ACS central science

Short Hydrogen Bonds and Proton Delocalization in Green Fluorescent Protein (GFP)

Luke M. Oltrogge and Steven G. Boxer*

Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States

Ⓢ Supporting Information

ABSTRACT: Short hydrogen bonds and specifically low-barrier hydrogen bonds (LBHBs) have been the focus of much attention and controversy for their possible role in enzymatic catalysis. The green fluorescent protein (GFP) mutant S65T, H148D has been found to form a very short hydrogen bond between Asp148 and the chromophore resulting in significant spectral perturbations. Leveraging the unique autocatalytically formed chromophore and its sensitivity to this interaction we explore the consequences of proton affinity matching across this putative LBHB. Through the use of noncanonical amino acids introduced through nonsense suppression or global incorporation, we systematically modify the acidity of the GFP chromophore with halogen substituents. X-ray crystal structures indicated that the length of the interaction with Asp148 is unchanged at ∼2.45 Å while the absorbance spectra demonstrate an unprecedented degree of color tuning with increasing acidity. We utilized spectral isotope effects, isotope fractionation factors, and a simple 1D model of the hydrogen bond coordinate in order to gain insight into the potential energy surface and particularly the role that proton delocalization may play in this putative short hydrogen bond. The data and model suggest that even with the short donor−acceptor distance (∼2.45 Å) and near perfect affinity matching there is not a LBHB, that is, the barrier to proton transfer exceeds the H zero-point energy.
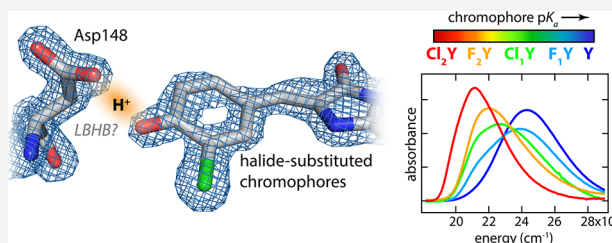
INTRODUCTION

Hydrogen bonds play a critical role in the structure and function of biological macromolecules. In this capacity they exist in a wide range of geometries (i.e., lengths and angles) and strengths. Of particular interest are the mechanisms whereby H-bonds facilitate enzymatic catalysis.[1,2] Many examples are known in which the removal of even a single H-bond can slow down reactions by several orders of magnitude.[3,4] Structural surveys of enzymes with transition state analogues bound have revealed an unusual prevalence of abnormally short H-bonds (<2.5 Å), and this has led to much speculation about the possible role of these interactions.[5−7] Low-barrier hydrogen bonds (LBHBs)—those in which the barrier to proton transfer is of the same order as the zero-point energy—have been proposed to stabilize transition states through quantum resonance.[5,8,9] However, comparisons to normal H-bonding in the uncatalyzed reactions in solution and considerations of charge solvation inside proteins have cast doubt on the original premise.[10−12] Nonetheless, the abundance of short H-bonds, whether low-barrier or otherwise, in enzyme active sites is probably not coincidental.

Definitive identification of LBHBs in proteins through experimental means has proven to be challenging. Most putative protein LBHBs are identified indirectly either by particularly short H-bond donor−acceptor distances or by the appearance of far downfield proton resonances in ^{1}H NMR. In several cases more quantitative metrics such as isotope fractionation factors have been measured,[13] and there is at

least one example of a direct observation of a LBHB with neutron diffraction crystallography.[14]

Many studies in small molecules have convincingly identified LBHBs in crystals and in the gas phase;[15,16] however solution studies—in some cases of the very same molecules—have revealed protons predominantly localized to either heteroatom.[17−19] This effect has been attributed to the inherently asymmetric solvation microenvironment experienced by each molecule in solution that tends to result in pK_a mismatches which bias the proton to one side or the other.[18] This is an illustration of the delicate balancing act between proton binding sites mediated by the intrinsic pK_a's and their interplay with the environment necessary to sustain a true LBHB. These considerations have led some to conclude that most inferred LBHBs are instead short ionic H-bonds absent very precise proton affinity matching.[12]

A short donor−acceptor distance is a necessary but not sufficient condition to form a LBHB. The other critical requirement, as mentioned above, is that there be close "pK_a matching" between the two proton binding sites. Protein environments will, in general, perturb the pK_a's of buried ionizable groups. To avoid confusion with the solution pK_a's we favor the term *differential acidity* (ΔpK_α, where we use the subscript α to distinguish this from the conventional pK_a) as a metric of the in situ mismatch of donor and acceptor proton binding energies in pK_a units.[20] (We adopt the convention that
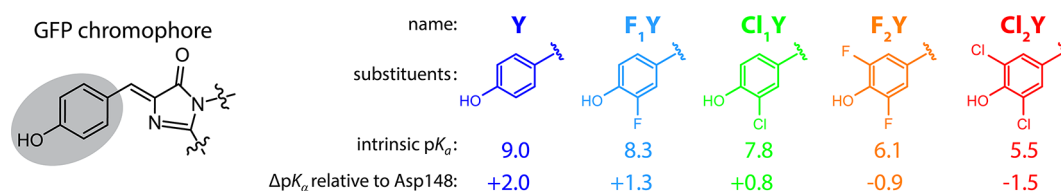
Figure 1. (Left) Structure of the GFP chromophore with the phenolic portion highlighted. (Right) Introduced halogen substituents with their corresponding shorthand names, intrinsic denatured pK_a values (see Figure S.3), and their differential acidities with respect to Asp148 in the folded protein deduced from the model described below.

$\Delta pK_\alpha = pK_{\alpha(Cro)} - pK_{\alpha(Asp148)}$, where Cro denotes Chromophore. This means that $\Delta pK_\alpha > 0$ implies greater proton stabilization on the chromophore while $\Delta pK_\alpha < 0$ implies greater stabilization on Asp148.) Thus, it is of interest to empirically vary ΔpK_α —ideally at fixed geometry—in order to answer two questions: (i) How much must the acidity of one site be altered to match that of the other (i.e., until $\Delta pK_\alpha = 0$)? (ii) When ΔpK_α is close to zero, is the interaction a LBHB? In pursuit of this goal we focus on a green fluorescent protein (GFP) mutant thought to contain a LBHB and take advantage of the unique sensitivity and specificity of the chromophore absorption to access a range of measurable (and in principle calculable) properties diagnostic of the underlying H-bond energetics.

GFP has long been an indispensable tool in cellular imaging due to its autocatalytically formed fluorescent chromophore. The single most important determinant of its visible absorbance spectrum is the protonation state of the phenolate portion of the chromophore. The neutral chromophore, so-called A-state, has an absorbance peaked near 400 nm (25,000 cm^{-1}) while the anion, the B-state, is peaked around 470 nm (21,200 cm^{-1}).[21] Wild-type GFP features an ultrafast excited-state proton transfer (ESPT) reaction in which excitation of the A-state results in fluorescence emission characteristic of the B-state in a matter of picoseconds.[22] The mutation S65T abrogates this process and dramatically increases the chromophore's sensitivity to solution pH having an in situ pK_a of ~5.7.[23,24] A second mutation, H148D, rescues ESPT by positioning an alternative terminal proton acceptor with an extremely short H-bond to the phenolic oxygen of <2.4 Å[25] and has a rate of ESPT less than 100 fs.[26,27] Furthermore, the A-state absorbance band was significantly red-shifted to 415 nm and, uniquely among GFPs, demonstrated an absorbance band shift upon exchange to D$_2$O.[26] All of these factors led to the proposal that the Asp148−chromophore interaction is a LBHB. In the present work we critically analyze this claim by incrementally increasing the chromophore acidity and investigating the spectral and structural consequences of these perturbations.

A key enabling technology for this undertaking is the introduction of noncanonical amino acids through the technique of nonsense suppression. This method utilizes a tRNA complementary to the Amber stop codon (UAG) which is charged with an exogenously added synthetic amino acid by an engineered aminoacyl-tRNA synthetase. In recent years as this technology has matured, large libraries of available noncanonical amino acids have been created and the production of proteins containing these has become increasingly commonplace.[28] Whereas conventional mutagenesis is limited to rather large structural and functional jumps among the 20 naturally occurring amino acids, nonsense suppression presents opportunities to introduce perturbations which are far

more subtle. Thus, we incrementally modify the acidity of the GFP chromophore through the introduction of halide substituted tyrosines at position 66. This residue together with Thr65 and Gly67 participates in the autocatalytic chromophore maturation reaction in the folded protein. Once maturation is complete, the halide-substituted tyrosyl group comprises the phenolic portion of the chromophore (Figure 1).

Methods. The halogen substituted tyrosines were prepared via chemical or enzymatic synthesis and introduced into recombinantly expressed GFP with nonsense suppression or global incorporation. All constructs are based on a circularly permuted Superfolder GFP and, unless otherwise indicated, contain the mutations S65T and H148D. The naming shorthand for proteins with modified chromophores is **Y**, Tyr66; **F$_1$Y**, Y66(3-fluoro-Tyr); **Cl$_1$Y**, Y66(3-chloro-Tyr); **F$_2$Y**, Y66(3,5-difluoro-Tyr); and **Cl$_2$Y**, Y66(3,5-dichloro-Tyr).

For detailed methods see Supporting Information (SI) S.2.

■ RESULTS AND DISCUSSION

Structural Evidence for Short Hydrogen Bonds. When evaluating the role of ΔpK_α it is necessary that the geometry of the interaction, particularly the O−O distances, be unchanged in order to isolate the effect. To this end we obtained X-ray crystal structures of **Y**, **Cl$_1$Y**, and **Cl$_2$Y** with PDB IDs 4ZF3, 4ZF4, and 4ZF5, respectively (see SI S.3).

The most important conclusion to draw from these structures is that r_{O-O} is indeed maintained very closely across this series at approximately 2.45 Å (Figure 2 and Table S.2). (Note that in **Cl$_1$Y** the chlorine atom could in principle occupy either position due to the free rotation of the tyrosine phenol group prior to chromophore maturation. However, we observe 100% occupancy of the shown isomer.) This result is assumed to hold also for **F$_1$Y** and **F$_2$Y**, for which we do not have
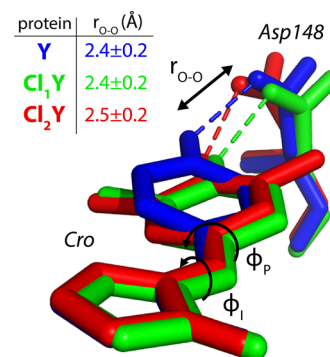
Figure 2. Overlay of the chromophore (Cro) and Asp148 for the aligned structures of **Y**, **Cl$_1$Y**, and **Cl$_2$Y** with tabulated r_{O-O} distances. **Y** and **Cl$_1$Y** are from the A chains which had lower B-factors. **Cl$_2$Y** is from chain B because chain A had an alternate deprotonated conformation (see SI S.6).

structures, because fluorine is sterically smaller and exerts a lesser effect on the acidity than chlorine. The structures for **Y**, **Cl$_1$Y**, and **Cl$_2$Y** adopt a different conformation of Asp148 than in the original structure of S65T, H148D from Shu et al. (PDB ID: 2DUF),[25] which was in an otherwise WT background and not circularly permuted. However, the phenomenology of **Y** in comparison to the original protein is nearly indistinguishable including the perturbed absorbance band, rapid ESPT, and the spectral isotope effect, providing strong evidence of functional similarity. Another difference is that **Cl$_1$Y** and **Cl$_2$Y** demonstrate a noticeable twist in the chromophore geometry relative to **Y** (Figure 2). The concern is that this twist, rather than the degree of proton sharing, may dominate the absorbance. This is likely not the case, however, because **Cl$_1$Y** and **Cl$_2$Y** have almost identical A- and B-state spectra under denaturing conditions (Figure S.1) and nearly the same twist in the structure. Yet, the natured proteins at low pH show dramatically different peak positions correlated with the chromophore acidity (Figure 3).
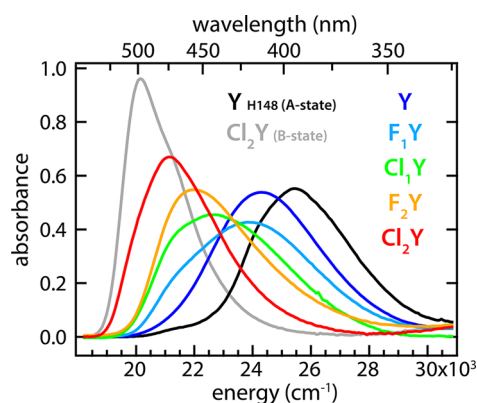
Figure 3. Spectra of natured substituted GFP protein in the limit of low pH (colored traces) shown with spectra for pure A- and B-states for comparison. All spectra are scaled by their associated B-state oscillator strength.

All three structures were obtained at pH 5.0 in order to favor the protonated form (Figure S.3). While the X-ray resolution is insufficient to visualize protons, we can infer the existence of a single proton between Asp148 and the chromophore on the basis of both the prohibitive energetic penalty a doubly deprotonated species would incur and the presence in all species of only two optical states as a function of pH (Figures S.2 and S.3). Unexpectedly, we observed in one of two chains of the **Cl$_2$Y** asymmetric unit an apparently deprotonated structure in which the ionized Asp148 was relocated away from the chromophore (Figure S.6). This is, to our knowledge, the first example of an alternative backbone geometry of β-strand 7 (Figure S.7). Strand 7 has been the object of considerable attention for its dynamical role in proton transfer,[29] FP sensor mechanisms,[30] and peptide photodissociation.[31] This new data may facilitate the development of specific structural models of these transitions in the future. The interested reader is referred to SI S.6.

pK_a Titration and Spectral Response. In order to assess the effects of the halogen substitutions on the absorbance bands and the intrinsic pK_a's of the chromophore, pH titrations were performed on protein denatured in 6 M guanidinium HCl to expose the chromophore to solution. As expected the chromophore pK_a's changed in response to halogen sub-

stitution ultimately spanning a range of 3.5 pK_a units (Figure S.3). Furthermore, the absorbance bands due to the protonated and deprotonated forms of all species were largely unaffected by the presence of the substituents (Figure S.1).

pH titrations of the natured protein were also measured and found to have clean isosbestic points. Like the denatured protein, the B-state absorbance bands showed only small shifts among the halogen substituted species (Figure S.2). In contrast, the A-state bands were massively perturbed by the halogen substituents causing the peaks to shift across the entire spectral range, with those with high intrinsic pK_a's resembling the usual A-state and those with lower intrinsic pK_a's resembling the B-state (Figure 3). This result is particularly striking when viewed in light of the exceptional consistency of the A-state absorbance band energy across the enormous variety of GFP mutants all peaked near 397 nm (25,200 cm^{-1}).

The series of halogen substituted chromophores can be considered as internally titrating against Asp148. In other words, progressively increasing the acidity of the chromophore with respect to Asp148 decreases Δp$K_α$ until matched (Δp$K_α$ = 0) and beyond (Δp$K_α$ < 0), at which point the proton is more stable on Asp148. Depending on the distance—which we know to be ~2.45 Å—and the degree of coupling, the potential energy surface along this bond will dictate the character of this proton "tug-of-war" between these sites.

Before undertaking a semiquantitative approach toward this interaction, it is important to clearly define what a LBHB is in unambiguous energetic terms.[12] We define a LBHB as an H-bonding interaction in which the barrier to proton transfer between the donor and acceptor atoms is equal to or lower than the proton zero-point energy (H-ZPE). This definition implicitly contains the requirement of Δp$K_α$ near to zero. A further consequence is that the proton probability density of a LBHB thus defined will have a single broad peak at or near the bond midpoint.

Physical Model. In the analysis of this short H-bond we have utilized the coupled Morse potential model put forward by Ross McKenzie.[32,33] This model has the advantage of using only two fit parameters and retaining physical transparency while describing experimental H-bond properties with reasonable accuracy. We have adapted the linear bond model in order to calculate approximate protein absorbance spectra as a function of chromophore acidity and can additionally make predictions for the trends in spectral isotope effects and isotope fractionation factors. The basic method is described below, and additional information can be found in SI S.5.

The model takes as input two Morse potential diabatic states for the proton donor and acceptor sites. These potential energy functions, representing the respective proton-bound states at infinite separation, serve as the diagonal of a 2 × 2 Hamiltonian matrix with the scalar coupling energy (Δ$_{DA}$) on the off-diagonal. Matrix diagonalization yields as eigenvalues the ground- and excited-state adiabatic potential energy surfaces (PESs). The associated eigenvectors describe the relative contribution of the two input diabats at each point along the bond. These are used to calculate the mapping between a particular proton position and the corresponding excitation energy. For example, if at some location there is 50% contribution from both A- and B-state diabats, then the excitation energy would be halfway between the A- and B-state basis states. Next, the proton (or deuteron) probability distribution functions on this PES are determined using a finite difference solution to the 1D Schrödinger equation. The
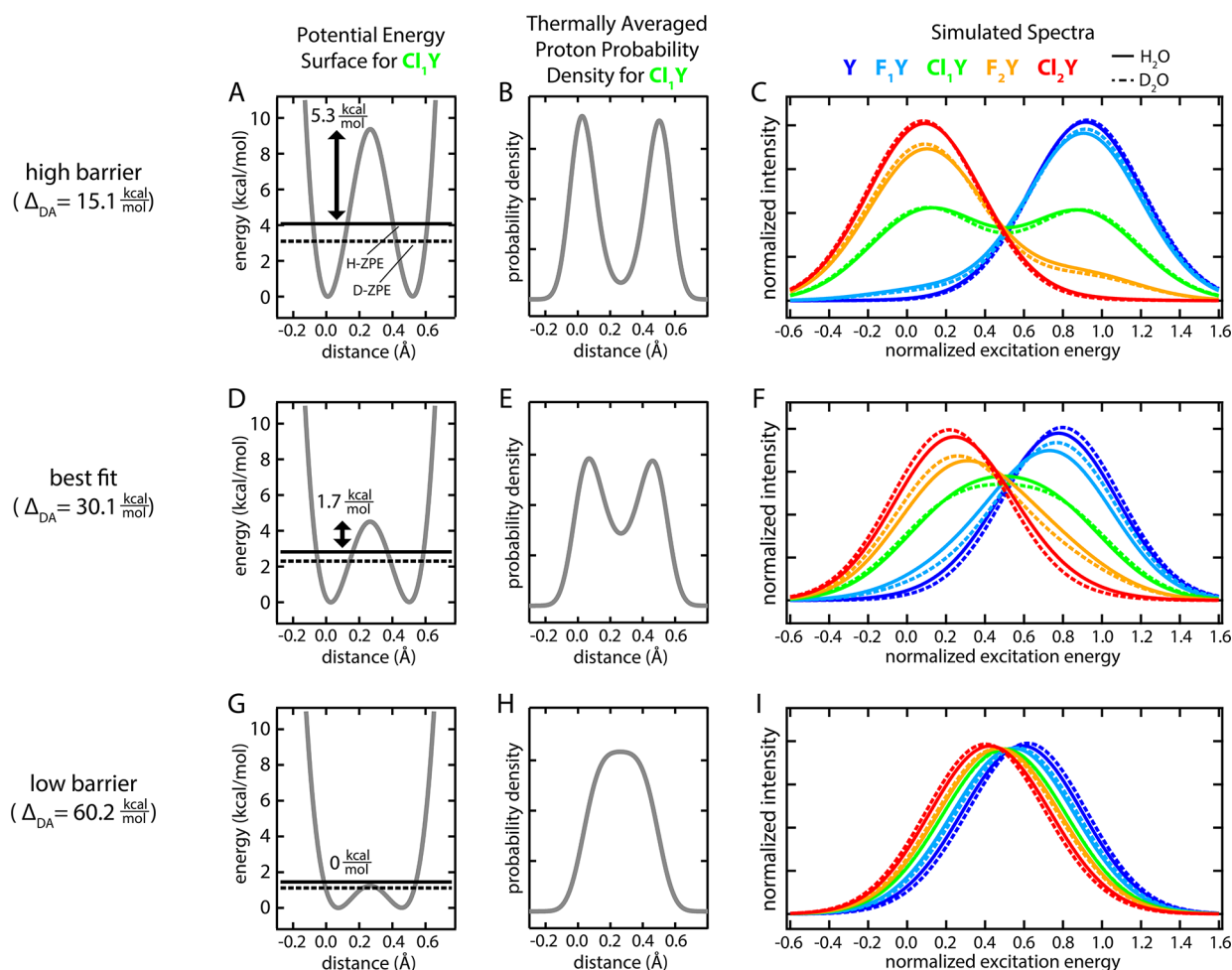
Figure 4. Comparison of the model results for the PESs, thermal probability densities, and simulated absorbance spectra among coupling energies (Δ_{DA}) for high-barrier, best fit, and low-barrier proton transfer. Note that while the PESs show only the ground state, the effect of low lying excited states are taken into account for both the thermal probability densities and simulated spectra.

resulting eigenstates are thermally populated according to Maxwell−Boltzmann statistics. The final step is to convert these probability densities into a spectrum using the mapping described above. This was accomplished by performing a probability-weighted histogram of excitation energies and then convolving the result with normalized Gaussians with width characteristic of the pure A-/B-state absorbance bands to approximate the homogeneous and inhomogeneous broadening. The process was repeated for each of the substituted chromophores where the input diabat for Asp148 remained the same while the chromophore diabats have proton binding energies scaled appropriately for their intrinsic pK_a differences. [The differences in proton binding free energy are calculated by $\Delta\Delta G° = \ln 10 \times RT\Delta pK_a$. Due to the chemical similarities among the substituted chromophores, we make the assumption that $\Delta\Delta G°$ is dominated by the enthalpic portion ($\Delta\Delta H°$). Thus, the input chromophore Morse potential binding energies are shifted by $\Delta\Delta G°$ directly. Furthermore, when transferred to the protein environment, the chromophore's $\Delta H°$ of proton binding will surely change, however, we assume that the relative enthalpy ($\Delta\Delta H°$) among the substituted chromophores will remain the same.] Apart from these energy shifts, the input diabats are all identical. The final model fixes the O−O distance to 2.45 Å and varies global parameters for (i) the energy offset between Asp148 and the set of chromophore diabats and (ii)

the magnitude of the coupling parameter between the sites (Δ_{DA}) to output simulated absorbance spectra. Lastly, the parameters giving the best overlap to energy-normalized experimental spectra were obtained using global nonlinear least-squares optimization in Matlab.

Figure 4 illustrates the results of applying this model with a range of coupling energies (Δ_{DA}). The best fit was obtained with the coupling parameter Δ_{DA} = 30.1 kcal/mol. For comparison the parameters fit by McKenzie for the distance dependence of the coupling predict Δ_{DA} = 38.9 kcal/mol at 2.45 Å. The offset energy parameter was such that $\Delta pK_\alpha = 0$ was found to occur between Cl_1Y and F_2Y. More specifically, a hypothetical substituted chromophore with an intrinsic denatured pK_a of 7.0 would be expected to share the proton equally with Asp148 (Figure 1). One might be tempted to directly compare the intrinsic pK_a of aspartic acid (~3.9) with that of the substituted chromophores in solution or to conclude that Asp148 has a pK_a of 7.0 in the protein. These comparisons, however, are not correct because they neglect the protein solvation environment which will, in general, affect the two sites differently. Consequently, when referring to this interaction within the protein we can speak only of differential acidities (ΔpK_α).

The simulated spectra with the best-fit value of Δ_{DA} accurately capture the shift in energy as well as the changes

in width of the "A state" absorbance band among the substituted chromophores (Figure 4F). This shift in peak position would seem to suggest a low-barrier or single-welled potential, however, the best-fit PESs indicate that there is a barrier between the sites with a height exceeding the H-ZPEs (Figure 4D; see also Figure 7A). Consequently, the proton probability density appears as two distinct peaks (Figure 4E). To critically assess the uniqueness of the coupling energy fit (Δ_{DAfit}) and to better understand the relationship between the barrier height and resulting spectra, we performed simulations for $0.5\Delta_{DAfit}$, Δ_{DAfit}, and $2\Delta_{DAfit}$ (Figure 4). For small coupling energy ($0.5\Delta_{DAfit}$, Figure 4A,B,C), and thus large barriers, the spectra present two distinct bands of fixed position and width which are essentially the A- and B-state basis spectra. With large coupling energy ($2\Delta_{DAfit}$, Figure 4G,H,I) the H-ZPE exceeds the barrier height, thus representing the LBHB case, and the absorbance bands shift position over a relatively small energy range. However, the bandwidth is similar to that of the A- or B-state basis spectra. At the best fit coupling energy (Δ_{DAfit}, Figure 4D,E,F) the peak position shifts over a greater energy range and is accompanied by a dramatic change in the peak width with the greatest width occurring for matched affinities (Cl_1Y, Figure 4E). This behavior is most consistent with the experimental data (Figure 3). The spectral isotope effects (SIEs)—the changes in the spectrum upon H to D exchange— are also quite distinct between the different coupling energies and will be discussed further below.

Isotope Effects. Equilibrium isotope effects are a powerful tool to probe potential energy surfaces. We utilize two equilibrium isotope phenomena in this study in order to obtain additional information and constraints on the nature of the Asp148-chromophore H-bond: spectral isotope effects (SIEs)—the spectral change induced by isotopic substitution— and isotope fractionation factors—the energetic isotope preference. We consider each of these in comparison to the model results across the substituted chromophore series. It should be emphasized that the global parameter fitting was performed only against the absorbance spectra in H_2O and that all the isotope effects emerged naturally from this fit.

In the model above we have described a method for mapping the hydron probability density to a predicted absorbance spectrum. Hydrogen to deuterium substitution, insofar as it alters this probability density, will lead to changes in the resulting spectra. Generally electronic absorbance spectra of molecules in water are insensitive to the mole fraction of deuterium even when possessing a titratable site. wtGFP's spectrum, for instance, has no detectable change. This can be understood through the fact that the chromophore participates only in "normal" H-bonds. In these longer interactions the hydron potential well is deep and is reasonably approximated as a harmonic oscillator. Thus, even though H to D substitution changes the ZPE, the position expectation value remains unchanged and so also the spectrum. In contrast, S65T H148D GFP has a short ionic H-bond, and the coupling between the proton binding sites can strongly perturb the PES. In particular, by introducing a much greater degree of anharmonicity, H may spread out toward the H-bond acceptor while D, by virtue of its lower ZPE, will remain more localized to the donor. With this qualitative picture we can evaluate the experimental SIEs and the model predictions.

Each of the halogen substituted proteins was exchanged into solutions of identical hydron activity (i.e., pH or pD) with varying mole fractions of deuterium, and the absorbance spectra
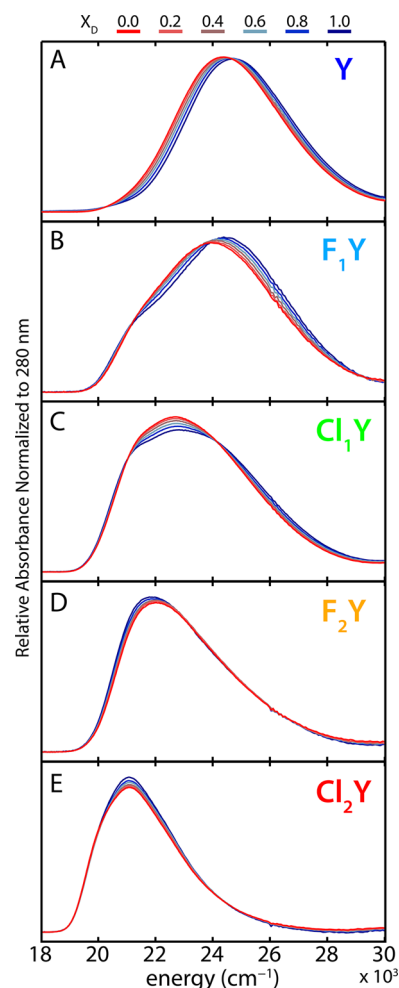

Figure 5. Spectral isotope effects (SIEs) for all proteins at pH(D) 5.0. 100% H_2O is shown in red transitioning to 100% D_2O in blue. Organic acids typically exhibit slightly higher pK_a's in D_2O, however, since most of these species are dominated by the protonated form at pH(D) 5.0, this effect should be negligible. Cl_2Y is 86% protonated at pH 5.0 and thus may have slight error.

were measured (Figure 5). The peak spectral isotope shifts (H to D) change sign through the acidity progression and decrease in magnitude out on the extrema to make an S-shaped curve (Figure 7C). (We note that there may also be secondary geometric isotope effects which would add some complication to the interpretation by subtly altering the bond length.[33] Due to the approximate nature of the model we neglect this in the present work.) This is broadly consistent with the logic described above if we assume the proton affinity of Asp148 to be constant while titrating it against the series of chromophores with variable proton affinities. For large $|\Delta pK_a|$ the shift tends toward zero because the PES is only weakly perturbed. As ΔpK_a approaches zero, the magnitude of the shift increases as the anharmonicity becomes more pronounced. However, once ΔpK_a changes sign, so also does the spectral isotope shift since H to D exchange leads to greater localization in the opposite well.

In addition to the spectral isotope shift there is also useful information in the changes in peak shapes. This is particularly true of Cl_1Y, the protein with the closest Asp148-chromophore proton affinity matching. The experimental data show a significant decrease in peak intensity along with some

broadening to both red and blue edges (Figure 5C). This same effect is captured in the best-fit spectra (solid and dashed green in Figure 4F). When the ZPE is slightly above the barrier there is very little SIE (solid and dashed green in Figure 4I) even though significant shifts are obtained for affinity mismatched species. In the event of an extremely strongly coupled single-welled potential one would actually anticipate the opposite effect, that is, an affinity matched D-bound species would have a sharper peak.

Further information about the energetics of the H-bond comes from the isotope fractionation factor (ϕ_{HD}). ϕ_{HD} is defined as the equilibrium constant for the exchange reaction in which an H in a bond or complex is swapped for a D from the solvent. Strongly coupled H-bonds energetically favor H over D and lead to H enrichment beyond the solvent composition (Figure 6). This effect is due to a reduction in the H/D ZPE
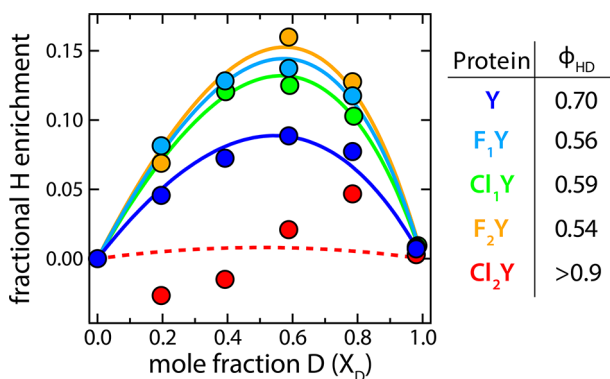
Figure 6. Fits to the isotope enrichment as a function of mole fraction deuterium with the calculated ϕ_{HD}'s tabulated on the right. **Cl$_2$Y** had only small differences in the H and D basis spectra, which led to considerably more uncertainty in the calculated enrichment. Thus, we indicate the fit with a dashed line and can only say that ϕ_{HD} is greater than 0.9.

Protein	ϕ_{HD}
Y	0.70
F$_1$Y	0.56
Cl$_1$Y	0.59
F$_2$Y	0.54
Cl$_2$Y	>0.9

difference in the perturbed PES relative to water thus creating a thermodynamic preference for H. Theoretical considerations suggest that ϕ_{HD} should be minimized for LBHBs, and, indeed, the lowest experimental values (~0.3) have been measured in putative LBHB-containing complexes.[13,34]

The existence of an SIE was critical to the determination of ϕ_{HD} since it allowed us to accurately decompose the absorbance spectrum into a linear combination of H- and D-bound basis spectra. We observed a U-shaped trend for ϕ_{HD} with a stronger effect for those species with ΔpK_α closer to zero. This trend is consistent with that calculated from the best-fit model but not the LBHB results, which depend only weakly on ΔpK_α (Figure 7D). The calculation of ϕ_{HD} from our 1D model, described in SI S.5, is known to significantly underestimate the value of ϕ_{HD} because it neglects the bending degrees of freedom,[34] however the trends with respect to ΔpK_α should remain robust. The differences in magnitude between the data and best fit may be ascribed to this underestimation tendency. Kreevoy and Liang estimated that the true ϕ_{HD} for a LBHB should be greater than that derived from a 1D model by a factor of ~1.7,[34] which would put the model into closer quantitative agreement with the data.

Absence of Low-Barrier Hydrogen Bonds. In summary, this very simple model with only two global fit parameters makes surprisingly accurate predictions of trends in absorbance spectra, spectral isotope effects, and isotope fractionation
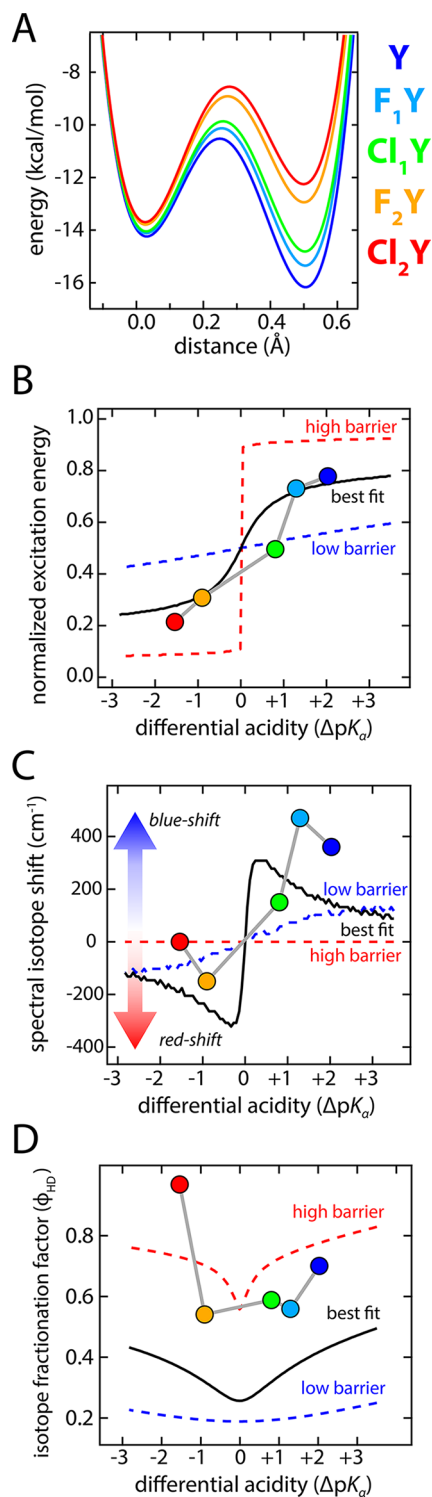
Figure 7. Results of modeling and comparisons of trends with experimental observations. The solid black curves in B, C, D, and E are calculated from the best fit parameters while the dashed red and blue curves correspond to high ($0.5\Delta_{DA}$) and low ($2\Delta_{DA}$) barrier predictions, respectively. We note that the calculated fractionation factors are certainly underestimates since they are derived from a 1D model and neglect the bending mode contributions.[34]

factors. The very fact that the model can successfully capture these properties appears to justify the most important assumption in which the substituted chromophore proteins were all modeled identically save the proton binding energy

offsets calculated from pK_a values for the chromophores in the denatured protein. The results strongly suggest against a LBHB interaction even though the prerequisite donor–acceptor distance and affinity matching conditions are met. This conclusion is drawn from several lines of evidence. First, the trends in the absorbance spectra with changing acidities conform most closely to those calculated from the best-fit coupling energy (Δ_{DA}) in detail of peak top and width (Figure 4F and Figure 7B). Likewise this Δ_{DA} also captures the trends in the spectral isotope effects and isotope fractionation factors which have poor correspondence to the predicted LBHB results (Figure 7C,D). Second, upon H to D exchange the SIE of the most closely affinity-matched species, Cl_1Y, shows a significant decrease in absorbance near the middle of the band (compare Figure 5C to Figure 4F solid and dashed green). This effect is anticipated by the best fit model but is absent for the LBHB case and emphasizes how D pulls away from the central barrier. Lastly, the magnitude of ϕ_{HD} at ~0.6 while suggesting a perturbed H-bond does not reach the low values, ~0.3, expected for a true LBHB (Figure 6).

The model presented above is entirely static. Another interpretation of the data is possible in which fluctuations in the protein and nearby solvent bias a low-barrier or barrierless PES toward one site or another even with perfect affinity matching on average. In this case one might view the model results presented as a time-averaged effective potential. Such a scenario, though beyond the scope of this paper, would also likely result in similar predictions for changes in spectra, spectral isotope effects, and isotope fractionation factors with respect to chromophore acidities and cannot be ruled out. This idea would be consistent with the many observations of Perrin et al. in which molecules forming symmetric H-bonds in crystallo are not symmetric in solution.[17,18] Furthermore, the immediate environment of the GFP chromophore is quite polar and contains numerous water molecules inside the β-barrel. A priori such surroundings would be expected to better stabilize a concentrated charge than the more diffuse charge associated with a LBHB.[11,12] Thus, the microscopic conditions for a largely symmetric LBHB may be only rarely encountered.

The lack of a LBHB in this poised model system may hold general implications for the existence and/or function of such interactions in proteins. The catalytic LBHB proposal posits that, in the course of an enzymatic reaction, the transition state transiently matches proton affinity with a protein H-bond partner and via a LBHB affords differential stabilization of 10–20 kcal/mol relative to the bound substrate.[8] Were such a large energetic preference present, one would expect LBHB character to dominate short protein H-bonds with affinity matched partners. Our inability to create such an interaction even while finely tuning the chromophore acidity suggests that there is no particular stabilization associated with a LBHB relative to a short ionic H-bond. The generality of this conclusion could be tested in the future in a number of proteins which share a common H-bonding motif—that is, a phenolate based chromophore engaged in a short H-bond. Among these are photoactive yellow protein (PYP) and ketosteroid isomerase (KSI). PYP has been shown by neutron diffraction crystallography to possess singly peaked deuteron density equidistant between Glu46 and the *p*-coumaric acid chromophore.[14] Moreover, it also features a significant SIE[35] and thus represents one of the best candidates for a genuine protein LBHB. KSI, despite sharing a similar interaction between Tyr16 and phenolate or naptholate transition state analogues, displays

electronic[36] and vibrational[37] spectroscopic data which quite clearly indicate well resolved protonation states suggestive of a double-welled potential. Systematic acidity perturbation of these systems and comparison to the results presented herein may help to clarify how the environmental context leads to diverse behavior among these otherwise geometrically similar situations.

CONCLUSIONS

We have successfully utilized methods of nonsense suppression and global incorporation to introduce a series of synthetic halide-substituted tyrosines into GFP that go on to make up the phenolic portion of the autocatalytically formed chromophore. Through inductive effects the substituents decreased the chromophore pK_a in the order $Y > F_1Y > Cl_1Y > F_2Y > Cl_2Y$ for a total span of 3.5 pK_a units. This enabled us to test the effect of proton affinity matching on a short protein H-bond by expressing these species in an S65T, H148D GFP mutant thought to harbor a LBHB between the chromophore and Asp148. X-ray crystal structures of the unmodified and the two chloro-substituted species revealed a largely conserved bond geometry with an O–O distance of ~2.45 Å, thus suggesting that the origins of the spectral shifts are largely isolated to the variation in chromophore acidities. The electronic absorbance spectra of the natured proteins were all found to titrate between a clearly deprotonated state at high pH and a highly unusual seemingly mixed state at low pH. This low pH mixed state contained the putative LBHB and gave rise to absorbance bands smoothly shifting between the protonated and deprotonated basis states with increasing acidity.

A 1D coupled Morse potential model provided a simple framework through which to interpret the experimental results and, with only two global fit parameters, was able to robustly model trends in the spectra, spectral isotope effects, and isotope fractionation factors. From the modeling results we draw two major conclusions. First, we predict that a modified chromophore having a solution pK_a of ~7.0 would, in the folded protein, be perfectly affinity matched to Asp148. Second, even under conditions of minimal differential acidity the predicted barrier to proton transfer exceeds the H-ZPE. The experimental observations most directly supporting this second claim are a marked decrease in intensity in the middle of the spectrum for near affinity matched sites upon H to D exchange, and higher isotope fractionation factors than expected for a LBHB. This evidence, however, cannot rule out an alternative model in which the dynamic solvation environment leads to bond asymmetry. In either case the fact remains that, relative to short ionic H-bonds, there appears to be no particular stabilization due to LBHBs and the accompanying proton delocalization across the bond. In fact, the absence of LBHBs suggests that the protein may actively avoid such a configuration. Distinguishing between these hypotheses will be the work of more sophisticated calculations. Regardless, our results suggest that even under seemingly ideal conditions for symmetric LBHB formation (i.e., O–O distances <2.5 Å and near perfect proton affinity matching) they are not observed.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The following file is available free of charge on the ACS Publications website at DOI: 10.1021/acscentsci.5b00160.

Additional information on the synthesis and incorporation of unnatural amino acids, X-ray crystallography statistics, the details of the 1D potential energy model, and ionization state dependent structural changes (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sboxer@stanford.edu. Phone: (650) 723-4482.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Shan, S.-o.; Herschlag, D. Hydrogen bonding in enzymatic catalysis: Analysis of energetic contributions. *Methods Enzymol.* **1999**, *308*, 246−276.

(2) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106* (8), 3210−3235.

(3) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* **1985**, *314* (6008), 235−238.

(4) Pollack, R. M. Enzymatic mechanisms for catalysis of enolization: ketosteroid isomerase. *Bioorg. Chem.* **2004**, *32* (5), 341−353.

(5) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. The Low Barrier Hydrogen Bond in Enzymatic Catalysis. *J. Biol. Chem.* **1998**, *273* (40), 25529−25532.

(6) Katz, B. A.; Spencer, J. R.; Elrod, K.; Luong, C.; Mackman, R. L.; Rice, M.; Sprengeler, P. A.; Allen, D.; Janc, J. Contribution of Multicentered Short Hydrogen Bond Arrays to Potency of Active Site-Directed Serine Protease Inhibitors. *J. Am. Chem. Soc.* **2002**, *124* (39), 11657−11668.

(7) Rajagopal, S.; Vishveshwara, S. Short hydrogen bonds in proteins. *FEBS J.* **2005**, *272* (8), 1819−1832.

(8) Cleland, W.; Kreevoy, M. Low-barrier hydrogen bonds and enzymic catalysis. *Science* **1994**, *264* (5167), 1887−1890.

(9) Frey, P.; Whitt, S.; Tobin, J. A low-barrier hydrogen bond in the catalytic triad of serine proteases. *Science* **1994**, *264* (5167), 1927−1930.

(10) Warshel, A.; Papazyan, A.; Kollman, P. On low-barrier hydrogen bonds and enzyme catalysis. *Science* **1995**, *269* (5220), 102−106.

(11) Warshel, A.; Papazyan, A. Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (24), 13665−13670.

(12) Schutz, C. N.; Warshel, A. The low barrier hydrogen bond (LBHB) proposal revisited: The case of the Asp ⋯ His pair in serine proteases. *Proteins: Struct., Funct., Bioinf.* **2004**, *55* (3), 711−723.

(13) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, P. A. Fractionation factors and activation energies for exchange of the low barrier hydrogen bonding proton in peptidyl trifluoromethyl ketone complexes of chymotrypsin. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (25), 14664−14668.

(14) Yamaguchi, S.; Kamikubo, H.; Kurihara, K.; Kuroki, R.; Niimura, N.; Shimizu, N.; Yamazaki, Y.; Kataoka, M. Low-barrier hydrogen bond in photoactive yellow protein. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (2), 440−444.

(15) Emsley, J. Very strong hydrogen bonding. *Chem. Soc. Rev.* **1980**, *9* (1), 91−124.

(16) Caminati, W.; Grabow, J.-U. The C2v Structure of Enolic Acetylacetone. *J. Am. Chem. Soc.* **2005**, *128* (3), 854−857.

(17) Perrin, C. L. Symmetries of Hydrogen Bonds in Solution. *Science* **1994**, *266* (5191), 1665−1668.

(18) Perrin, C. L.; Lau, J. S. Hydrogen-Bond Symmetry in Zwitterionic Phthalate Anions: Symmetry Breaking by Solvation. *J. Am. Chem. Soc.* **2006**, *128* (36), 11820−11824.

(19) Perrin, C. L. Are Short, Low-Barrier Hydrogen Bonds Unusually Strong? *Acc. Chem. Res.* **2010**, *43* (12), 1550−1557.

(20) Fried, S. D.; Boxer, S. G. Thermodynamic framework for identifying free energy inventories of enzyme catalytic cycles. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (30), 12271−12276.

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

(22) 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* (16), 8362−8367.

(23) Heim, R.; Cubitt, A. B.; Tsien, R. Y. Improved green fluorescence. *Nature* **1995**, *373* (6516), 663−664.

(24) Kneen, M.; Farinas, J.; Li, Y.; Verkman, A. S. Green fluorescent protein as a noninvasive intracellular pH indicator. *Biophys. J.* **1998**, *74* (3), 1591−1599.

(25) Shu, X.; Kallio, K.; Shi, X.; Abbyad, P.; Kanchanawong, P.; Childs, W.; Boxer, S. G.; Remington, S. J. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies. *Biochemistry* **2007**, *46* (43), 12005−12013.

(26) Shi, X.; Abbyad, P.; Shu, X.; Kallio, K.; Kanchanawong, P.; Childs, W.; Remington, S. J.; Boxer, S. G. Ultrafast Excited-State Dynamics in the Green Fluorescent Protein Variant S65T/H148D. 2. Unusual Photophysical Properties. *Biochemistry* **2007**, *46* (43), 12014−12025.

(27) Kondo, M.; Heisler, I. a.; Stoner-Ma, D.; Tonge, P. J.; Meech, S. R. Ultrafast dynamics of protein proton transfer on short hydrogen bond potential energy surfaces: S65T/H148D GFP. *J. Am. Chem. Soc.* **2010**, *132* (5), 1452−1453.

(28) Liu, C. C.; Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **2010**, *79* (1), 413−444.

(29) Shinobu, A.; Agmon, N. The Hole in the Barrel: Water Exchange at the GFP Chromophore. *J. Phys. Chem. B* **2015**, *119* (8), 3464−3478.

(30) Wang, Q.; Shui, B.; Kotlikoff, M. I.; Sondermann, H. Structural Basis for Calcium Sensing by GCaMP2. *Structure* **2008**, *16* (12), 1817−1827.

(31) Do, K.; Boxer, S. G. Thermodynamics, Kinetics, and Photochemistry of beta-Strand Association and Dissociation in a Split-GFP System. *J. Am. Chem. Soc.* **2011**, *133* (45), 18078−18081.

(32) McKenzie, R. H. A diabatic state model for donor-hydrogen vibrational frequency shifts in hydrogen bonded complexes. *Chem. Phys. Lett.* **2012**, *535*, 196−200.

(33) McKenzie, R. H.; Bekker, C.; Athokpam, B.; Ramesh, S. G. Effect of quantum nuclear motion on hydrogen bonding. *J. Chem. Phys.* **2014**, DOI: 10.1063/1.4873352.

(34) Kreevoy, M. M.; Liang, T. M. Structures and isotopic fractionation factors of complexes, A1HA2. *J. Am. Chem. Soc.* **1980**, *102* (10), 3315−3322.

(35) Kaledhonkar, S.; Hara, M.; Stalcup, T. P.; Xie, A.; Hoff, W. D. Strong ionic hydrogen bonding causes a spectral isotope effect in photoactive yellow protein. *Biophys. J.* **2013**, *105* (11), 2577−2585.

(36) Childs, W.; Boxer, S. G. Proton affinity of the oxyanion hole in the active site of ketosteroid isomerase. *Biochemistry* **2010**, *49* (12), 2725−2731.

(37) Fried, S. D.; Boxer, S. G. Evaluation of the Energetics of the Concerted Acid−Base Mechanism in Enzymatic Catalysis: The Case of Ketosteroid Isomerase. *J. Phys. Chem. B* **2011**, *116* (1), 690−697.

Supporting Information for:

"Short Hydrogen Bonds and Proton Delocalization in Green Fluorescent Protein (GFP)"

Luke M. Oltrogge and Steven G. Boxer
Department of Chemistry, Stanford University, Stanford, CA 94305-5012, United States

Contents

S.1 Protein Sequences

All proteins used in this study are based on a circular permutant of Superfolder GFP.[1] The native N- and C-termini are fused with the linker sequence, GGTGGS (highlighted in blue). These constructs were designed by us and synthetically prepared by Genscript. This protein is the same as the one called *ih*:GFP used in Refs. [2] and [3]. "*ih*" indicates that the interior helix bearing the chromophore is the N-terminal structural element.

The residue triad forming the chromophore on the central helix is highlighted in green, Asp148 (on β-strand 7) is highlighted in yellow, the original terminal linker is highlighted in cyan, and the hexa-histidine purification tag is highlighted in gray.

ih:GFP S65T, H148D

MGHHHHHHSSGGKLPVPWPTLVTTL TY^aG VQCFSRYGTRGSGSIEGRHSGSGSPDHMKR
HDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGH
KLEYNFNS D^b NVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPD
NHYLSTQTVLSKDPNEKRDHMVLLEFVTAAGITHGMDELY GGTGGS ASQGEELFTGVV
PILVELDGDVNGHKFSVRGEGEGDATIGKLTLKFISTT

a. or halogen substituted Tyr66

b. or His148

S.2 Experimental Methods

Synthesis of Halide-Substituted Tyrosines

Of the four halide-substituted tyrosines used in this study only the 3-chloro-L-tyrosine was commercially available (Sigma Aldrich, St. Louis). The 3,5-dichloro-L-tyrosine, 3-fluoro-L-tyrosine, and 3,5-difluoro-L-tyrosine were all prepared synthetically. 3,5-dichloro-L-tyrosine was synthesized by exhaustive chlorination of L-tyrosine while the fluoro-substituted tyrosines were prepared by enzymatic synthesis from fluoro-substituted phenols.

3,5-dichloro-L-tyrosine synthesis

A 50 mL round-bottom flask was charged with a stirbar and 1.0 g (5.5 mmol) of L-tyrosine (Sigma) suspended in 10 mL of glacial acetic acid. Under nitrogen, 2.0 mL (25 mmol) of sulfuryl chloride was added dropwise and allowed to stir at room temperature for 20 hours.[4] The suspended reaction product was vacuum filtered, washed with cold diethyl ether, and dried. The product was confirmed pure by ^{1}H-NMR and used without further purification.

3-fluoro- and 3,5-difluoro-L-tyrosine synthesis

The plasmid for *Citrobacter freundii* tyrosine phenol lyase (TPL) was kindly provided by Prof. Robert Phillips. The TPL was recombinantly expressed and crudely purified by an ammonium sulfate salt cut using previously described methods.[5] To a beaker containing one liter of 100 mM pyruvic acid, 200 mM ammonium acetate, pH 8.25 buffer were added 13.7 mg pyridoxal-5'-phosphate (the TPL cofactor), 0.35 mL of β-mercaptoethanol, ~20 units of purified TPL, and 10 mmol of either 2-fluoro-phenol or 2,6-difluoro-phenol (Sigma). The reaction was allowed to sit covered overnight at room temperature. At 24 hour intervals, 5 mmol of the appropriate fluoro-phenol was added up to a final total of 30 mmol. The reaction progress was monitored daily by analytical HPLC prior to additions. The reaction mixture was purified on a column of approximately 200 mL of Dowex-50 cation exchange resin (Sigma) equilibrated with 2 M HCl. After washing with deionized water, the product was eluted with 2 M NH$_4$OH. The eluate was lyophilized and later redissolved in ~10 mL of boiling water before recrystallizing at

4°C. The product was vacuum filtered, washed with ice cold water and ethanol, and dried under vacuum prior to analysis by ^{1}H-NMR.

*Global incorporation of 3-fluoro-L-tyrosine (**F_1Y**)*

No nonsense suppression system currently exists for 3-fluoro-L-tyrosine due to its chemical similarity to L-tyrosine causing cross reactivity with the endogenous tyrosyl-tRNA synthetase. Therefore the incorporation of 3-fluoro-L-tyrosine was performed globally[6, 7]; this introduces fluorine into the chromophore and the other 8 tyrosines, but these have little or no effect on the properties of GFP. BL21(DE3) *E. coli* containing the vector pET-15b with *ih*:GFP C48S, S65T, H148D were grown in one liter of M9 minimal media with 100 mg/L ampicillin at 37°C to a cell density of approximately 0.5 OD_{600}. To the media was added 1 mmol each of L-phenylalanine and L-tryptophan along with 1 mmol of 3-fluoro-L-tyrosine.[a] After 5 min., protein expression was induced with 10 mM isopropyl thiogalactopyranoside (IPTG). The cells were harvested after 4 hours, resuspended in lysis buffer (50 mM HEPES, 300 mM NaCl, pH 8.0), and homogenized. The cellular debris was removed by centrifugation and the supernatant purified by hexahistidine-tag nickel affinity chromatography and further by anion exchange chromatography.

*Nonsense suppression non-canonical amino acid incorporation (**Cl_1Y, Cl_2Y, and F_2Y**)*

In the *ih*:GFP S65T, H148D pET-15b vector Tyr66 was mutated to the amber stop codon (TAG) using a Stratagene Quikchange Lightning mutagenesis kit according to manufacturer protocols. BL21(DE3) *E. coli* were co-transfected with GFP pET-15b vector with ampicillin resistance and each of the required tRNA/aminoacyl tRNA synthetase pEVOL vectors bearing chloramphenicol resistance. The two mono- and dichloro tyrosine plasmids were provided by Prof. Jiangyun Wang (ClYRS and Cl2YRS from Liu *et al.*)[8] while the difluoro-tyrosine plasmid was provided by Prof. Joanne Stubbe (E3 from Minnihan *et al.*).[9]

---

[a] Phenylalanine and tryptophan were added in order to generally suppress aromatic amino acid metabolism to keep levels of unlabeled tyrosine down.

To one liter of rich LB growth media was added 0.5 mmol of the corresponding substituted tyrosine prior to autoclaving. 5-mL starter cultures grown overnight were transferred to 100 mL of media and allowed to grow for 2 hours at 37°C. This 100-mL culture was transferred into a flask containing the remaining media in addition to IPTG at 0.25 g/L (for GFP induction) and D-arabinose at 0.2 g/L (for tRNA/aminoacyl tRNA synthetase induction) at 25°C. The cells were grown overnight at 25°C and harvested. The protein purification protocol was the same as described above. The purity and identity of all proteins were confirmed with electrospray ionization mass spectrometry.

pH Titrations

All pH titrations were performed using master buffers of 30 mM citrate, 30 mM phosphate, and 200 mM NaCl with small amounts of NaOH added to adjust the pH. Titrations of denatured protein included 6 M guanidinium hydrochloride. In these solutions a correction of +0.72 was applied to the electrode pH reading according to Garcia-Mira and Sanchez-Ruiz.[10] Concentrated solutions of protein were diluted at least 30-fold into the corresponding buffer and allowed to equilibrate for several minutes prior to absorbance measurements.

Equilibrium isotope fractionation

Two buffers were prepared with identical hydron activity with a composition of 30 mM citrate, 30 mM phosphate, 200 mM NaCl, and pH or pD 5.0. The pD was determined using the standard pH electrode correction (*i.e.* $pD = pH_{app} + 0.42$).[11] These two buffers were mixed in the proper ratios to give solutions of 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 mole fraction deuterium (χ_D). Concentrated protein solutions were diluted more than 30-fold into each of these buffers to give a maximum optical density of about 0.2 at 1 cm pathlength. The absorbance spectra were measured using a Cary 6000i UV/Vis spectrometer.

Protein crystallization

The proteins **Y**, **Cl₁Y**, and **Cl₂Y** were all crystallized by hanging drop vapor diffusion. **Y** and **Cl₂Y** were crystallized under the same conditions with 50 mM sodium acetate, 100 mM NaCl, 5% (wt/vol) PEG 3550, pH 5.0. The conditions for **Cl₁Y** were 50 mM sodium acetate, 100 mM NaCl, 15% (w/vol) PEG 3550, pH 5.0. In all cases protein of concentration 10 mg/mL was mixed 1:1 with the mother liquor in the initial drop. All proteins gave initial crystals of rather poor quality which were improved considerably by microseeding. This was performed by crushing the low quality crystal by vortexing with a glass bead, serially diluting the resulting nanocrystals, and using a cat whisker to transfer the microseeds to another set of equilibrated hanging drops prepared with lower precipitant concentrations. The cryoprotectant was the mother liquor with sucrose added to 2 M.

X-ray data collection and structure refinement

The x-ray diffraction data for **Cl₁Y** was collected on BL 7-1 at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA) while the data for **Y** and **Cl₂Y** was acquired at BL 5.0.2 at the Advanced Light Source (Berkeley, CA).[b] All data was obtained at 100 K. The data were indexed and reduced with iMosflm[12] and scaled with CCP4.[13] Molecular replacement was carried out in PHENIX[14] using the original S65T, H148D structure (PDB entry: 2DUF) as the search model. Mutations and cycles of model building and refinement were performed with Coot[15] and PHENIX. The substituted chromophore structures and parameters were modified with the Phenix REEL program. To reduce model bias in the location of D148 the sidechain was removed and then modeled back in to the electron density following simulated annealing. The overall fold of the protein, despite being a circular permutant, was virtually identical to those of both Superfolder GFP (PDB ID: 2B3P) and the original S65T, H148D GFP (PDB ID: 2DUF).

^{1}H-NMR spectra of protein

1D ^{1}H-NMR spectra were collected for **Y**, **Cl₁Y**, and **Cl₂Y** using a Varian 800 MHz spectrometer with a cryoprobe. All samples were approximately 300 μM protein with 5% D_2O for locking. Water suppression was performed using the 1-3-3-1 binomial sequence[16] in order to minimize spin saturation for protons in chemical exchange.

S.3 Crystallographic Data

	Y	Cl$_1$Y	Cl$_2$Y
Data Collection			
Space group	P 1 2$_1$ 1	P 1 2$_1$ 1	P 1 2$_1$ 1
Cell dimensions			
a, b, c (Å)	46.37, 68.26, 57.87	51.84, 68.59, 60.88	46.14, 68.13, 58.00
α, β, γ (°)	90.0, 102.2, 90.0	90.0, 100.5, 90.0	90.0, 102.33, 90.0
Total Reflections	136,600	125,462	269,998
Unique Reflections	27,729	37,043	38,429
Resolution (Å)	37.8-1.9	36.3-1.8	56.7-1.7
R$_{merge}$	0.069	0.107	0.079
I/σI	11.0	6.8	9.8
Completeness (%)	99.3	98.6	99.3
Redundancy	4.9	3.4	7.0
Refinement			
Resolution (Å)	1.9	1.8	1.7
R$_{work}$ / R$_{free}$	0.22/0.26	0.20/0.25	0.20/0.23
No. molecules	2	2	2
No. non-H atoms			
protein	3344	3535	3458
ligand	44	46	48
solvent	26	99	58
B-factors			
protein (Å^3)	38.0	22.8	28.5
ligand (Å^3)	28.2	18.9	21.2
solvent (Å^3)	28.7	21.0	23.5
rmsd bond lengths (Å)	0.008	0.007	0.009
rmsd bond angles (°)	1.15	1.06	1.25

Table S.1

X-ray diffraction data collection and refinement statistics.

protein	chain	r_{O-O} (Å)	ϕ_P (°)	ϕ_I (°)
Y	A	2.41±0.22	12.8	16.8
	B	2.74±0.26	8.5	15.6
Cl$_1$Y	A	2.41±0.20	24.9	10.0
	B	2.52±0.20	18.6	7.9
Cl$_2$Y	A[c]	---	9.6	7.3
	B[d]	2.51±0.19	18.6	7.9
S65T, H148D pH 5.6 (PDB ID: 2DUF)[17]	A	2.32	7.2	16.8
S65T, H148D pH 10.0 (PDB ID: 2DUE)[17]	A	---	4.2	5.7
sfGFP[e] (PDB ID: 2B3P)[1]	A	---	0.4	1.4

Table S.2

Chromophore geometric parameters. Distance uncertainties were estimated using the method of Gurusaran *et al.*[18]

---

[c] This A-chain is deprotonated, not to be confused with the protonated form called the A-state. Asp148 is pointed outward to the solvent thus an r_{O-O} distance is not given. See section S.6 and Fig. S.5 for details.

[d] This B-chain has a protonated chromophore.

[e] This protein is deprotonated in this structure. Also no r_{O-O} distance is provided since this is His148.

S.4 Spectroscopic Data
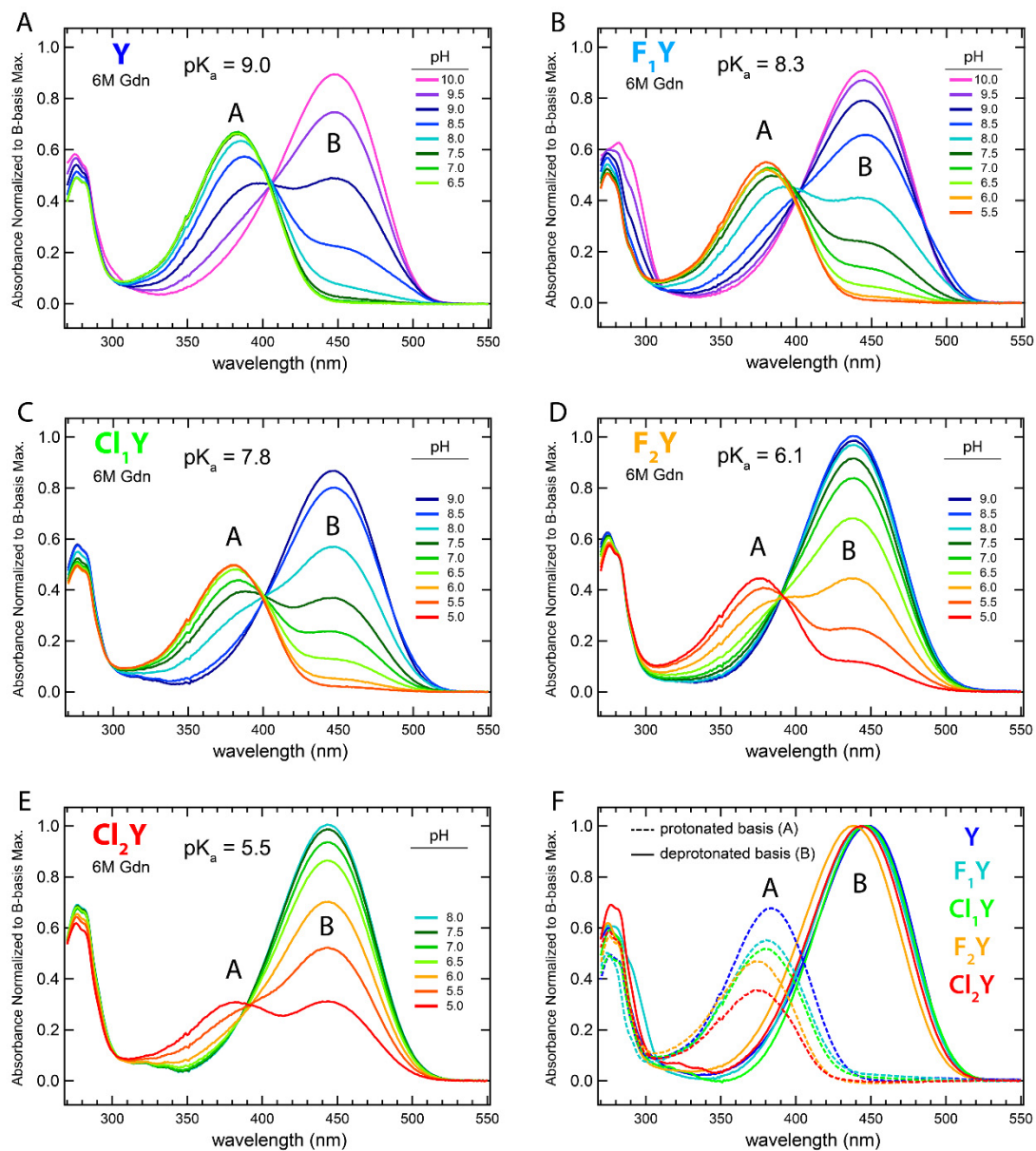
Figure S.1

A-E) UV/Vis pH titrations of substituted GFPs denatured in 6M guanidinium HCl. F) Comparison of denatured protonated and deprotonated basis spectra normalized to the deprotonated maximum absorbance.
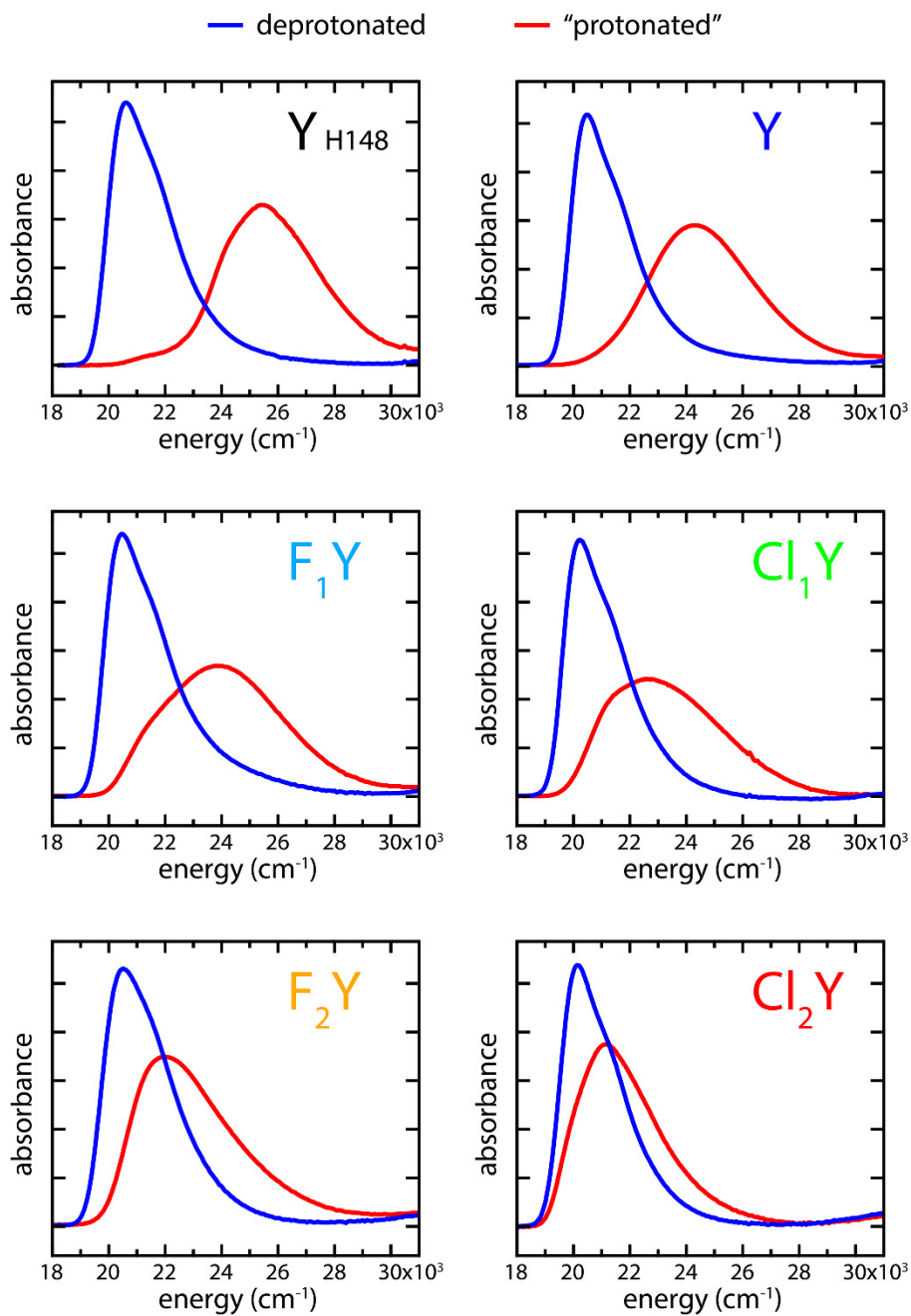
Figure S.2

Absorbance basis spectra for natured protein in low (red, "protonated") and high (blue, deprotonated) pH limits. "Protonated" is in quotes to indicate that rather than simply being protonated the chromophore is instead sharing the proton with Asp148. The extent to which the chromophore owns the proton depends on the relative pK_a.
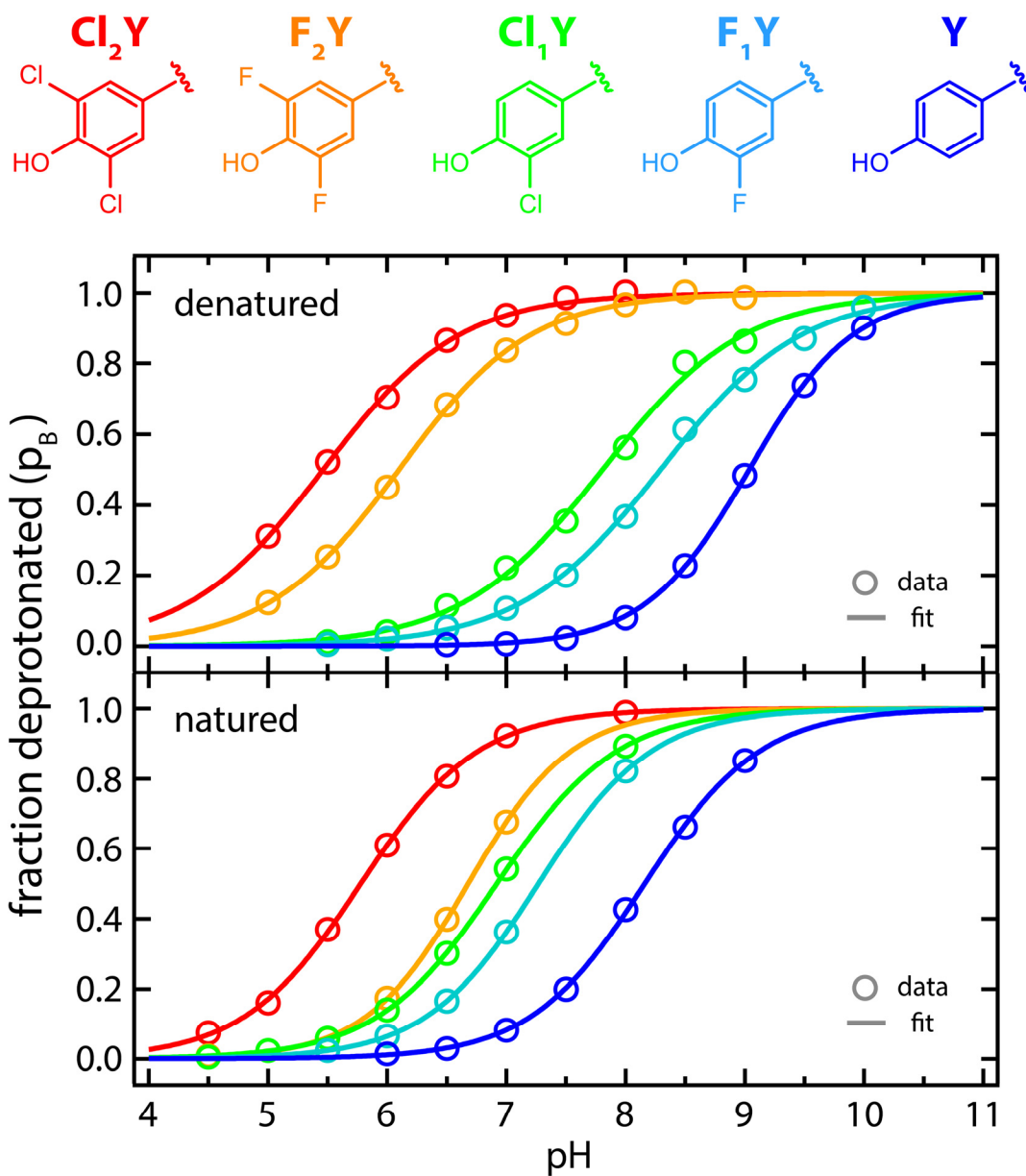
Figure S.3

pH titrations of substituted chromophores under denaturing and naturing conditions. It should be noted that the denatured titrations report directly on the acid dissociation of the substituted chromophore itself, and thus we refer to the resulting pK_a's as the *intrinsic pK_a's* (see Table S.3). In contrast, the natured titrations report on the acid dissociation of the proton shared by the chromophore and Asp148 and thus represents the collective behavior of these strongly coupled sites. This does not tell us anything about the relative affinity of the Asp148 and chromophore proton binding sites with respect to one another.

protein	denatured (energy in cm^{-1})				natured (energy in cm^{-1})			
	pK_a	A$_{max}$	B$_{max}$	ΔE	pK_a	A$_{max}$	B$_{max}$	ΔE
Y (H148)	9.0	26,040	22,320	3,720	5.9	25,450	20,620	4,830
Y	9.0	26,040	22,320	3,720	8.2	24,300	20,490	3,810
F$_1$Y	8.3	26,320	22,470	3,850	7.3	23,810	20,450	3,360
Cl$_1$Y	7.8	26,320	22,320	4,000	6.9	22,570	20,200	2,370
F$_2$Y	6.1	26,740	22,780	3,960	6.7	21,980	20,490	1,490
Cl$_2$Y	5.5	26,740	22,570	4,170	5.8	21,140	20,160	980

Table S.3

pK_a's , peak maxima, and energy differences between A- and B-states (ΔE) for denatured and natured substituted protein species.

S.5 Coupled Morse Potential Model

The basic framework for modeling the PES for the short hydrogen bonds and the effects of chromophore pK_a changes was adapted from that of Ross McKenzie.[19, 20] In this model the PES is generated by coupling two Morse potentials of the form,

Equation S.1
$$V(r) = D_e[\exp(-2a(r-r_0)) - 2\exp(-a(r-r_0))]$$

where D_e is the bond dissociation energy, r_0 is the equilibrium bond length, and a is a parameter controlling the well width. We use parameters characteristic of a typical O-H bond with D_e = 110 kcal/mol, r_0 = 0.96 Å, and a = 2.12 Å^{-1}. This yields a bond with a 3500 cm^{-1} stretch frequency. The coupling energy we use is that provided by McKenzie for a linear hydrogen bond (*i.e.* with the proton falling on the line between the donor and acceptor),

Equation S.2
$$\Delta_{DA}(R) = \Delta_1 \exp\left(-b\left(R - 2r_0 - \frac{1}{a}\right)\right)$$

where R is the distance between donor and acceptor atoms, r_0 and a are constants as given above for the O-H bond, and Δ_1 and b are fit parameters. In the present work we only consider R = 2.45 Å as determined from crystal structures of the S65T H148D variants so Δ_{DA} is itself just a constant throughout. McKenzie determined that Δ_1 = 0.4 D_e and b = a provided reasonable global parameters for hydrogen bonds of type -O··H-O-.[19] Thus the Hamiltonian may be written as,

Equation S.3
$$\hat{H} = \begin{pmatrix} V_{Asp148}(r) & \Delta_{DA} \\ \Delta_{DA} & V_{Cro}(R_{O-O} - r) \end{pmatrix}$$

By diagonalizing Equation S.3 the adiabatic potential energy surfaces are obtained from the eigenvalues with the one corresponding to the ground-state given by,

Equation S.4
$$E(r,R) = \tfrac{1}{2}\left(V_{Asp}(r) + V_{Cro}(R-r)\right) - \tfrac{1}{2}\left(\left(V_{Asp}(r) - V_{Cro}(r-R)\right)^2 + 4\Delta_{DA}(R)\right)$$

where $V_{Asp}(r)$ and $V_{Cro}(R-r)$ are the Morse potentials corresponding to the proton bound to Asp148 and the chromophore, respectively. In order to account for the pK_a differences in the

various substituted chromophores the dissociation energy (D_e) was adjusted by amounts equal to the differences acid dissociation free energies,

Equation S.5 $$\Delta\Delta G° = \ln 10 * RT \, \Delta pK_a$$

This makes the assumption that $\Delta\Delta G°$ is dominated by the enthalpy difference ($\Delta\Delta H°$) rather than any contributions from the entropy ($\Delta\Delta S°$) which is likely near zero for all substituted chromophores due to their structural similarity. The eigenvector associated with the ground-state eigenvalue describes the contribution of each of the diabats to the final PES. We make the approximation that the excitation energy is given by the linear combination of A- and B-state bases weighted according to the relative contribution due to the Asp148 and chromophore diabats. In this formulation a proton bound 100% to the chromophore would result in a pure A-state absorbance while a proton bound 100% to Asp148 would result in a pure B-state absorbance. Note that we specifically are not considering transitions between the adiabatic states obtained from diagonalization of Equation S.3 which would represent local excitation of the bond itself. Rather, the collective behavior of the entire chromophore must be taken into account. The simple model we use cannot encompass this complexity and so we take a semi-classical approach which recognizes the primacy of the protonated state in determining the spectrum and abstracts all other degrees of freedom into this coordinate. We use a normalized energy scale such that the excitation energy of the B-state is arbitrarily set to zero while the A-state is set to one. In this scale the excitation energy is given by,

Equation S.6

$$E_{ex}(r) \cong \frac{v_{Cro}^2}{v_{Asp}^2 + v_{Cro}^2} = \frac{1}{1 + \left(\dfrac{V_{Asp}(r) - V_{Cro}(R-r) - \sqrt{\left(V_{Asp}(r) - V_{Cro}(R-r)\right)^2 + 4\Delta_{DA}(R)}}{2\Delta_{DA}(R)} \right)^2}$$

where v_{Cro} and v_{Asp} are the eigenvector elements associated with the chromophore and Asp148 diabats respectively. This equation provides a mapping function with which to calculate the excitation energy as a function of the proton transfer coordinate (r).

Now with the functional form of the PES and the excitation energy the spectral prediction can be made with a knowledge of the proton (or deuteron) probability distribution. This was

calculated numerically using a finite difference method. Starting with the 1D time-independent Schrödinger equation,

Equation S.7
$$\hat{H}\psi(x) = \frac{-\hbar}{2m}\frac{\psi^2(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

with *V(x)* as the coupled Morse potential described above and *m* as the reduced mass for the asymmetric stretch mode of O-H-O or O-D-O, the expression was discretized with $x = n\Delta x$ which makes,

Equation S.8
$$\frac{\psi^2(x)}{dx^2} = \frac{[\psi((n+1)\Delta x) - \psi(n\Delta x)] - [\psi(n\Delta x) - \psi((n-1)\Delta x)]}{\Delta x^2}.$$

With *n* from 1 to *N* the nuclear Hamiltonian can be written as an *NxN* matrix,

Equation S.9

$$\hat{H} = \begin{pmatrix} \frac{\hbar}{m\Delta x^2} + V(\Delta x) & \frac{-\hbar}{2m\Delta x^2} & \cdots & 0 \\ \frac{-\hbar}{2m\Delta x^2} & \frac{\hbar}{m\Delta x^2} + V(2\Delta x) & \vdots & \vdots \\ \vdots & \cdots & \frac{\hbar}{m\Delta x^2} + V((N-1)\Delta x) & \frac{-\hbar}{2m\Delta x^2} \\ 0 & \cdots & \frac{-\hbar}{2m\Delta x^2} & \frac{\hbar}{m\Delta x^2} + V(N\Delta x) \end{pmatrix}.$$

The nuclear Hamiltonian matrix was numerically diagonalized in Matlab. The eigenvalues represent the vibrational energy levels and the associated eigenvectors are the corresponding wavefunctions. In order to obtain equilibrium populations, Maxwell-Boltzmann statistics were used to thermally average the probability density.

Equation S.10
$$p(x) = \frac{\sum_{i=1}^{N}|\psi_i(x)|^2 e^{-E_i/k_BT}}{\sum_{i=1}^{N} e^{-E_i/k_BT}}$$

E_i is the energy and $|\psi_i(x)|^2$ is the probability density of the *i*th vibrational eigenstate.

The probability density from Equation S.10 was then used in combination with Equation S.6 to produce a distribution of excitation energies (*p(E$_{ex}$)*). In order to generate a more realistic

spectrum with which to compare directly to experimental data, the calculated excitation energy distribution was convolved with a normalized Gaussian for the final output spectrum, $S(E_{ex})$.

Equation S.11
$$S(E_{ex}) = p(E_{ex}) \otimes \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-E_{ex}^2}{2\sigma^2}\right)$$

where σ specifies the peak width and was set to 0.289 in order approximate the experimental A- and B-state absorbance band widths.

In final form the model takes as global fit parameters the coupling energy (Δ_{DA}) and the proton binding energy offset whereby the set of substituted chromophore diabats could be collectively adjusted relative to the Asp148 diabat. These parameters together with the constants characteristic of the O-H bond (r_0, a, and D_e) fully specify the model potential energy surfaces from which the spectra are calculated. The workflow for the process is shown schematically in Figure S.4. It should be emphasized that for a given set of global parameters the diabats from the substituted chromophores are identical in every respect except for the small deviations in well depth due to the pK_a offset. The best fit global parameters were calculated using nonlinear least-squares optimization in Matlab. The fit parameters obtained thusly were then used to recalculate the vibrational eigenvalues, probability densities, and spectra with the deuteron reduced mass. The spectral isotope effects were obtained directly from these spectra. The isotope fractionation factor is defined as the equilibrium constant for the H/D exchange reaction with water.[21]

Equation S.12 ^{f} $\qquad$ ^{f} $\qquad\qquad$ $H\text{-}X + D(L_2O) \;\rightleftharpoons\; D\text{-}X + H(L_2O)$

Equation S.13 ^{f} $\qquad\qquad\qquad\qquad$ $\phi_{HD} = \frac{[D-X]}{[H-X]}\left(\frac{H}{D}\right)_{L_2O}$

We assume that the chemical potential difference for this reaction is given by the differences in zero-point energies for the O-H (O-D) stretch in the complex relative to solvent. From this the isotope fractionation factor may be calculated directly.

---

^{f} L_2O is being used as a shorthand for the particular solvent composition which may contain a mixture of H and D. For example the appearance of $D(L_2O)$ in Equation S.12 indicates that the deuteron in the reaction is obtained from the solvent.

Equation S.14 $$\phi_{HD} = \exp\left(\frac{-1}{k_B}\left(ZPE_{DX} - ZPE_{HX} + ZPE_{H(L_2O)} - ZPE_{D(L_2O)}\right)\right)$$

This treatment is most certainly an overestimation of the magnitude of the effect (*i.e.* the calculated ϕ_{HD} is too small) because it neglects all other degrees of freedom which experience considerably less perturbation. Kreevoy and Liang estimate that in LBHBs the true value for ϕ_{HD} should be ~1.7 times higher.[21] This correction factor puts our observed ϕ_{HD} in near quantitative agreement with the model. Regardless, the primary motivation was to compare trends rather than absolute magnitudes.
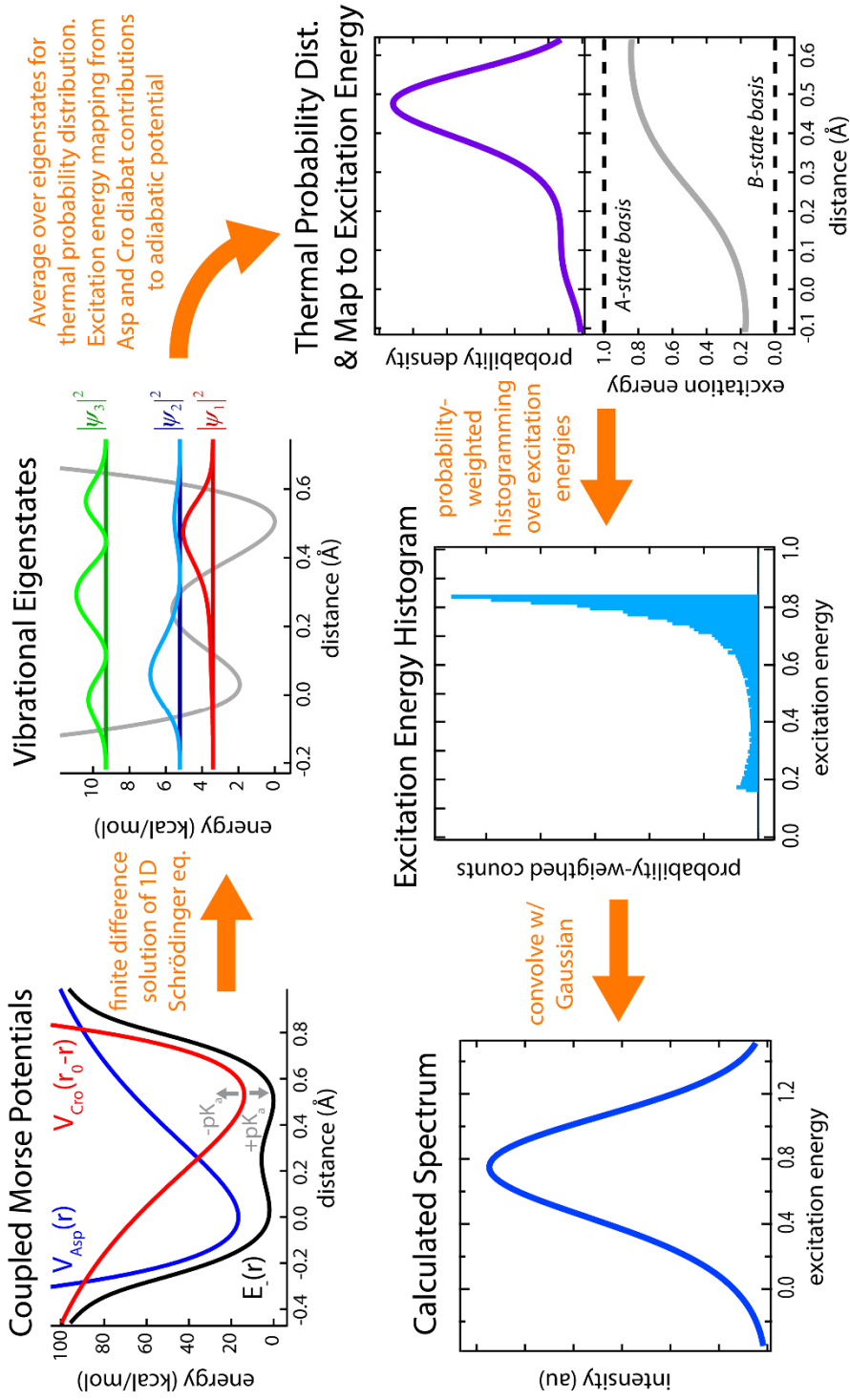
Figure S.4

Workflow for calculating trial absorbance spectra from input coupled Morse potentials.

S.6 Structural Changes Dependent on Ionization State

While this study is concerned primarily with the specific interaction between the chromophore and Asp148 it has also provided new structural information that bears on the issues of ground-state proton transfer and attendant conformational change which will be of interest to those more concerned with properties of GFP itself. In the course of structural characterization of these GFPs, two forms of **Cl₂Y** differing by chromophore ionization state were found in the same crystal. Analysis of the differences between these structures brings some insight into the structural changes accompanying proton transfer and also a plausible mechanistic explanation for the very slow (> 10 ms) chemical exchange rates between these states.

UV/Vis absorbance spectra from pH titrations of natured protein of all S65T, H148D mutants (including **Y**, **F₁Y**, **Cl₁Y**, **F₂Y**, **Cl₂Y**, and S65T, H148D from Shu *et al.*[17]) have found clean isosbestic points between a deprotonated state at high pH with absorbance peaked around 490 nm and another peak at low pH with a variable absorbance maximum dependent on the chromophore pK_a (see Figure S.2 and Table S.3). As discussed above and in the text this state is hypothesized to be a result of proton delocalization across a short hydrogen bond. The behavior of this state is *not* pH-dependent, however, the transitions between this state and the deprotonated state are pH-dependent.

The proximity of Asp148 to the chromophore presents a structural difficulty in the deprotonated state since it would entail the energetically prohibitive formation of two anions within 2.5 Å of one another. It is reasonable therefore to suspect that the chromophore deprotonation is achieved either by some pH-dependent modification of the environment which increases the proton affinity of Asp148 and thus allows complete proton transfer, or that there is a significant structural change that increases the distance to accommodate the two anions. Our data clearly support this latter hypothesis. The evidence for this claim is two-fold with NMR spectra providing evidence for slow inferred dynamics and x-ray crystallography providing an atomistic perspective on the mechanism.

In previous work we utilized an NMR method to follow proton transfer in GFP by introducing a ^{13}C label with high sensitivity to the chromophore ionization state.[3] The resulting

1D ^{13}C-NMR spectra contain information on both the thermodynamics and kinetics of this process. We found that β-strand 7 (residue 148 is in the middle of strand 7) plays an important role in setting the timescale of proton transfer to and from the solution and that this process can be quite slow (>10 ms) thus suggesting large-scale structural motion. We employed the same methodology to **Y** and found very slow chemical exchange between protonation states (Figure S.5). However, in addition, we have been able to complement this data with structural information on the closely related **Cl₂Y** protein.
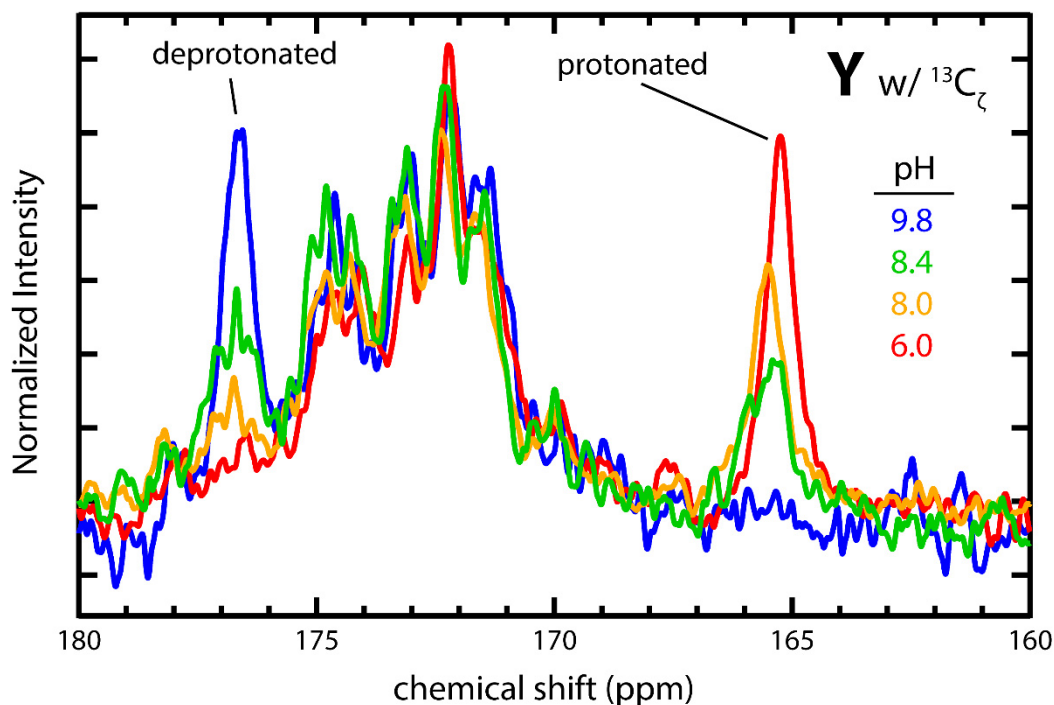
Figure S.5

1D ^{13}C-NMR of **Y** for a pH titration series. The presence of two very clear peaks implies that the chemical exchange rate is far less than the inverse frequency difference between the two basis states.

$\tau \gg \dfrac{1}{\left| v_B - v_A \right|}$. These spectra were collected on a 300 MHz NMR (75 MHz for ^{13}C) such that the

relevant reference timescale is 11.6 ms. The broad cluster of peaks between 170 and 175 ppm is due to backbone carbonyls at natural abundance.

Shu *et al.* solved crystal structures for GFP S65T, H148D in conditions of pH 5.6 (PDB ID: 2DUF) and pH 10.0 (PDB ID: 2DUE) for pure populations of the two observed optical

states.[17] The pH 5.6 structure was the one first demonstrating the close interaction between the chromophore and Asp148. The pH 10.0 structure, however, had poorly ordered electron density in the region of strand 7 containing Asp148 and could not be modeled. This was the first evidence that significant structural changes might be accompanying changes of ionization state.

In the present work all proteins were crystallized at pH 5.0 in the spacegroup P 1 2_1 1 and had a unit cell comprised of two GFP protein chains. For **Y** and **Cl$_1$Y** chains A and B were essentially identical. However, the crystal of **Cl$_2$Y** was fortuitously such that chain A had a deprotonated chromophore and chain B had a protonated chromophore (or more accurately a proton shared between the chromophore and Asp148). The measured solution pK_a for natured **Cl$_2$Y** was 5.8. At pH 5.0 one would expect there to be a relatively significant fraction of the deprotonated form of the chromophore, however due to increased protein destabilization and denaturation at low pH we were unable to push to lower pH for a more homogenous population. Ordinarily this scenario would be expected to lead to static disorder in the crystal. In this case however, presumably aided by the very slow chemical exchange process deduced from NMR, the crystal formed in such a way that alternating protein chains were in the two different conformations enabling simultaneous solution of both structures.

The resolution of the **Cl$_2$Y** structure at 1.7 Å is not sufficient to directly observe hydrogen atoms, therefore the positions of protons must be indirectly inferred. A number of factors informed the assignment. First, as mentioned above, the proximity of anions that a deprotonated assignment to chain B would necessitate creates an energetically unrealizable situation such that we may be confident this chain has a proton shared between these groups. Conversely, the conformational change in chain A which relocates Asp148 to the exterior of the protein would alleviate this destabilization and allow both positions to carry a negative charge. Secondly, the putative deprotonated chromophore (chain A) has three potential hydrogen bonding partners within 2.9 Å (Ser147, Thr203, and an ordered water) while the putative protonated chromophore (chain B) has just two (Asp148 and Thr203). This situation is consistent with that from other studies on the structural changes upon chromophore titration.[22] Lastly, chain A has a chromophore structure which is much more planar than the chain B chromophore. This is also an observation in accord with previous experimental measurements[10, 14] and theoretical calculations concerning the chromophore deformability.[17, 22, 23]
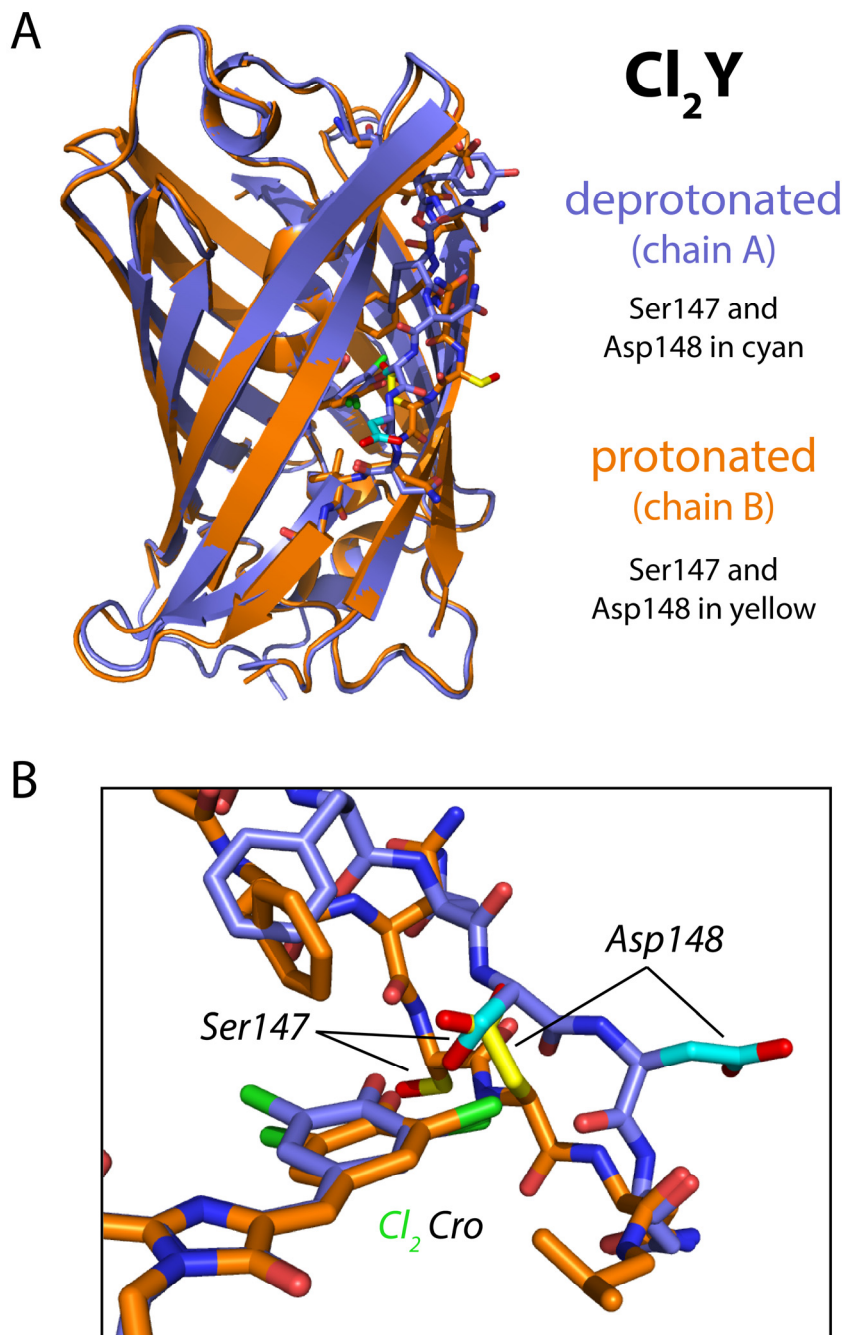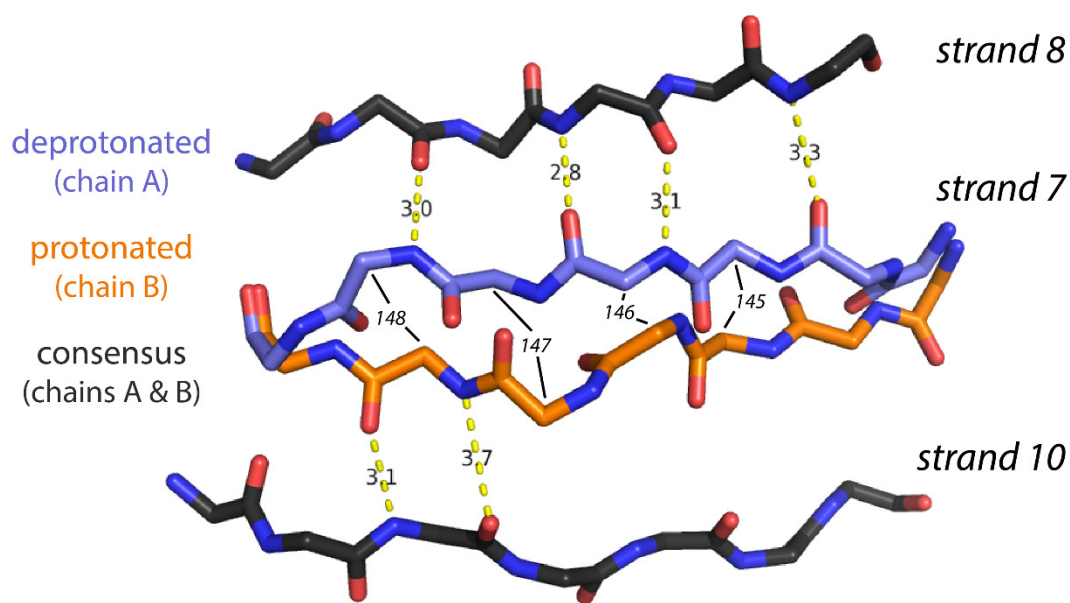
Figure S.6

A) Cartoon structure of aligned structures of chain A (deprotonated) and chain B (protonated) of **Cl₂Y** with the residues of greatest deviation shown in stick form. B) A close-up of the dichloro-chromophore and the positional changes of the key residues Ser147 and Asp148 on strand 7.

Alignment of the two **Cl₂Y** chains in PyMOL (Schrödinger, LLC) reveals that the structural transition linking the two conformations is largely localized to β-strand 7 (see Figure S.6). The two chains were essentially superimposed elsewhere having a backbone RMSD of 0.238 Å excluding strand 7. The significant deviations begin at position 143 and continue on to 149 with C_α distances reaching a maximum difference of 4.4 Å at Ser147. The most important functional difference between the two conformations is the swap of residues interacting with chromophore. In the protonated structure (chain B) Asp148 is participating in the short hydrogen bond with the chromophore while Ser147 is directed to the protein exterior. Upon deprotonation (chain A) the Asp148 sidechain has a net motion of 8.6 Å for the interacting oxygen which positions it on the exterior while the Ser147 hydroxyl group has a net motion of 6.0 Å as it establishes a hydrogen bond to the chromophore phenolate (see Figure S.6B).

Figure S.7

Backbone representation of the structural changes occurring between protonated (chain B) and deprotonated (chain A) structures. The regions of consensus structure are superimposable between the two chains. Backbone hydrogen bonds are indicated in yellow with the N-O distances given in Ångstroms. Residue C_α positions on strand 7 are indicated.

Uniquely among the 11 β-strands comprising the GFP structure, strand 7 possesses a conventional anti-parallel β-sheet for only about half the length of barrel. Beyond this point the backbone is twisted and, instead of interacting with the adjacent strands, satisfies backbone amide and carbonyl hydrogen bonds with sidechains, ordered water, or bulk solvent. The bordering β-strands, 8 and 10, are actually separated by a distance which makes it impossible for strand 7 to simultaneously engage in backbone hydrogen bonding with both. These factors have been implicated as giving rise to dynamic structural heterogeneity of strand 7 in solution computationally[24, 25] and also through NMR experiments[26] including H/D exchange.[27] It is interesting therefore to examine the structural change accompanying deprotonation of **Cl₂Y** in this context. The protonated form (chain B) has very close structural homology to Superfolder GFP (PDB ID: 2B3P) and S65T, H148D GFP (PDB ID: 2DUF) and the backbone conformation is essentially superimposable on both of these structures. The deprotonated **Cl₂Y** structure (chain A), however, is to our knowledge unique. The exclusion of Asp148 from the protein interior requires considerable changes in the backbone geometry. Unexpectedly this accommodation resulted in the establishment of four additional backbone hydrogen bonds to strand 8 relative to the common strand 7 conformation which makes none of these contacts but rather shares two backbone hydrogen bonds to strand 10 (Figure S.7).

The preceding discussion has been focused primarily on H148D. We hypothesize, however, that structural transitions involving this toggling between strands 8 and 10 may be a general phenomenon in fluorescent proteins. Beyond the role of gating proton transfer between the solution and chromophore this structural mechanism could be important to the large class of GFP-based sensors which incorporate sensing domains inserted within strand 7.

S.7 ^{1}H-NMR in search of LBHBs

One of the most common experimental observables used to identify LBHB's in proteins is the presence of ^{1}H-NMR peaks occurring abnormally far downfield (~12 – 20 ppm).[28-30] This method was used on **Y** and **Cl$_1$Y** and no diagnostic downfield proton resonances were observed. [g] Using the empirical correlation between O-O distance and proton chemical shift from Berglund and Vaughan we would have expected peaks around 18 ppm for the measured distance of ~2.45 Å.[28] A possible explanation for the missing peaks is that there is rapid proton exchange with water on the NMR timescale. In this event the peak of interest would disappear into the solvent peak provided that chemical exchange were taking place in approximately a millisecond or less. This result may be related to the observation of Stoner-Ma *et al.*[31] that no vibrational signature of neutral Asp148 could be detected following ultrafast ESPT in visible pump/ IR probe experiments on GFP S65T, H148D whereas such a signal was clearly seen for Glu222 in wtGFP.[32] Perhaps the proton is efficiently shuttled through an alternative relay which may also be operative for proton exchange in the ground state. Alternatively, it is possible that, on account of a smaller than expected coupling energy, the O-H distance is closer to its normal length and thus results in a peak that is further upfield and obscured by the protein background. **Cl$_1$Y** is most closely affinity matched and therefore should be expected to have the furthest downfield proton resonance. However, our model for **Cl$_1$Y** finds an O-H bond lengthening of about 0.07 Å relative to r_0 [h] which is consistent with empirical findings of Berglund and Vaughan and should still give rise to a downfield peak[28] so we favor the hypothesis of rapid exchange to solvent.

---

[g] ^{1}H-NMR was also performed on GFP S65T, H148D and likewise no unusual downfield proton resonance was observed.[17]

[h] These length change was calculated by using a value for the bond length determined by the maximum in the proton probability density.

S.8 References

(1) Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein, *Nat. Biotechnol. 24*, 79-88.

(2) Kent, K. P., Oltrogge, L. M., and Boxer, S. G. (2009) Synthetic Control of Green Fluorescent Protein, *J. Am. Chem. Soc. 131*, 15988-15989.

(3) Oltrogge, L. M., Wang, Q., and Boxer, S. G. (2014) Ground-State Proton Transfer Kinetics in Green Fluorescent Protein, *Biochemistry 53*, 5947-5957.

(4) McCubbin, J. A., Maddess, M. L., and Lautens, M. (2008) Enzymatic Resolution of Chlorohydrins for the Synthesis of Enantiomerically Enriched 2-Vinyloxiranes, *Synlett 2008*, 289-293.

(5) Chen, H., Gollnick, P., and Phillips, R. S. (1995) Site-Directed Mutagenesis of His343→Ala in Citrobacter freundii Tyrosine Phenol-Lyase, *Eur. J. Biochem. 229*, 540-549.

(6) Brooks, B., and Benisek, W. F. (1994) Mechanism of the Reaction Catalyzed by .DELTA.5-3-Ketosteroid Isomerase of Comamonas (Pseudomonas) testosteroni: Kinetic Properties of a Modified Enzyme in Which Tyrosine 14 Is Replaced by 3-Fluorotyrosine, *Biochemistry 33*, 2682-2687.

(7) Brooks, B., Phillips, R. S., and Benisek, W. F. (1998) High-Efficiency Incorporation in Vivo of Tyrosine Analogues with Altered Hydroxyl Acidity in Place of the Catalytic Tyrosine-14 of Δ5-3-Ketosteroid Isomerase of Comamonas (Pseudomonas) testosteroni: Effects of the Modifications on Isomerase Kinetics, *Biochemistry 37*, 9738-9742.

(8) Liu, X., Jiang, L., Li, J., Wang, L., Yu, Y., Zhou, Q., Lv, X., Gong, W., Lu, Y., and Wang, J. (2014) Significant Expansion of Fluorescent Protein Sensing Ability through the Genetic Incorporation of Superior Photo-Induced Electron-Transfer Quenchers, *J. Am. Chem. Soc. 136*, 13094-13097.

(9) Minnihan, E. C., Young, D. D., Schultz, P. G., and Stubbe, J. (2011) Incorporation of Fluorotyrosines into Ribonucleotide Reductase Using an Evolved, Polyspecific Aminoacyl-tRNA Synthetase, *J. Am. Chem. Soc. 133*, 15942-15945.

(10) Garcia-Mira, M. M., and Sanchez-Ruiz, J. M. (2001) pH corrections and protein ionization in water/guanidinium chloride, *Biophys. J. 81*, 3489-3502.

(11) Glasoe, P. K., and Long, F. A. (1960) Use of glass electrodes to measure acidities in deuterium oxide, *The Journal of Physical Chemistry 64*, 188-190.

(12) Leslie, A. G. W., and Powell, H. R. (2007) Processing Diffraction Data with Mosflm, In *Evolving Methods for Macromolecular Crystallography*, pp 41-51.

(13) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallographica Section D 67*, 235-242.

(14) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D 66*, 213-221.

(15) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D 66*, 486-501.

(16) Hors, P. J. (1983) A new method for water suppression in the proton NMR spectra of aqueous solutions, *Journal of Magnetic Resonance (1969) 54*, 539-542.

(17) Shu, X., Kallio, K., Shi, X., Abbyad, P., Kanchanawong, P., Childs, W., Boxer, S. G., and Remington, S. J. (2007) Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies, *Biochemistry 46*, 12005-12013.

(18) Gurusaran, M., Shankar, M., Nagarajan, R., Helliwell, J. R., and Sekar, K. (2014) Do we see what we should see? Describing non-covalent interactions in protein structures including precision, *IUCrJ 1*, 74-81.

(19) McKenzie, R. H. (2012) A diabatic state model for donor-hydrogen vibrational frequency shifts in hydrogen bonded complexes, *Chem. Phys. Lett. 535*, 196-200.

(20) McKenzie, R. H., Bekker, C., Athokpam, B., and Ramesh, S. G. (2014) Effect of quantum nuclear motion on hydrogen bonding, *J. Chem. Phys. 140*, -.

(21) Kreevoy, M. M., and Liang, T. M. (1980) Structures and isotopic fractionation factors of complexes, A1HA2, *J. Am. Chem. Soc. 102*, 3315-3322.

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

(23) Olsen, S., Lamothe, K., and Martínez, T. J. (2010) Protonic Gating of Excited-State Twisting and Charge Localization in GFP Chromophores: A Mechanistic Hypothesis for Reversible Photoswitching, *J. Am. Chem. Soc. 132*, 1192-1193.

(24) Shinobu, A., and Agmon, N. (2015) The Hole in the Barrel: Water Exchange at the GFP Chromophore, *J. Phys. Chem. B 119*, 3464-3478.

(25) Helms, V., Straatsma, T. P., and McCammon, J. A. (1999) Internal Dynamics of Green Fluorescent Protein, *J. Phys. Chem. B 103*, 3263-3269.

(26) Seifert, M. H. J., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B., and Holak, T. A. (2003) Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution, *Biochemistry 42*, 2500-2512.

(27) Huang, J.-r., Craggs, T. D., Christodoulou, J., and Jackson, S. E. (2007) Stable Intermediate States and High Energy Barriers in the Unfolding of GFP, *J. Mol. Biol. 370*, 356-371.

(28) Berglund, B., and Vaughan, R. W. (1980) Correlations between proton chemical shift tensors, deuterium quadrupole couplings, and bond distances for hydrogen bonds in solids, *J. Chem. Phys. 73*, 2037-2043.

(29) Jeffrey, G. A., and Yeon, Y. (1986) The correlation between hydrogen-bond lengths and proton chemical shifts in crystals, *Acta Crystallographica Section B 42*, 410-413.

(30) Frey, P., Whitt, S., and Tobin, J. (1994) A low-barrier hydrogen bond in the catalytic triad of serine proteases, *Science 264*, 1927-1930.

(31) Stoner-Ma, D., Jaye, A. a., Ronayne, K. L., Nappa, J., Meech, S. R., and Tonge, P. J. (2008) An alternate proton acceptor for excited-state proton transfer in green fluorescent protein: rewiring GFP, *J. Am. Chem. Soc. 130*, 1227-1235.

(32) Stoner-Ma, D., Jaye, A. A., Matousek, P., Towrie, M., Meech, S. R., and Tonge, P. J. (2005) Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy, *J Am Chem Soc 127*, 2864-2865.