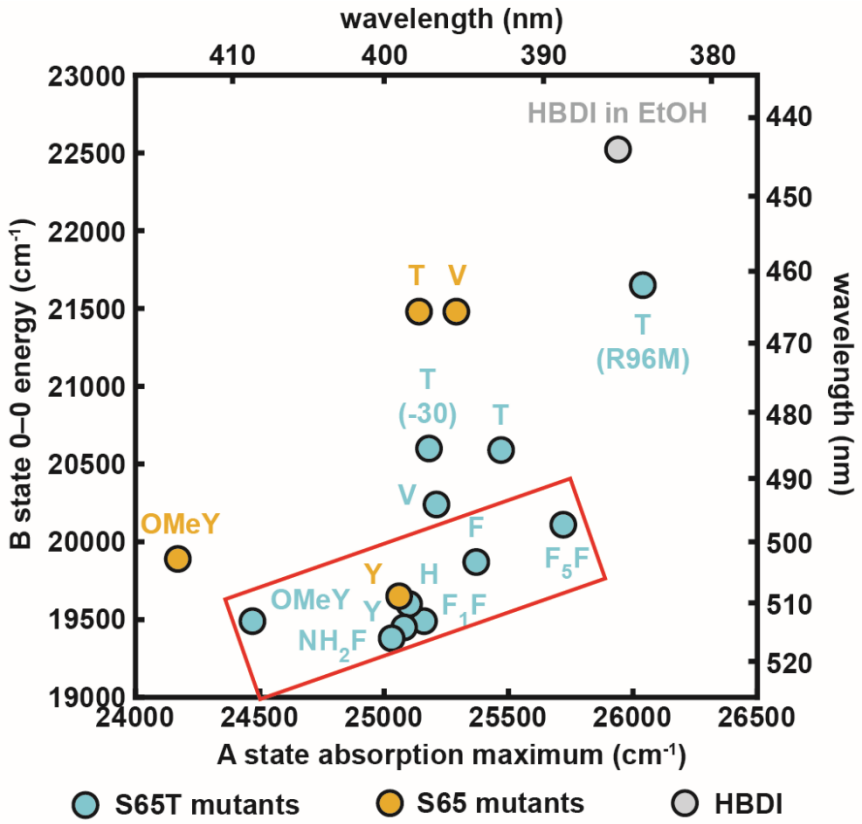


Figure S4. Correlation plot between the B state 0–0 energy (which is approximately the absorption maximum [2]) and A state absorption maximum for GFP mutants and HBDI in ethanol (Table S3). The data points are labeled with the identities of residue 203, and additional mutations are noted in the parentheses. The overall trend shows a positive correlation of these two quantities, but is not even roughly linear for the following reasons. First, from the color tuning mechanisms discussed in Figure 4, it is easier to tune the A state for redder mutants while the B state is more tunable when blue. Second, there are significant hydrogen bonding network rearrangement in the vicinity of the chromophore between A and B states, especially for mutants with S65, T203, or V203 [21] or HBDI in ethanol. This causes an extra stabilization in the P form of the B state, leading to a bluer B state absorption than expected and deviating from the trend of S65T π–π stacking series (red box), for which no significant structural rearrangement should occur. Given the nonlinearity of the absorption maximum to electrostatics (e.g. Equation S5), it is better to study the correlation in terms of driving forces (Figure S5).

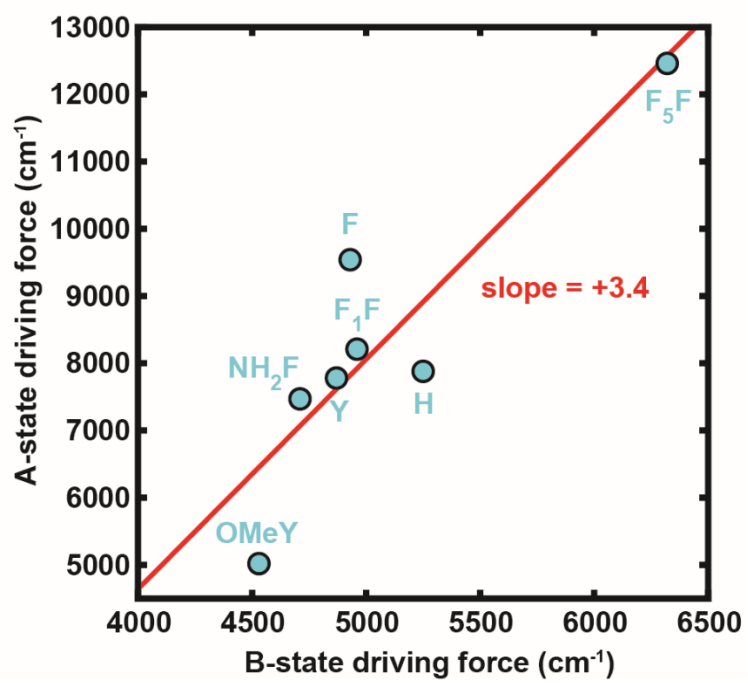


Figure S5. Correlation plot between the A-state and B-state driving forces from π–π stacking S65T mutants (aromatic residue at position 203), whose values are highlighted in blue in Table S5. The data points are labeled with the identities of residue 203 for the mutants. This series is chosen since π–π stacking allows for minimal structural rearrangement between the environments of the protonated and deprotonated chromophore, rendering the comparison between A and B states valid within these mutants. A linear fit (red) has a slope of 3.4, which agrees well to the ratio of the diabatic dipole moments between the two protonation states ($\frac{\mu_{CT}}{\Delta\mu_{CT}} \approx 3.1$ from Figure 4). From this plot, we can estimate the absorption maximum for the protonated GFP chromophore in vacuum, given the experimentally determined gas phase value for the deprotonated chromophore (479 nm from Nielsen and colleagues [22]). The latter corresponds to a B-state driving force of 8100 cm^{-1}. By extrapolation, the A-state chromophore in vacuum should possess a driving force of 18700 cm^{-1}, which translates to an absorption maximum at 382 nm according to Equation S5 using the parameters determined from Figure 4. The exact experimental value is still under debate. The Nielsen group determined it to be 370 ± 5 nm [23], which is close to our estimation; however, 340 ± 5 nm was obtained using multiphoton ionization from Greenwood *et al.* [24]. In our opinion, the latter seems unlikely as it is bluer than the bluest possible absorption $\Delta\bar{\nu}_{LE}$ (= 366 nm) obtained from Figure 4A, unless the chromophore geometry in the gas phase deviates substantially from that in the GFP environment.

S7 Supplementary Tables

Table S2. Summary of the second-derivative analysis of GFP mutants and the model chromophore HBDI in ethanol at 77 K in the protonated A state (Figures 3D and S2), presented in order of the 0–0 position). The numbering scheme follows Table S12 in [2]. Even though the difference between 0–1 and 0–0 energies for HBDI in ethanol is an apparent outlier, it is curiously half as much as a typical BLA vibrational frequency (~ 1320 cm^{-1}), as also evidenced by its 0–2 energy at 25800 cm^{-1} (Figure S2).

number	species	0–0 position (nm)	0–0 energy (cm^{-1})	0–1 energy (cm^{-1})	difference between 0–1 and 0–0 energies (cm^{-1})
S65T GFP mutants					
2	ih^{mat}(65T) · ~~ih:loop~~:GFP R96M	403.5	24783	26144	1361
12	s10(203F_5F) · ~~s10:loop~~:GFP S65T	410.4	24366	25700	1334
16	s10(203F) · ~~s10:loop~~:GFP S65T	413.1	24207	25465	1258
8	ih:GFP S65T	415.2	24085	25368	1283
15	s10(203(4-F_1F)) · ~~s10:loop~~:GFP S65T	417.0	23981	25233	1253
10	supercharged -30	420.0	23810	25214	1405
13	s10:loop:GFP S65T T203V	420.0	23810	25157	1348
18	s10(203(4-NH_2F)) · ~~s10:loop~~:GFP S65T	420.6	23776	25138	1363
17	s10:loop:GFP S65T T203Y	423.3	23624	24931	1308
14	s10:loop:GFP S65T T203H	424.5	23557	24988	1430
19	ih:GFP S65T T203(3-OMeY)	428.1	23359	24618	1259
S65 GFP mutants					
24	s10:loop:GFP T203V	413.1	24207	25484	1277
20	s10:loop:GFP	417.9	23929	25195	1266
27	s10:loop:GFP T203Y	424.5	23557	24950	1393
26	ih:GFP T203(3-OMeY)	434.4	23020	24242	1222
GFP model chromophore					

66	HBDI in ethanol	408.6	24474	25082	608

Table S3. A-state and B-state absorption maxima (the latter cited from Table S14 in [2]) for GFP mutants at 77 K, listed in the order of decreasing A-state absorption energy. The numbering scheme follows Table S12 in [2]. Mutants colored with shades of orange carry threonine at the 203 position, which assume different rotamers in two different protonation states (Figure 1 in the main text [21]).

number	species	A-state absorption maximum (nm)	A-state absorption maximum (cm^{-1})	B-state absorption maximum (nm)	B-state absorption maximum (cm^{-1})
S65T GFP mutants					
2	ih^{mat}(65T) · ~~ih:loop~~:GFP R96M	384.0	26040	462.0	21650
12	s10(203F_5F) · ~~s10:loop~~:GFP S65T	388.8	25720	497.4	20110
8	ih:GFP S65T	392.7	25470	485.7	20590
16	s10(203F) · ~~s10:loop~~:GFP S65T	394.2	25370	503.4	19870
13	s10:loop:GFP S65T T203V	396.6	25210	494.1	20240
10	supercharged -30	397.2	25180	485.4	20600
15	s10(203(4-F_1F)) · ~~s10:loop~~:GFP S65T	397.5	25160	513.0	19490
14	s10:loop:GFP S65T T203H	398.4	25100	510.3	19600
17	s10:loop:GFP S65T T203Y	398.7	25080	514.2	19450
18	s10(203(4-NH_2F)) · ~~s10:loop~~:GFP S65T	399.6	25030	516.0	19380
19	ih:GFP S65T T203(3-OMeY)	408.6	24470	513.0	19490
S65 GFP mutants					
24	s10:loop:GFP T203V	392.4	25290	465.6	21480
20	s10:loop:GFP	397.8	25140	465.6	21480
27	s10:loop:GFP T203Y	399.0	25060	508.8	19650
26	ih:GFP T203(3-OMeY)	413.7	24170	502.8	19890

Table S4. Absorption maxima and Stark tuning rates for A states of GFP mutants at 77 K listed in the order of decreasing absorption maximum. The numbering scheme is the same as Table S12 in [2]. For mutants where only B-state properties are measured (Figure S3), the corresponding reasons are given below. The energy gap between the CT and LE forms (denoted as $\Delta\bar{\nu}$) and degree of CT character in S_1 state is estimated using Equation S5 and S6, respectively. Note that for all mutants we have accessed, the CT form is consistently higher in energy than the LE form and none of the estimated CT degree exceeds 50%.

number	species	A-state absorption maximum (cm^{-1})	A-state Stark tuning rate (D)	energy gap between CT and LE $\Delta\bar{\nu}$ (cm^{-1})	degree of CT character in S_1 state
S65T GFP mutants					
2	ih^{mat}(65T) · ~~ih:loop~~:GFP R96M	26040	6.1	16360	6.7%
12	s10(203F_5F) · ~~s10:loop~~:GFP S65T	25720	10.6	12460	10.1%
8	ih:GFP S65T	25470	10.6	10250	13.2%
16	s10(203F) · ~~s10:loop~~:GFP S65T	25370	14.1	9540	14.4%
13	s10:loop:GFP S65T T203V	25210	12.4	8550	16.4%
10	supercharged -30	25180	11.0	8320	16.9%
15	s10(203(4-F_1F)) · ~~s10:loop~~:GFP S65T	25160	13.9	8210	17.2%
14	s10:loop:GFP S65T T203H	25100	13.4	7880	17.9%
17	s10:loop:GFP S65T T203Y	25080	15.9	7780	18.2%
18	s10(203(4-NH_2F)) · ~~s10:loop~~:GFP S65T	25030	17.0	7470	18.9%
19	ih:GFP S65T T203(3-OMeY)	24470	20.8	5020	26.5%
1	ih^{mat}(65T)·~~ih:loop~~:GFP R96E E222K	no A state observed [2]			
3	ih:GFP S65T E222Q				
4	s10:loop:GFP S65T				
5	supercharged +36 s10-				
7	supercharged +36				

9	ih:GFP S65T H148D	A state with a short hydrogen bond with D148 [25]			
S65 GFP mutants					
24	s10:loop:GFP T203V	25290	13.6	9030	15.4%
20	s10:loop:GFP	25140	11.3	8100	17.4%
27	s10:loop:GFP T203Y	25060	12.1	7670	18.4%
26	ih:GFP T203(3-OMeY)	24170	24.3	3960	30.6%
22	s10:loop:GFP E222Q	no A state observed [2]			
23	avGFP	significant overlap between A and B states [2]			
25	s10:loop:GFP T203V E222Q	no A state observed [2]			
28	s10:loop:GFP T203Y E222Q				
GFP model chromophore					
66	HBDI in ethanol	25940	6.3	14950	7.7%

Table S5. The energy gaps (or driving forces) between diabatic forms for the neutral and anionic GFP chromophores evaluated in the GFP environments. The latter is reproduced here from Table S13 in [2]. The numbering scheme follows Table S12 in [2]. Mutants with a S65T chromophore and an aromatic residue at the 203 position are highlighted in blue. Since within this set of mutants, the π–π interaction stays relatively rigid and S65T chromophore disallows hydrogen bonding network rearrangement upon chromophore protonation, the comparison between A-state and B-state values is meaningful unlike others (Figure S5). In contrast, there must be a significant change in solvation structures of ethanol when interacting with the neutral and anionic chromophores, so the comparison between the two corresponding driving forces is unrealistic.

number	species	energy gap between CT and LE for A state (cm^{-1})	energy gap between I and P for B state (cm^{-1})
S65T GFP mutants			
2	ih^{mat}(65T) · ~~ih:loop~~:GFP R96M	16360	10270
12	s10(203F_5F) · ~~s10:loop~~:GFP S65T	12460	6320
8	ih:GFP S65T	10250	7880
16	s10(203F) · ~~s10:loop~~:GFP S65T	9540	4930
13	s10:loop:GFP S65T T203V	8550	6170
10	supercharged -30	8320	7540
15	s10(203(4-F_1F)) · ~~s10:loop~~:GFP S65T	8210	4960
14	s10:loop:GFP S65T T203H	7880	5250
17	s10:loop:GFP S65T T203Y	7780	4870
18	s10(203(4-NH_2F)) · ~~s10:loop~~:GFP S65T	7470	4710
19	ih:GFP S65T T203(3-OMeY)	5020	4530
S65 GFP mutants			
24	s10:loop:GFP T203V	9030	7600
20	s10:loop:GFP	8100	9770
27	s10:loop:GFP T203Y	7670	5680
26	ih:GFP T203(3-OMeY)	3960	6340

GFP model chromophore			
66	HBDI in ethanol	14950	12200

S8 References

[1] Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science* **2020**, *367*, 76-79.

[2] Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified model for photophysical and electro-optical properties of green fluorescent proteins. *J. Am. Chem. Soc.* **2019**, *141*, 15250-15265.

[3] Zuris, J. A.; Thompson, D. B.; Shu, Y.; Guilinger, J. P.; Bessen, J. L.; Hu, J. H.; Maeder, M. L.; Joung, J. K.; Chen, Z.-Y.; Liu, D. R. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat. Biotechnol.* **2015**, *33*, 73-80.

[4] Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. Mechanism and bottlenecks in strand photodissociation of split green fluorescent proteins (GFPs). *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2146-E2155.

[5] Bublitz, G. U.; Boxer, S. G. Stark spectroscopy: applications in chemistry, biology, and materials science. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213-242.

[6] Andrews, S. S.; Boxer, S. G. A liquid nitrogen immersion cryostat for optical measurements. *Rev. Sci. Inst.* **2000**, *71*, 3567-3569.

[7] Pauszek, R. F.; Stanley, R. J. A “How-To” Guide to the Stark Spectroscopy of Flavins and Flavoproteins. In *Methods in Molecular Biology (Methods and Protocols) Volume 1146: Flavins and Flavoproteins*; Weber, S.; Schleicher, E., Eds.; Humana Press: New York, 2014; pp 443-466.

[8] Olsen, S. Locally-excited (LE) versus charge-transfer (CT) excited state competition in a series of para-substituted neutral green fluorescent protein (GFP) chromophore models. *J. Phys. Chem. B* **2015**, *119*, 2566-2575.

[9] Goldstein, H.; Poole, C.; Safko, J. *Classical Mechanics*, 3rd ed.; Addison-Wesley: Boston, Massachusetts, 2002; pp 483-509.

[10] Strogatz, S. H. *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*, 1st ed.; Perseus: Cambridge, Massachusetts, 1994.

[11] Drobizhev, M.; Tillo, S.; Makarov, N. S.; Hughes, T. E.; Rebane, A. Color hues in red fluorescent proteins are due to internal quadratic Stark effect. *J. Phys. Chem. B* **2009**, *113*, 12860-12864.

[12] Hasegawa, J.; Fujimoto, K. J.; Hakatsuji, H. Color tuning in photofunctional proteins. *ChemPhysChem* **2011**, *12*, 3106-3115.

[13] Drobizhev, M.; Hughes, T. E.; Stepanenko, Y.; Wnuk, P.; O'Donnell, K.; Scott, J. N.; Callis, P. R.; Mikhaylov, A.; Dokken, L.; Rebane, A. Primary role of the chromophore bond length alternation in reversible photoconversion of red fluorescence proteins. *Sci. Rep.* **2012**, *2*, 688.

[14] Heckmann, A.; Lambert, C. Organic mixed-valence compounds: a playground for electrons and holes. *Angew. Chem. Int. Ed.* **2012**, *51*, 326-392.

[15] Gouterman, M. Study of the effects of substitution on the absorption spectra of porphin. *J. Chem. Phys.* **1959**, *30*, 1139-1161.

[16] Gouterman, M. Spectra of porphyrins. *J. Mol. Spectrosc.* **1961**, *6*, 138-163.

[17] Kanchanawong, P.; Dahlbom, M. G.; Treynor, T. P.; Reimers, J. R.; Hush, N. S.; Boxer, S. G. Charge delocalization in the special-pair radical cation of mutant reaction centers of *Rhodobacter sphaeroides* from Stark spectra and nonadiabatic spectral simulations. *J. Phys. Chem. B* **2006**, *110*, 18688-18702.

[18] Zhou, H.; Boxer, S. G. Charge resonance effects on electronic absorption line shapes: application to the heterodimer absorption of bacterial photosynthetic reaction centers. *J. Phys. Chem. B* **1997**, *101*, 5759-5766.

[19] Fano, U. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.* **1961**, *124*, 1866-1878.

[20] Drobizhev, M.; Callis, P. R.; Nifosì, R.; Wicks, G.; Stoltzfus, C.; Barnett, L.; Hughes, T. E.; Sullivan, P.; Rebane, A. Long- and short-range electrostatic fields in GFP mutants: implications for spectral tuning. *Sci. Rep.* **2015**, *5*, 13223.

[21] Craggs, T. D. Green fluorescent protein: structure, folding and chromophore maturation. *Chem. Soc. Rev.* **2009**, *38*, 2865-2875.

[22] Nielsen, S. B.; Lapierre, A.; Andersen, J. U.; Pedersen, U. V.; Tomita, S.; Andersen, L. H. Absorption spectrum of the green fluorescent protein chromophore anion *in vacuo*. *Phys. Rev. Lett.* **2001**, *87*, 228102.

[23] Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Rocha-Rinza, T.; Christiansen, O.; Bravaya, K. B.; Erokhin, A. V.; Bochenkova, A. V.; Solntsev, K. M.; Dong, J.; Kowalik, J.; Tolbert, L. M.; Petersen, M. Å.; Nielsen, M. B. Photoabsorption studies of neutral green fluorescent protein model chromophores *in vacuo*. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9996-10002.

[24] Greenwood, J. B.; Miles, J.; De Camillis, S.; Mulholland, P.; Zhang, L.; Parkes, M. A.; Hailes, H. C.; Fielding, H. H. Resonantly enhanced multiphoton ionization spectrum of the neutral green fluorescent protein chromophore. *J. Phys. Chem. Lett.* **2014**, *5*, 3588-3592.

[25] Oltrogge, L. M.; Boxer, S. G. Short hydrogen bonds and proton delocalization in green fluorescent protein (GFP). *ACS Cent. Sci.* **2015**, *1*, 148-156.